

РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
ИНСТИТУТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ
И МАТЕМАТИЧЕСКОЙ ГЕОФИЗИКИ

Валерий Павлович Ильин

Математическое моделирование. Часть I
Непрерывные и дискретные модели

НОВОСИБИРСК
2017

РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
ИНСТИТУТ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ
И МАТЕМАТИЧЕСКОЙ ГЕОФИЗИКИ

Валерий Павлович Ильин

Математическое моделирование. Часть I
Непрерывные и дискретные модели

НОВОСИБИРСК
2017

ПРЕДИСЛОВИЕ

Математическое моделирование — это относительно новое научное направление. Хотя новое — это хорошо забытое старое, но в нашем веке естественно рассматривать только компьютерное моделирование, про которое вполне определенно можно сказать, что эта научная дисциплина уже существует официально. Именно под таким названием фигурирует одна из программ фундаментальных исследований РАН. Математическое моделирование входит в список актуальных направлений научно-технологического развития Российской Федерации. И наконец, в номенклатуре Высшей Аттестационной Комиссии имеется специальность 05.13.18. “Математическое моделирование, численные методы и комплексы программ”, содержанием которой, согласно приведенной официальной формуле, “является разработка фундаментальных основ и применение математического моделирования, численных методов и комплексов программ для решения научных и технических, фундаментальных и прикладных проблем”.

Можно предложить следующее научное определение: математическое моделирование — это совокупность фундаментальных и прикладных научных направлений по изучению процессов и явлений с помощью математических методов, вычислительных алгоритмов и информационных технологий, реализуемых в программных комплексах, с помощью которых проводится машинный эксперимент и принимается решение по результатам исследований.

Данная краткая формулировка является очень ёмкой и требует развернутых комментариев. Прежде всего, рассматриваемые процессы и явления охватывают самые разные сферы окружающего мира и человеческой деятельности, а их сущности представляются физическими, химическими, биологическими и другими предметными моделями. Общим является лишь то, что все они описываются с помощью математических моделей на языке дифференциальных, интегральных и/или дискретных уравнений. Изучение этих объектов обязательно начинается средствами теоретической математики, которыми устанавливаются необходимые вопросы существования, единственности, коррект-

ности и основные свойства решений исследуемых задач, которые в реальных проблемах являются междисциплинарными и описываются как прямыми, так и обратными постановками. Для сформулированных задач конструируются, обосновываются и оптимизируются численные алгоритмы, которые на современном уровне обеспечивают высокое разрешение прогностического моделирования с гарантированной точностью результатов, но требуют при этом экстремальных вычислительных ресурсов. Построенные численные модели затем реализуются в прикладных программных комплексах на современных многопроцессорных вычислительных системах с распределенной и иерархической общей памятью, при квалифицированном использовании существующих операционных систем и огромного объема программных инструментов и приложений. Резюмировать такую квалифицированную характеристику можно следующей формулой: “Математическое моделирование = модели + алгоритмы + прикладные программы + компьютерный эксперимент”. Особенности именно текущего момента в этом достаточно старом понятии являются три аспекта: актуальные модели представляют собой сверхзадачи, подступиться к которым было невыносимым еще пару десятилетий назад, алгоритмы являются принципиально параллельными, без чего невозможно достичь результата за астрономически короткое время, программное обеспечение и расчеты реализуются в рамках гетерогенных кластерных систем на основе облачных вычислений и сетевых (grid) технологий.

Математическое моделирование призвано стать третьим орудием познания, наравне с теоретическими и экспериментальными исследованиями, а также явиться реальной производственной силой и средством качественного повышения производительности труда в эпоху реиндустриализации и нового экономического уклада.

Обеспечение высокопроизводительного вычислительного оборудования и профессионального программного обеспечения — это прерогатива промышленных технологий, однако методология моделирования составляет именно математическое научное направление.

Настоящая книга представляет первую часть дисциплины “Математическое моделирование”, посвященную непрерывным и дискретным моделям широкого класса задач. Введение дает об-

щее представление о предметной области, которой посвящена книга. Первые две главы посвящены классическим и обобщенным постановкам начально-краевых задач. Следующие три главы содержат последовательно описания методов конечных разностей, конечных объемов и конечных элементов, являющихся основными универсальными подходами к построению численных решений. Изложение в данном случае акцентировано на аппроксимационных аспектах, а вопросы алгебраических свойств и технологических способов реализации вычислительных схем отнесены к последующим разделам. Эффективность всех рассматриваемых алгоритмов в значительной степени зависит от способа дискретизации исходной задачи и от качества получаемой сетки. Именно этой проблеме посвящена последняя глава книги.

Содержание книги охватывает в целом очень широкий круг вопросов, многие из которых заслуживают отдельных монографий. Принятая в книге концепция изложения материала исходит из положения Козьмы Пруткова, что “никто и обнимет необъятное”, а требовать от читателя стать энциклопедистом нереально в наш век глубокой профессионализации. Специалиста по математическому моделированию можно сравнить с многоборцем по спортивным дисциплинам, успех которого складывается из суммы показателей в различных видах соревнований. Из этой аналогии следует, что “математик-моделлер” (от появившегося относительно недавно английского слова *modeller*) должен иметь широкую область компетенций. Это не отрицает его более глубокий профессионализм в отдельных направлениях, но предполагает наличие определенного необходимого минимума по всем сопричастным научным областям.

Исходя из данной парадигмы, подача математического материала ведется на понятийном, или описательном, а отчасти справочном уровне. В различных главах даются все основные определения и утверждения, но теоремы приводятся без доказательств. Читателю рекомендуется проявить общесистемный подход к усвоению достаточно разнородных знаний, устанавливая внутренние связи и закономерности в той многообразной картине мира, которую представляет рассматриваемая нами предметная область.

Значительная часть рассматриваемых в книге вопросов относится к прерогативе научных направлений, в целом называе-

мых Computer Science, а формой общения здесь объективно стал английский язык, который терминологически активно проникает и в русскую литературу. Борьба с этой исторической интернациональной тенденцией бесполезно и даже вредно. Поэтому мы зачастую приводим общеупотребительные англоязычные термины, которые, кстати говоря, помогают находить необходимую информацию в Интернете.

В списке литературы мы приводим только книги, в том числе на английском языке, которые еще не переведены на русский. Тексты большинства из них можно найти с помощью интернетовских поисковиков. Ссылки на журнальные статьи и научно-технические отчеты, издаваемые большинством зарубежных университетов, мы не приводим (за некоторыми исключениями в отдельных параграфах), так как их изучение привнесло бы читателю слишком тяжелую нагрузку.

Приведенный библиографический материал по тематике очень обширен и разнороден, как и сам изучаемый предмет. Здесь есть книги выдающихся математиков В.И. Арнольда [6] и Ю.И. Манина [119] с глубокими размышлениями о роли абстрактной математики в познании реального мира. Большое количество монографий посвящено проблемам прикладной математики, связанным с конкретными приложениями и различными физико-математическими моделями. Перечень затрагиваемых направлений можно представить следующим образом: процессы тепло-массопереноса, гидро-газодинамики и фильтрации многофазных сред, проблемы упругости и вязкопластичности, электромагнетизм и биофизика, окружающая среда, природопользование и экология, экономика и многообразные дискретные системы, обратные задачи и оптимизация.

Мы приводим в списке литературы также книги по численным методам математического моделирования, преимущественно по конечно-разностным, конечно-объемным и конечно-элементным алгоритмам аппроксимации уравнений в частных производных. При этом неизбежно затрагиваются и другие направления вычислительной математики, такие как интегрирование систем обыкновенных дифференциальных уравнений, методы граничных интегральных уравнений, методы частиц в ячейках, алгоритмы локальной и глобальной минимизации функционалов. Конечно, все эти вопросы невозможно осветить должным образом в пределах одного тома.

Библиография содержит достаточно обширные материалы по концепции и технологиям организации крупномасштабного вычислительного эксперимента, включая вопросы геометрического моделирования и взаимодействия с системами автоматического проектирования, генерации адаптивных сеток, современных методологий программирования, распараллеливания алгоритмов и их отображение на архитектуру многопроцессорных компьютерных систем, проблемы искусственного интеллекта в математическом моделировании, а также неизбежные аспекты организации крупных проектов по вычислительно-информационным технологиям.

Данная книга является первой, и неизбежно наиболее математизированной, частью общей дисциплины “математическое моделирование”, предваряя такие насущные проблемы, как производительность вычислений, автоматизация построения алгоритмов, их распараллеливание и отображение на архитектуру многопроцессорных компьютерных систем, а также технологии программирования и самого вычислительного эксперимента. Эти вопросы, как и наиболее критические по ресурсоемкости методы решения больших алгебраических задач, относятся к последующей части издания.

Следует сказать, что написание учебника по математическому моделированию, а точнее говоря, серии учебников по его различным направлениям — это очень важная и непростая научно-образовательная и методологическая проблема, ждущая пока своего решения. Но именно от нее зависит успех и качество новых специалистов, призванных воплощать вычислительно-информационные технологии компьютерного эксперимента в инновации грядущего научно-технического прогресса.

Введение в предметную область

Математическое моделирование, в современном понимании, — это наука, сложившаяся в последние 30-40 лет двадцатого столетия и продолжающая свое бурное развитие в начале 21-го века. Истекший период является небольшим в историческом плане, но и немалым, примерно столько лет пророк Моисей водил свой народ по пустыне после исхода из Египта, прежде чем пришло на смену поколение людей, готовое к возвращению на землю обетованную и к восприятию нового учения, изложенного им в своих заповедях.

Данное фундаментальное научное направление есть порождение компьютеризации, и оно же является средством вовлечения математических знаний и вычислительно-информационных технологий во все сферы человеческой деятельности. Сами эти технологии непрерывно эволюционируют, в силу постоянно меняющихся архитектур электронных вычислительных машин (ЭВМ). Если в 1960-70 гг. уверенно можно было классифицировать ЭВМ первого поколения (ламповые), второго (полупроводниковые) и т.д., то после пятого поколения все критерии перемешались, и такое последовательное разделение потеряло смысл. В конце прошлого века произошли два главных события, определивших тенденции развития вычислительного оборудования (hardware): появление персональных компьютеров и многопроцессорных вычислительных систем (МВС), — между производителями которых началась азартная гонка за наращиванием оперативной памяти и быстродействия.

В 2008 г. мировое сообщество вступило в эпоху петафлопсных компьютеров (10^{15} арифметических операций в секунду, или флопс), появление которых было предсказано, еще за 20 лет, но казалось тогда совершенно фантастическим. Согласно закону Мура, одного из основателей компании ИНТЕЛ, хорошо выполняющемуся уже более 40 лет, за каждые 11 лет происходит 1000-кратное увеличение компьютерной производительности (именно с таким периодом происходило появление первого гигафлопсника (10^9 флопс), первого терафлопсника (10^{12} флопс) и первого петафлопсника), а в 2020 г. реально прогнозируется создание эксафлопсного суперкомпьютера (10^{18} флопс). На текущий мо-

мент — середина 2017 г. — в мире уже работает более 30 петафлопсников в Китае, США, Европе, Японии и России. Самый быстрый из них имеет скорость более 100 петафлопс. Ожидаемые “экзафлопсники” будут включать миллионы процессоров, миллиарды вычислительных ядер, огромную память и потребляемую электроэнергию в десятки мегаватт, что означает переход количественных изменений компьютеров в кардинально новое качество и ставит перед человечеством не только профессиональные и социальные, но и общегуманитарные проблемы.

Одновременно эволюционировало и программное обеспечение (software), обусловленное последовательным появлением все новых и новых версий операционных систем Windows и Unix, языков программирования Fortran 77, 90, 95, Python, Fortran 2008, C, C++ и т.д., а также бесчисленных программных библиотек и инструментариев. Рядовыми явлениями уже стали системы распараллеливания алгоритмов на распределенной и общей памяти, а также кластерные МВС, сетевые технологии (GRIDS) и облачные вычисления (Cloud Computing).

В 2008 г. под руководством Дж.Донгарры (Окриджская лаборатория, США) создан Международный комитет IESP (International Exascale Software Project) с участием более 50 экспертов различных стран, которые подготовили “дорожную карту” — развернутый документ RoadMap с Программой международных исследований на ближайшее десятилетие. Предлагаемый план включает ряд принципиально новых моментов:

- существующее программное обеспечение (ПО) хотя и позволяет сейчас распараллеливать алгоритмы на сотнях или тысячах процессоров, является своими корнями “однопроцессорным” и при переходе к экстремальным (петафлопсным и более) вычислениям должно быть полностью и срочно(!) обновлено, включая операционные системы, компиляторы, всевозможные приложения и инструментарии;
- тотальная замена старого ПО на новое по грандиозности своих масштабов беспрецедентно и возможно только при всеобщей интеграции и тесной координации международных работ;
- новый “софтвр” должен быть адаптируемым к вновь появляемым компьютерным архитектурам и общедоступным

(Open Source) в главном, что не исключает создание на его базе корпоративных “закрытых” приложений;

- эффективность масштабируемого распараллеливания алгоритмов и их отображения на архитектуру систем МВС становится не только математическим критерием, но и важным экономическим фактором организации крупномасштабных вычислительных процессов.

Согласно общепринятым представлениям, архитектура МВС в ближайшее десятилетие не претерпит революционных изменений (хотя работы по созданию квантовых компьютеров активно продолжаются) и будет эволюционировать в направлении гетерогенных многопроцессорно-многоядерных вычислительно-информационных сетей, использующих в узлах ускорители типа графических процессорных элементов (GPU, которые уже успешно применяются для арифметических задач) или/и программируемых логических интегральных схем (реконфигурируемые ПЛИС, или FPGA –Field Programable Gate Array). Что касается аппаратной инфраструктуры, то здесь на крайних позициях находятся многоядерные “персональные суперкомпьютеры” и голиафы – “датацентры” (Data Center, или ЦОД – центр обработки данных, или же ВЦКП – вычислительный центр коллективного пользования) с супер-МВС экстремальной производительности. Программное окружение ЦОДа базируется на информационных технологиях высокого уровня, предоставляющих пользователям универсальный доступ через интернет в технологиях “облачных” вычислений (Cloud Computing). В этом случае специалисту в конкретной предметной области для решения своей суперзадачи достаточно войти, например – с помощью мобильного устройства, на ВЦКП, который предоставляет не только мощные расчетные ресурсы, но и соответствующее программное обеспечение, в соответствии с SaaS-технологиями (Software as a Service).

Все эти кардинальные изменения компьютерных архитектур и программных платформ с каждым годом меняют технологию работы прикладных математиков и программистов, которые решают практические задачи. Тем не менее какие-то общие принципы математического моделирования выкристаллизовались, и цель настоящего введения — дать читателю и краткий исторический экскурс, и основные составляющие, а также когнитивные

и философско-методологические аспекты новой научной дисциплины.

В этой связи важно выделить ее главные квалификационные признаки. Данное понятие чрезвычайно многогранно, а его рассмотрение и классификация могут осуществляться в различных координатных плоскостях. Одно из возможных направлений анализа – с точки зрения типов изучаемых задач: электромагнетизм, теплофизика, упругопластичность, гидро-газодинамика, многофазная фильтрация, химическая кинетика, квантовые явления, динамические системы. Важно отметить, что особую актуальность приобретают междисциплинарные прямые и обратные задачи. При этом математические формулировки могут описываться самыми разными постановками: дифференциальными и/или интегральными уравнениями, классическими или обобщенными, вариационными и смешанными.

Другой взгляд на это же разнообразие возможен с точки зрения отраслевых приложений: машиностроение, металлургия, энергетика, геофизика, погода и климат, экология и катастрофы, биомедицина, материаловедение, экономика и социум.

Следующей независимой от предыдущих категорий являются вычислительные методы и технологии, применяемые на различных стадиях компьютерного эксперимента: геометрическое и функциональное моделирование, генерация сеток, аппроксимация исходных уравнений, решение получаемых алгебраических систем, методы оптимизации для обратных задач, обработка и визуализация численных результатов и т.д.

Главная идея приведенных рассуждений — математическое моделирование процессов и явлений занимает все большее место в разных областях общечеловеческой деятельности, включая научно-технические, производственные и социальные сферы, а проблемы его развития обретают и фундаментальные, и прикладные, и технологические аспекты, взаимосвязи или разграничение которых требует своего философско-методологического осмысления.

От пророчества жрецов – к тайнам мироздания

Известно, что математика — служанка всех наук. Еще Галилей отмечал, что “природа формирует свои законы на языке математики”. Теоретические исследования в физике, химии, биологии,

экономике, социологии и т.д. начинаются с того, что изучаемые процессы или явления описываются математическими средствами, например, на языке алгебраического, дифференциального или интегрального исчисления. Существует формальное определение математики, как науки о математических объектах, однако “чистая” наука — это от лукавого, и сам математический аппарат возник из потребностей отображения окружающего мира. До середины 20-го века, которую можно назвать эпохой последних энциклопедистов Л.Д.Ландау и Р.Фейнмана, фундаментальные результаты получались и доказывались “на кончике пера”, т.е. исключительно благодаря активности “серого вещества” в голове ученого. Конечно, научные изыскания невозможны без экспериментальных исследований, которые или предварительно обнаруживают новое явление, или впоследствии его подтверждают (практика — критерий истины).

С прогрессом науки новые знания даются все труднее, поскольку сами изучаемые объекты становятся сложнее и сложнее. Постепенно человек стал привлекать себе в помощь компьютеры — как для быстрых вычислений, так и для хранения больших объемов информации, не уместяющихся в памяти “гомо сапиенс”. С развитием ресурсов, мощностей и интеллекта ЭВМ их роль и искусство применения непрерывно возрастают, и вот однажды происходит переход количества в качество, ознаменовавший появление новой научной дисциплины — математического моделирования процессов и явлений, ставшего третьим путем, или компонентой, человеческого познания, наряду с теоретическими и экспериментальными исследованиями. Ее ветвями стали направления с недавно появившимися названиями “вычислительная физика”, “математическая биология” и т.д.

Если под математическим моделированием понимать просто применение вычислений, то такое искусство известно с незапамятных времен, по крайней мере древнеегипетских, когда жрецы уже умели предсказывать солнечные и лунные затмения. А наиболее эффектный пример в 18-м веке — это открытие планеты Плутон, сделанное Леверье на основе расчета возмущения орбит близлежащих планет.

Следует заметить, что бурное развитие вычислительной техники в 1950-е и 60-е годы было инициировано в США и СССР прежде всего насущными проблемами национальной безопасности, поскольку успехи атомного и космического проектов не мог-

ли быть достигнуты без огромных объемов расчетных работ. Из примеров последних лет актуальность моделирования подтверждается следующим фактом: передовые страны подписали договор о всеобщем запрещении ядерных испытаний только после того, как были созданы надежные вычислительные средства расчетов воздействия новейших видов вооружений, работа над которыми, естественно, не прекращается. Кстати говоря, выдающиеся физики Л.Д.Ландау и Р.Фейнман принимали непосредственное участие в атомных проектах СССР и США соответственно и были здесь пионерами применения математических ядерных расчетов. Более того, Л.Д.Ландау вместе со своим учениками Н.Н.Мейманом и И.М.Халатниковым во время проведения этих секретных работ в 1951-52 гг. предложил, исследовал и применил неявную разностную схему для решения дифференциальных уравнений, названную ими “нелокальной”. Этот знаменательный для вычислительной математики результат был, естественно, долгое время неизвестен и опубликован ими значительно позже. В годы холодной войны аналогичная судьба ждала и многие другие научные открытия. Однако есть мнения, что нет худа без добра и что политическое противостояние двухполярного мира явилось катализатором научно-технического прогресса и стимулировало многие разработки не только военного, но и общечеловеческого значения.

Поразительно, что за последние 50 лет рост мощностей ЭВМ не ослабевает. Однако остается справедливым и инвариант академика Н.Н.Яненко, установленный им в 70-е годы: большая задача решается примерно одинаковое время на ЭВМ любого поколения. Здесь неформально принимается, что малые задачи — это те, которые решаются до нескольких минут, средние — десятки минут, а большие — часы, сутки и более того. Фактически происходит трансформация понятия “большая задача”. Понятно, что с каждым годом растут практические требования и к адекватности математических моделей, и к точности численных решений, что приводит к росту объемов расчетов в соответствии с известной поговоркой “аппетит приходит во время еды”. Наиболее ресурсоемкими являются современные междисциплинарные задачи. Данное понятие относится к моделированию взаимосвязанных явлений различной природы, когда приходится решать системы дифференциальных уравнений со многими неизвестными функциями. Здесь можно привести две иллюстрации из

актуальных прикладных областей. Первая касается мира нанотехнологий и связана с учетом самых разных процессов в субмикронных элементах электроники: электромагнитные поля и токи, квантовые эффекты, теплоперенос с фазовыми переходами, термоупругость и прочность, — представляющих солидный джентльменский набор задач матфизики. Второй пример — это комплексное моделирование полей георазведки: сейсмических, электромагнитных, гравитационных и магнитных аномалий Земли, совместное использование которых наиболее перспективно для повышения качества интерпретации геофизических данных при поиске полезных ископаемых.

С прагматической точки зрения можно выделить два уровня моделирования. Нижний является в определенном смысле чисто познавательным и сводится к решению прямых задач, в которых при заданных исходных условиях исследуются свойства или поведение изучаемого объекта. Второй уровень можно назвать конструктивным, и он заключается в решении обратных задач, предполагающих, что некоторые из исходных величин зависят от параметров, значения которых надо определить по задаваемым дополнительно условиям, или ограничениям. Одна из типичных ситуаций здесь — проектирование прибора или устройства с требуемыми техническими характеристиками. Вторая иллюстрация — идентификация параметров модели с недоопределенными данными, которые необходимо конкретизировать путем сопоставления результатов расчетов и натурных измерений, мера разности которых принимается за целевой функционал. Решение серьезной обратной задачи “в лоб” удается исключительно редко, и реально оно приводит к многократному решению прямых задач, при последовательной коррекции параметров с помощью оптимизационных алгоритмов, заключающихся в условной минимизации целевого функционала. В свою очередь, здесь необходимо выделить два класса проблем принципиально различного уровня сложности — задачи локальной и глобальной минимизации, последние из которых связаны с поиском различного числа минимумов.

Исключительная роль и универсальность математического аппарата заключается в том, что одинаковыми или похожими уравнениями описываются процессы и явления из самых разных сфер окружающего мира. В силу этого и возникла счастливая возможность создания унифицированных вычислительно-

информационных алгоритмов и технологий, обеспечивающих эффективное применение ЭВМ в машиностроении и металлургии, химии и биологии, геофизике и экологии, экономике, социологии, медицине, финансовой деятельности и т.д.

Термин “моделирование” имеет очень широкий смысл. Существует физическое, или аналоговое, моделирование, основанное на законах подобия разномасштабных или вообще принципиально отличающихся процессов и явлений. Например, в середине прошлого века широко применялись электрические сетки, мгновенно (со скоростью света) определявшие приближенное решение уравнения Пуассона, которое описывает многие стационарные поля (тепловые, магнитные и др.). Такие устройства не выдержали конкуренции с ЭВМ из-за своей низкой точности и трудоемкой настройки на конкретную задачу. Другая иллюстрация — аэродинамические трубы, в которых успешно и сейчас изучают поведение летательных аппаратов (хотя, возможно, вскоре эти испытания полностью или частично также заменят математическим моделированием).

Для изучения различных сложных систем широко применяется имитационное моделирование. Объектом исследования может быть сама вычислительная система, состоящая из огромного количества элементов, оптимальное проектирование которой может значительно повысить ее быстродействие. Другие примеры — телекоммуникации и различные игры (не только развлекательные но и военные). Для этих исследований главным средством служит дискретная математика, проблематика которой выходит за рамки данной книги.

Традиционное понимание моделирования естественно-научных процессов связано с нахождением непрерывных величин, подчиняющихся дифференциальным и интегральным уравнениям, описывающим континуум. Хотя мир имеет дискретную иерархическую структуру (микромир, кристаллическое вещество, макрокосмос), Лейбниц и Ньютон создали идеальные математические объекты, которые при всей своей абстракции дают человеку очень мощный аппарат для исследования конкретных природных явлений. Например, общеизвестен физический закон Фика, согласно которому поток тепла пропорционален градиенту температуры. Производная от функции $T(\vec{r})$ в точке с радиус-вектором $\vec{r}$ вдоль направления

$\delta\vec{r}$ определяется как предел отношения $\lim_{|\delta\vec{r}| \rightarrow 0} \frac{T(\vec{r} + \delta\vec{r}) - T(\vec{r})}{|\delta\vec{r}|}$, но только если этот предел существует. Понятно, что если длина вектора $\delta\vec{r}$ сравнима с межатомным расстоянием, то понятие производной от температуры не существует. Однако при переходе к макроскопическим длинам, т.е. порядка тысяч или миллионов атомных размеров, определение градиента имеет смысл и закон Фика хорошо описывает реальность. Парадоксально, что численные методы решения функциональных уравнений базируются на их дискретизации, т.е. от бесконечно малых величин и их приращений мы приходим назад к конечным, а от бесконечномерных объектов — к конечномерным.

Типичные задачи математического моделирования описываются системами уравнений тепломассопереноса, фильтрации многофазной жидкости в пористых средах, термоупругости и пластичности, электродинамическими уравнениями Максвелла и квантомеханическим уравнением Шредингера, системами уравнений гидро-газодинамики и т.д. Искомое решение требуется искать в некоторой расчетной области определенной размерности, ограниченной или неограниченной, на внешних границах которой задаются какие-то краевые условия, определяемые исходной физической моделью. Область может состоять из подобластей с различными материальными свойствами сред, отображаемых в коэффициентах функциональных уравнений, на внутренних границах раздела которых формулируются условия сопряжения. Внешние и внутренние границы могут быть многосвязными, иметь сложную конфигурацию с углами и сильно разномасштабными деталями, которые приносят свои сложности в изучаемую проблему и в численные алгоритмы, особенно в многомерных случаях с динамическими объектами.

Эквивалентные математические постановки могут записываться в виде вариационных формулировок, определяющих не классические, а обобщенные решения. Для “хороших” задач функциональный анализ позволяет установить свойства существования, единственности и корректности решений. Это очень важно, чтобы не искать кошку в темной комнате, которой там нет. Тем не менее, зачастую практика заставляет браться за проблемы, о которых теоретически мало что известно, и здесь грамотно поставленный вычислительный эксперимент может быть

самодостаточным и дать обоснование гарантированной точности полученного результата.

Остановимся теперь на нескольких терминологических вопросах. Несколько перефразируя сформулированное академиком А.Н.Тихоновым в “Математической энциклопедии” определение, *математической моделью* назовем описание какого-либо явления или процесса, выраженное с помощью математических средств. Соответственно, *математическое моделирование* означает изучение реального явления или процесса с помощью его математической модели.

Другими словами, построение модели, т.е. обоснованная постановка математической задачи, ее исследование и решение в совокупности и составляют суть моделирования. Следует заметить, что для одного и того же изучаемого объекта, например, физической проблемы, или физической модели, могут быть сформулированы различного типа математические модели, в большей или меньшей степени эквивалентные между собой (соответствующие иллюстрации будут рассмотрены в первой главе). Поскольку научный интерес представляют достаточно сложные проблемы, записываемые с помощью функциональных соотношений (например, систем дифференциальных и/или интегральных уравнений — СДУ, СИУ или СДИУ), точное аналитическое решение которых найти невозможно, на помощь приходят приближенные численные методы. Таким образом, формулируется *вычислительная модель*, или совокупность алгоритмов, которые должны обеспечить адекватность соответствующего численного решения, т.е. точность и еще какие-либо необходимые свойства по отношению к исходному решению. Если предыдущая область деятельности относится к прикладной математике, то в последнем случае наступает очередь математика-вычислителя. И это еще не все. Разработанные алгоритмы необходимо запрограммировать и исполнить в конкретном операционном окружении на некоторой ЭВМ (все это является прерогативой прикладного программиста). Именно получаемый исполняемый код (execution code, или “экзешник” на русском профессиональном жаргоне), который можно назвать *компьютерной моделью*, и определяет полностью реализацию и окончательный результат вычислительного процесса по решению задачи.

Необходимо подчеркнуть, что для выбранной математической модели могут быть построены весьма отличающиеся *вычис-*

лительные модели, а для каждой из последних — разнообразные компьютерные модели, и от всей сконструированной цепочки зависит успех или неуспех математического моделирования. А поскольку в этой последовательности ключевыми факторами являются численные методы и мощная ЭВМ, то зачастую употребляются словосочетания *компьютерное моделирование* и *численное моделирование*. Эти формулировки будем считать синонимами математического моделирования, разве что с акцентами на программную реализацию или на алгоритмы соответственно. Отметим, что в англоязычной литературе, наряду со словом *modeling*, употребляется его аналог *simulation*, обычно при рассмотрении конкретных процессов или явлений.

Важно понимать, что моделирование — это всего лишь инструмент для изучения объектов реального мира, и последнее слово остается за экспертом, владеющим не только предметными знаниями, но и передовыми компьютерными средствами, умеющим ставить и решать математические задачи для достижения практических целей. Здесь сродни сопоставить науку с искусством и провести сравнение с резцом скульптора, который в руках мастера способен из глыбы камня высечь одухотворенное произведение.

Автору пришлось слышать из уст известного геолога такую фразу: “Я не доверяю “модельерам”, у них что заложишь, то и получишь”. Речь же в данном случае идет о фундаментальной проблеме исследования тысячелетних геодинамических процессов в литосфере и мантии Земли. Экспериментальные измерения на больших глубинах слишком дорогие, а зачастую получить необходимый материал и вообще невозможно. В условиях жесткого дефицита исходных данных проблема определения эволюции геологических структур представляет собой головоломную обратную задачу, включающую уравнения упругопластичности и многофазных течений со сложными термодинамическими условиями. Ее решение имеет не только теоретическое значение, но и определяет историю формирования природных кладовых, знание которой необходимо для прогнозирования залежей полезных ископаемых. Очевидно, что идеальная стратегическая организация в такой ситуации заключается в тесном сотрудничестве геофизика и математика по разработке модели, планировании вычислительных и натурных экспериментов к интерпретации расчетных данных.

В этой связи следует сказать о понятии точности моделирования, связанной с такой общей категорией, как истина и познаваемость реальности. В силу разных причин, на каждой стадии изучения объекта привносятся свои погрешности: допущения физической модели, математические приближения, алгоритмические погрешности и ошибки машинных операций. Кроме того, используемые для расчетов исходные числовые данные, получаемые в итоге измерений или из других соображений, известны с большой степенью неопределенности. В этих условиях получение достоверного результата моделирования – это путь между Сциллой и Харибдой, требующий определенной решительности в условиях риска. Здесь необходимо тщательное проведение этапов валидации моделей и верификации кода, чтобы гарантировать адекватность итоговых расчетных данных, которые при неквалифицированном исполнении могут содержать в основном привнесенные артефакты.

Имеется еще такой важный философско-методологический момент: моделирование не является волшебной лампой Алладина, способной по мановению открыть все тайны мироздания. Природа неисчерпаема, а наука не всесильна. Однако предназначение ученого сообщества — не впадать в пессимистический агностицизм, а браться за новые все более сложные задачи, и в их решении роль математического моделирования будет становиться определяющей.

Технологии решают всё

При всем многообразии прикладных проблем моделирования вычислительные методы и технологии их решения укладываются в стройную структуру, каждая компонента которой имеет свою ясную цель и способы ее достижения. В 70-80 годы прошлого века благодаря исследованиям в научных школах Г.И. Марчука, А.А. Самарского и Н.Н. Яненко сложилась методология крупномасштабного вычислительного эксперимента, определяющая в этой многопрофессиональной деятельности последовательность достаточно автономных стадий. Согласованная программная реализация каждой из них в совокупности и определяют успешность и востребованность вычислительно-информационных технологий математического моделирования. И здесь очень уместно напомнить тезис классика Никлауса Вир-

та: “Программа — это алгоритмы + структура данных”.

Первый этап технологической цепочки при решении большой реальной задачи связан с геометрическим и функциональным моделированием, заключающимся в создании, редактировании и модификации информационной модели решаемой проблемы. Здесь надо описать (аналитически и/или графически) все геометрические объекты: объемные фигуры, поверхностные сегменты, ребра, вершины, — а также их топологические взаимосвязи и возможные перемещения. Аналогичные вопросы возникают в существующих многочисленных графических библиотеках и пакетах программ, а также в системах автоматизации проектирования: САПР, или CAD (Computer Aided Design), CAE (Engineering), PLM (Product Lifecycle Management, т.е. управление жизненным циклом продуктов) и др., — широко применяющихся в промышленности, но цели и средства там другие.

По большому счету, в мире происходит конвергенция программных систем моделирования и САПР, в идеале исходные и выходные данные компьютерных расчетов должны представляться в форме конструкторской документации, непосредственно применимой для производственных процессов. Эти тенденции только усиливаются в наступающую эпоху постиндустриализации, символами которой становятся “интернет вещей”, “аддитивные технологии”, “3D-принтеры” и т.п.

К функциональным объектам относятся фрагменты задания решаемых уравнений и их коэффициентов, граничных условий и начальных данных, целевых функционалов и специальных требований к расчетам. Другими словами, исходная информация этого вида определяет материальные свойства среды, форма задания которой варьируется от тривиальных числовых коэффициентов до решения сложной проблемы идентификации параметров модели, требующей вовлечения автоматизированной системы измерений и фактической реализации обратных задач. От гибкости используемых средств зависит и уровень интеллектуальности моделирующей системы, и ее универсальность, и потенциальная широта использования в производстве.

Кроме исходных данных, необходимо указать, что и в какой форме требуется получить. Могут быть также предписаны и применяемые методы или даже детальные вычислительные схемы, определяющие однозначно процесс математического моделирования в конкретной операционной обстановке. Акцен-

тируясь на перечисленных аспектах, мы фактически приходим к автоматизации построения моделей и алгоритмов. Итоговая проблема в данном случае состоит в создании гибкого входного и выходного интерфейса для конечного пользователя, который в идеальном случае представляется средствами специализированных языков программирования естественного типа, ориентированных на конкретную профессиональную подготовку, с возможным активным использованием как текстовых, так и графических средств. А поскольку пользователями одного и того же разрабатываемого обеспечения могут быть специалисты самого разного профиля: математики, программисты, инженеры, техники и т.д., – то становится очевидной необходимость образования богатой гаммы прикладных когнитивных подходов. “Сухим остатком” выполнения первого этапа являются геометрическая и функциональная структуры данных (ГСД и ФСД), содержащие полную информацию о постановке исходной задачи. Важно отметить, что такие структуры должны быть расширяемыми и совместимыми по интерфейсам с внешними программными продуктами, в которых вопросы стандартизации геометрической информации уже достаточно хорошо проработаны.

Если компьютер уже “знает” поставленную задачу, то ее численное решение начинается с дискретизации расчетной области, другими словами, с построения сетки. Многообразие возникающих здесь вопросов позволяет оценить простое перечисление часто применяемых типов сеток: адаптивные, структурированные, неструктурированные и квазиструктурированные, согласованные, несогласованные и мортарные, регулярные и нерегулярные, статические и динамические, и т.д. Важно отметить, что современные наиболее эффективные подходы связаны с достаточно сложными дискретными объектами и их преобразованиями, включая последовательности иерархических сеток и их локальные сгущения, декомпозиции сеточных областей на подобласти, динамическую перестройку сеток, апостериорный и/или априорный учет свойств искомым решений и т.д.

Имеется большое количество критериев качества сеток, с фундаментальными вопросами которых связаны имена знаменитых русских ученых Б.Н.Делоне и Г.Ф.Вороного, но понятие их оптимальности — это пока тайна за семью печатями. Проблема даже определения оптимальной сетки слишком сложна и пока что в практических разработках даже не формулируется. Основ-

ной наиболее распространенный принцип заключается в эмпирическом подходе к использованию плотностей распределения сеточных узлов из общих качественных соображений. Отдельные имеющиеся методические результаты об оптимизации сеток в частных случаях являются только исключениями, подтверждающими общее правило. Академик С.К.Годунов предложил термин “хорошие” сетки, в которых неожиданно гармонируют и чисто математические, и эстетические качества. Действительно, изображения сконструированных с помощью высокоинтеллектуальных алгоритмов сеток для экзотических областей ассоциируются с магическими рисунками великого графика Эшера. Фундаментальным вопросам и методам генерации сеток для двумерных и трехмерных областей посвящено огромное количество статей и монографий. На мировом рынке прикладного программного обеспечения широко распространены и коммерческие, и свободно доступные инструментальные пакеты, которые являются самостоятельными продуктами и могут применяться в разных приложениях. Однако открытых теоретических и прагматических вопросов здесь еще великое множество, а по их решению непрерывно ведутся разработки и методического, и технологического характера.

Результатом работы по дискретизации исходной задачи является сеточная структура данных (ССД), которая в совокупности с ГСД и ФСД также полностью отображает постановку решаемой проблемы, но уже не на непрерывном, а на дискретном уровне. Здесь также необходимо подчеркнуть важность совместимости ССД с программными интерфейсами уже имеющихся сеточных генераторов, и как правило – предусмотреть наличие конверторов входной / выходной информации.

Следующий вычислительный этап — это аппроксимация исходной функциональной задачи, основными орудиями которой являются методы конечных объемов и конечных элементов (МКО и МКЭ). Первые из них реализуют дискретизацию интегральных законов сохранения и представляют собой фактически давно известные балансные конечно-разностные схемы (новое – хорошо забытое старое). МКЭ, или вариационные методы, имеют сейчас наибольшее распространение, во-первых, в силу построения уникальной теории, обеспечивающей обоснование алгоритмов различных порядков точности для очень широкого круга задач на разных типах конечных элементов (парал-

лелепипеды, тетраэдры, призмы и т.д.). Второй, не менее важный, фактор состоит в изобретении унифицированных поэлементных технологий вычисления локальных и сборки глобальных матриц, с помощью которых намного повышается производительность программных реализаций, допускающих эффективное распараллеливание алгоритмов. Итогом этого этапа является полная алгебраизация исходной задачи, т.е. формирование систем линейных или нелинейных алгебраических уравнений (СЛАУ или СНАУ).

На выходе аппроксимационной стадии должна быть универсальная алгебраическая структура данных (АСД), охватывающая все основные типы матриц, встречающихся на практике. Исторически вычислительная алгебра – наиболее продвинутая в технологическом плане область математики, и здесь уже разработан спектр форматов данных для матриц, в том числе разреженных (например, разреженный строчный формат CSR – Compressed Sparse Row).

Решение алгебраических задач – “узкое горлышко” численного моделирования, поскольку объем арифметических операций на данном этапе нелинейным образом растет с увеличением числа узлов сетки и размерности алгебраических систем, а их порядки в современных задачах достигают сотен миллионов и многих миллиардов, что представляет серьезную проблему даже для суперкомпьютеров. Разработка здесь быстрых алгоритмов и программ – это бурно развивающаяся в нашей стране и за рубежом область Computer Science, а главное направление ее развития состоит в создании новых параллельных итерационных методов в подпространствах Крылова. Важно отметить, что в мире уже имеется большое количество программных “решателей”, в том числе доступных через Интернет. Реализация новых алгоритмов и программ здесь очень актуальна и сопряжена с многочисленными технологическими проблемами, так как реальные СЛАУ имеют сложные блочные конфигурации и многоуровневые представления матричных графов, а для их экономичного хранения и выполнения на иерархической памяти требуется использовать громоздкие структуры данных, значительно снижающие производительность программного кода.

Если для многомерной задачи получено векторное решение на сложной нерегулярной сетке, то его человеческое понимание и анализ невозможны без предварительной обработки и визуа-

лизации численных результатов. Целью расчетов являются не числа, а их осмысливание. Например, для компактного и наглядного изображения многомерных векторных полей соответствующее прикладное программное обеспечение (ППО) должно иметь развитый набор инструментов для формирования таких типичных представлений, как изоповерхности, силовые линии, сечения, всевозможные графики и т.д. В последние годы здесь активно развиваются новые математические методы и технологии, сформировавшиеся в самостоятельное направление “изометрических анализ”.

Кроме того, всего заранее предусмотреть нельзя, и в интеллектуальной системе моделирования должны содержаться средства автоматизации программирования различных возможных характеристик итоговых данных отражающих “виртуальную реальность”. И наконец, необходимо понять возможное разнообразие профессий конечных пользователей, желающих в каждом конкретном случае иметь наиболее комфортное представление итогов использования компьютера, что и определяет его производственных эффект. Здесь могут помочь также существующие интеллектуальные графические системы, однако остается актуальной проблема значительного повышения их быстродействия, так как данный рутинный этап чрезвычайно ресурсоемок, и машинное время отображения расчетных данных может превышать затраты на математические вычисления. В последние десятилетия широкое распространение получили специализированные параллельные “графические ускорители” (GPU – от Graphic Processor Unit), а непрерывное совершенствование аппаратных средств влечет за собой и необходимость новых программных разработок.

Рассмотренные технологические этапы составляют основу решения прямых задач, в которых надо найти решение, однозначно и корректно определяемое, при полностью заданных исходных данных. Однако на практике конечной целью являются так называемые обратные задачи, в которых сама постановка содержит варьируемые параметры, которые надо подобрать из некоторых дополнительных требований. Здесь примерами могут служить оптимальное проектирование какой-то установки с нужными характеристиками или минимизация эксплуатационных потерь промышленного оборудования. Другого типа классическая обратная задача – это идентификация параметров ма-

тематической модели, где актуальнейшими приложениями являются проблемы поиска полезных ископаемых по результатам натурных измерений каких-то полей (сейсмических, электромагнитных и др.). Для таких обратных задач естественным образом добавляются оптимизационные алгоритмы условной минимизации целевых функционалов при заданных линейных или функциональных ограничениях. Пусть, например, прямая задача описывается уравнением

$$Lu(\vec{p}, x, t) = f(u, \vec{p}, x, t),$$

где L есть некий дифференциальный и/или интегральный оператор, x и t – пространственная и временная независимые переменные, а $\vec{p} = \{p_k\}$ – m -мерный вектор искомых параметров. Наиболее общими методами решения таких проблем являются оптимизационные, которые основываются на минимизации определяемого в описании обратной задачи целевого функционала:

$$\Phi_0(u(x, t, \vec{p}_{opt})) = \min_{\vec{p}} \Phi_0(u(x, t, \vec{p})),$$

при заданных m_1 линейных и m_2 нелинейных ограничениях на параметры задачи:

$$\begin{aligned} p_k^{min} &\leq p_k \leq p_k^{max}, \quad k = 1, \dots, m_1, \\ \Phi_l(u(x, t, \vec{p})) &\leq \delta_l, \quad l = 1, \dots, m_2, \quad m_1 + m_2 = m. \end{aligned}$$

Здесь Φ_l – некоторые функционалы, определяемые какими-либо дополнительными условиями на исходные данные рассматриваемой задачи условной минимизации.

Традиционные подходы к решению оптимизационных задач, основанные на введении лагранжиана и экстремальных условий Каруша–Куна–Такера (ККТ), претерпели в последние десятилетия революционное развитие, обусловленное появлением новых методов внутренних точек, последовательного квадратичного программирования и доверительных интервалов, кардинально повышающих надежность, скорость и точность поиска локальных или глобальных минимаксных точек. Здесь следует подчеркнуть исторически значимый математический результат – теорию и методы линейного программирования, которые были открыты в первой половине 20-го века лауреатом Нобелевской премии Л.В. Канторовичем.

Численное решение обратных постановок с помощью оптимизационных методов базируется на итерационных процедурах и зачастую требует многократного решения прямых задач. Особенные неприятности возникают при сложных зависимостях целевого функционала от параметров, например, при так называемых овражных особенностях. Заметим при этом, что пока мы говорим о так называемых задачах локальной минимизации [73], когда вычисляемые последовательные приближения сходятся к своему единственному минимуму. Гораздо сложнее является задача глобальной минимизации, когда целевой функционал имеет много минимумов, и при поиске минимального из них приходится несколько раз находить локальные минимальные значения.

Необходимо отметить, что комплекс оптимизационных алгоритмов может быть реализован независимо от того, к какой предметной области относится решаемая задача, поскольку вычисление целевого функционала и соотношений-ограничений в виде равенств или неравенств может быть осуществлено с помощью внешних программных процедур и/или функций, созданных сторонними разработчиками.

Важно подчеркнуть, что даже при наличии самого идеального ППО компьютерное моделирование достаточно сложных процессов или явлений — это сама по себе многогранная творческая деятельность. Например, для систематического изучения каких-то приложений сначала надо убедиться, что применяемые модели и методы соответствуют техническим требованиям, для чего предварительно проводятся пробные расчеты с анализом адекватности получаемых данных. Затем осуществляются сами исследования, которые могут составлять крупномасштабный машинный эксперимент, требующий своего планирования и составления методики. Последнее означает необходимость гибких возможностей составления расчетных схем, что требует создания соответствующих языков (декларативного или императивного типов) для управления вычислительными процессами, а также для создания дружественных и гибких интерфейсов, ориентированных как на конечных пользователей создаваемых программных продуктов, так и на их разработчиков. Характерным примером здесь может служить создаваемая под эгидой Департамента энергетики США система описания наукоемких интерфейсов SIDL (Scientific Interface Definition Language). Дру-

гой популярный проект формирует задачу создания “фабрики” специализированных под приложения языков и носит название DSL – Domestic Specific Language.

И наконец, поскольку моделирование является не самоцелью, а орудием познавательной или производственной сферы, то по результатам расчетов для принятия решения в состав ППО должны быть заложены или какие-то когнитивные принципы, или средства подключения к САПРовским инфраструктурам, или технологии поддержки и оптимизации эксплуатационных режимов конкретных процессов, или автоматизации управления и планирования натуральных экспериментов при проведении фундаментальных исследований окружающего мира. Но эта деятельность уже выходит за рамки самого математического моделирования, хотя сами алгоритмы и технологии принятия решений (см., например, [44]) имеют общеметодологическое значение.

В заключение о вычислительно-информационных технологиях следует сказать, что они должны базироваться на современных достижениях бурно развивающейся теоретической математики, активное внедрение которых в прикладные программные комплексы приводит к постоянным расширениям, обновлениям и требует высокого уровня автоматизации построения математических моделей и алгоритмов. А последний факт приводит к такому парадоксальному на первый взгляд выводу, что в эпоху “модных” супервычислений особенно востребованными становятся средства и подходы искусственного интеллекта. По одному образному выражению, наступает эра перехода от “палеоинформатики” к “неоинформатике”.

Коллективизация и индустриализация в программировании

Дискуссии о необходимости перехода практики прикладного программирования от кустарного подхода к индустриальному проводились еще лет 40 назад. Однако нынешняя ситуация с драматическим усложнением математических методов, а также компьютерных, программистских и вычислительных технологий делает уже невозможным ситуацию, когда один человек — “и швец, и жнец, и на дуде игрец”. Можно предположить, что чем уже специализация профессионала, тем выше будет качество его тру-

да. Но когда сильно возрастает количество узких специалистов, участвующих в одной общей разработке, возникает проблема эффективного согласования армии исполнителей, и здесь требуется нечто вроде гениального изобретения конвейера, сделанного Фордом в автомобилестроении.

В системном программировании осуществление крупных проектов с большим количеством исполнителей является обычной практикой, имеющей многочисленные успешные примеры в виде операционных систем, компиляторов, САПР-овских продуктов и т.д. Причем эти разработки имеют свою многолетнюю историю, массовую эксплуатацию с квалифицированным сопровождением и непрерывным развитием. Однако следует признать, что в этих разработках состав исполнителей и стоящие перед ними конкретные задачи являются достаточно однородными, в сравнении с созданием наукоемких разнородных комплексов, имеющих зачастую поисковый исследовательский характер. Тем не менее хотелось бы надеяться найти положительный ответ на такой сакраментальный вопрос: а можно ли построить программную систему математического моделирования с уровнем технологичности, приближающейся к широко эксплуатируемым операционным системам или компиляторам?

Итак, мы выделили следующие основные этапы, предполагающие соответствующие компоненты вычислительных технологий математического эксперимента: геометрическое и функциональное моделирование, генерация сетки, аппроксимация задач, решение алгебраических систем, обработка и визуализация результатов, оптимизация параметров модели. Каждая из этих подсистем может осуществляться автономно, если предварительно оговорить и зафиксировать информационные интерфейсы между ними, т.е. различные структуры данных: геометрическую и функциональную, сеточную, алгебраическую, графическую и оптимизационную.

Если прикладная программа пишется автором в исследовательских целях, например, для иллюстрации эффективности нового алгоритма или физического эффекта в статье или диссертации, то здесь вместо технологий действует простое правило “хозяин-барин”, и такой одноразовый продукт после использования обычно идет в корзину. Как правило, подобные программные разработки создаются под одну частную задачу: простейшая расчетная область, тривиальное дифференциальное урав-

нение и т.д.

Однако компьютерная программа — яркий образец интеллектуальной собственности, которая может иметь немалую материальную ценность, о чем свидетельствует биография Билла Гейтса, ставшего самым богатым человеком в мире благодаря системам DOS и WINDOWS. Лучшие программные комплексы для решения больших прикладных задач, получившие в 70-е годы популярное название пакетов прикладных программ (ППП), — это уже серьезные проекты, требующие месяцы работы и которые не по-хозяйски было бы выкидывать. Крупные ППП являются результатом коллективных усилий и заключают в себе человеко-годы овеществленного высококвалифицированного труда.

Следующим шагом к переходу от индивидуального ремесленничества к индустриальным принципам создания прикладного математического обеспечения для широких классов задач должна стать концепция разработки базовой системы моделирования (БСМ), в которой перечисленные выше шесть технологических компонент будут составлять проблемно независимое инструментальное ядро. Например, подсистема генерации сеток, если она включает достаточно содержательный набор сеточных типов (адаптивные, динамические, составные, вложенные и т.д.), в определенной степени никак не ориентирована на конкретную задачу и должна быть применимой в самых разных приложениях. Соответствующие алгоритмы дискретизации предпочтительно реализовать в форме расширяемой программной библиотеки, допускающей пополнение процедурами от различных разработчиков. В аналогичном стиле легко конструируется и комплекс алгебраических решателей, который для своей внутренней содержательности должен иметь набор как универсальных, так и специализированных современных методов (прямых и итерационных) для различных типов СЛАУ: симметричных и несимметричных, положительно определенных и знако-неопределенных, полного и неполного ранга, — с типовыми матричными форматами хранения элементов, никак не связанными с изначальной математической постановкой.

Разрабатываемые на таких модульных принципах подсистемы приобретают двойное назначение. С одной стороны — это неотъемлемые компоненты единой БСМ, в рамках которой происходит активная эксплуатация и одновременное развитие. Но

они также являются самодостаточным программным продуктом и могут независимо применяться в других прикладных системах.

В целях функциональной полноты БСМ должна содержать средства конфигурационного управления, предназначенные для сборки ППП, ориентированных на решение конкретных задач с помощью общего инструментального ядра и содержащих соответствующие надстройки для специалистов разной квалификации. Очень важно, чтобы архитектура БСМ обладала гибкостью, позволяющей естественное расширение всей системы за счет подключения новых математических моделей, численных методов и программных инструментов.

Концепция коллективного создания различными группами разработчиков интегрированных программных окружений для решения широкого класса приложений, в том числе междисциплинарных (или multi-physics) уже хорошо известна в мире, а примерами реализации таких подходов могут служить достаточно популярные системы OPEN FOAM и DUNE (Distributed Unified Numerical Environment).

Безусловно, БСМ с непрерывным обновлением современных методов и технологий представляет собой в высшей степени наукоемкий продукт, а его рождение и активное существование невозможно без внедрения фундаментальных результатов многообразных вычислительных наук (Computer Sciences). В то же время — это дорогостоящая промышленная разработка, весь жизненный цикл которой требует участия профессиональных программистов, инженеров и администраторов с производственным стилем организации труда, нацеленным на достижение экономического успеха. Понятно, что такая деятельность не является прерогативой академической или университетской науки, и здесь необходимо партнерство представителей различных трудовых сообществ. В советское время такое сотрудничество осуществляла отраслевая наука, занимавшаяся “внедрением фундаментальных достижений в народное хозяйство”.

Отображение алгоритмов на архитектуру ЭВМ

Заголовок данного раздела – это фактически лозунг, провозглашенный в 1970-е годы директором Вычислительного центра СО АН СССР Г.И.Марчуком. Ныне его непосредственный перевод

на английский: mapping of algorithms on the computer architecture,— является основным принципом разработчиков компьютерного оборудования (компании Intel, AMD и др.) для достижения высокой производительности расчетов (High Performance Computing). Сейчас в заголовке слово “ЭВМ” следовало бы заменить на “МВС” (многопроцессорная вычислительная система), поскольку современные большие задачи (а они представляют главный профессиональных интерес) надо решать на современных компьютерах, которые насчитывают десятки, сотни и тысячи процессоров.

Понятие “суперкомпьютер” меняется в наше время с калейдоскопической быстротой, и если еще несколько лет назад мы с уважением относились к ЭВМ с быстродействием в 1 терафлоп (10^{12} операций в секунду), то сейчас она далеко от попадания и во всемирный рейтинговый список самых мощных компьютеров ТОП-500, и в аналогичную российскую категорию ТОП-50. Детальная информация об архитектуре современного компьютера — “терра инкогнито” для простого смертного, однако существует определенный уровень знаний, который можно отнести к минимуму суперкомпьютерной грамотности с названием “это должен знать каждый”, кто хоть что-то желает представлять о математическом моделировании.

Традиционно МВС делятся на системы с распределенной и с общей памятью, хотя сейчас они на самом деле чаще бывают гибридными, т.е. каждый процессор имеет память, общую с несколькими “соседними” процессорами, но она же доступна остальным “дальним” процессорам только через более медленную соединительную шину. Типичная конфигурация суперкомпьютера — кластер, представляющий собой сеть из нескольких десятков или сотен вычислительных узлов, соединенных по принципу “каждый с каждым”. Отдельный узел состоит из двух, например, центральных процессоров (CPU, Central Processor Unit), имеющих общую многоуровневую оперативную память, а также графических вычислительных устройств с большим количеством (несколько сот) облегченных ядер для быстрого выполнения арифметических операций: ускорители типа GPGPU (General Purpose Graphic Processor Unit) или Intel Phi. Альтернативная архитектура МВС — это сервер, включающий до нескольких десятков процессоров (ядер) с общей памятью. Такой компьютер имеет меньшие коммуникационные потери при вычислениях, но его создание резко усложняется и дорожает

при увеличении количества процессоров.

Отдельный процессор содержит несколько арифметических устройств (сумматоры, умножители и т.п.) со сверхбыстрыми регистрами памяти, где могут размещаться два-три десятка вещественных чисел, непосредственно над которыми выполняются операции, а также очень быструю память (но медленнее регистров CPU) под названием кэш (новое русское слово от английского *cache*). Объем кэша ограничен и позволяет хранить всего лишь несколько тысяч вещественных чисел. На самом деле кэш имеет свои внутренние уровни, причем самый нижний уровень, через который происходит передача данных регистрам, имеет наибольшую скорость и наименьший объем, а с повышением уровня его скорость замедляется, но объем – увеличивается. Управление этими устройствами, как правило, осуществляется компилятором, и его средства прикладному программисту недоступны. Последние модели компьютеров являются многоядерными, а их главная особенность заключается в том, что несколько вычислительных ядер располагаются на одном кристалле(чипе) и имеют общий кэш.

Если говорить о последних новинках суперкомпьютеров, то надо сказать о ПЛИС – программируемых логических интегральных схемах, в англоязычной литературе называемых FPGA – Field Programable Gate Array, которые призваны осуществить давнюю мечту математиков-вычислителей и прикладных программистов – конструирование ЭВМ “под задачу или алгоритм”. Технологически эта проблема уже преодолена, и дело осталось только за коммерчески оправданными производственными решениями. В концептуальном плане данное явление означает конвергенцию развития компьютерных архитектур и структур алгоритмов, но оно еще имеет и философский подтекст в том смысле, что фактически машины уже начали создавать себе подобных.

Предсказать, как конкретная задача будет реально выполняться на МВС, заранее практически невозможно, однако существуют системные инструментарии, которые позволяют отслеживать и анализировать загрузку компьютерных устройств во время работы прикладной программы. Отметим еще, что компьютеры содержат обязательно и дисковую память (очень большую, но слишком медленную), которая используется в процессе решения задач только при исключительном дефиците оператив-

ной памяти, поскольку это катастрофически снижает производительность расчетов (для такого феномена уже имеется “русское” слово своппинг, от английского *swapping*).

В целях распараллеливания вычислительных процессов пользователю предоставляется программная система передачи сообщений MPI (Message Passing Interface), обслуживающая МВС с распределенной памятью и содержащая командный язык для информационных обменов между процессорами. Реализация расчетных схем в многопроцессорных компьютерах с общей памятью осуществляется с помощью программных средств типа OpenMP, принцип действия которых несколько другой. Если информационная структура алгоритмического фрагмента позволяет независимое исполнение в разных вычислительных потоках (*threads*), то они явно описываются программистом, и при этом каждый поток обрабатывается на своем процессоре без коммуникационных потерь (исключение составляют так называемые конфликты памяти, когда разные потоки обращаются к одинаковым данным). Кроме того на сверхбыстрых регистрах ядер с помощью их командных директив типа AVX можно организовывать векторизацию вычислений с детальным учетом особенностей доступа к памяти нижнего уровня.

Таким образом, продвинутое программирование больших задач основано на гибридной технологии, в которой формируется до четырех уровней распараллеливания. В данном случае пользователь должен сам провести модульный анализ алгоритмов и указать, какие их фрагменты нужно выполнять с помощью потоковых вычислений на многопроцессорных узлах с общей памятью, а какие чередуются с межузловыми обменами средствами MPI. Как видно из краткого изложения принципов распараллеливания алгоритмов, это занятие не для слабонервных и требует достаточно высокой специальной подготовки.

Качество распараллеливания алгоритмов характеризуется коэффициентом ускорения S , который означает, во сколько раз некоторая задача выполняется на P процессорах быстрее, чем на одном процессоре:

$$S_P = T_1(A)/T_P(A), \quad E_P = S_P/P.$$

Здесь через $T_P(A)$ обозначено время решения задачи A на P процессорах, а E_P означает эффективность использования многопроцессорной вычислительной системы.

Интуитивно напрашивается, что в лучшем случае величина S_P должна равняться P , а $E_P \approx 1$, и такое ускорение называется линейным. Однако на практике это оказывается недостижимым идеалом из-за коммуникационных потерь. Процессоры не могут все без остановки работать, так как для выполнения арифметических действий в каждом из них периодически приходится тратить драгоценное время на пересылку необходимых ему данных из других процессоров. Общее время решения задачи T складывается из времени выполнения арифметических операций T_a и времени T_c межпроцессорных информационных обменов, во время которых сами процессоры вынуждены простаивать. Решение большой задачи состоит из последовательного выполнения отдельных вычислительных этапов, для реализации каждого из которых обычно существует выбор из нескольких возможных алгоритмов. Искусство математика и программиста при оптимальном распараллеливании численных методов на МВС заключается в построении такой цепочки, которая обеспечивала бы наименьшую величину T_c , в сравнении с T_a . Зачастую наиболее быстрые классические алгоритмы очень плохо распараллеливаются, и на смену им приходится изыскивать новые подходы, адаптированные к особенностям конкретных вычислительных машин.

Отметим, что если отношение объема последовательных вычислений к общему количеству распараллеливаемых операций равно Θ , то ускорение в данном случае выражается следующим законом Амдала:

$$S_P = T_1 / [\Theta T_1 + (1 - \Theta)T_1 / P] = P / [1 + \Theta(P - 1)].$$

Отсюда следует, в частности, что даже при $P \rightarrow \infty$ величина S_P остаётся равной примерно $1/\Theta$.

Чтобы понять качество распараллеливания вычислительного процесса, надо уметь оценить величины T_a и T_c , что само по себе является отнюдь не простой задачей. В первом приближении суммарная длительность исполнения арифметических действий T_a равна произведению их общего количества N_a на среднее время τ_a , затрачиваемое на одну операцию. Однако толк от такого усреднения — все равно что определять среднюю температуру по больнице. Например, сложение чисел выполняется во много раз быстрее, чем их деление. Кроме того, работу имеющихся в процессорах сумматоров и умножителей можно, в

принципе, запустить в режиме конвейера (великое изобретение Форда востребовано и здесь), что намного увеличивает их итоговую производительность. Для простейшей модели машинных операций можно записать

$$T = T_a + T_c, \quad T_a \approx \tau_a N_a, \quad T_c \approx \tau_0 + \tau_c N_c,$$

где N_c – количество пересылаемых чисел, а τ_0 и τ_c – длительность “настройки” одного обмена и время передачи одного числа, причем на практике справедливы соотношения $\tau_0 \gg \tau_c \gg \tau_a$.

Из приведенных оценок очевидным образом следует, что надо стремиться к построению алгоритмов, которые требуют относительно малых информационных межпроцессорных обменов. И здесь дело не только в замедлении вычислений: ситуация усугубляется еще тем, что коммуникационные операции являются наиболее энергозатратными. Последнее обстоятельство настолько серьезно, что это становится главной проблемой для создания экзафлопсного компьютера, который по пессимистическим оценкам потребует для своего функционирования электроэнергии около 100 мегаватт. Инженеры ставят себе цель снизить эту цифру до 20 мВт, но отсюда возникает неожиданный вывод и для математиков: надо переходить на методы высоких порядков точности, которые как раз за счет усложнения самих вычислений требуют для своей реализации значительно меньших объемов данных!

Следует заметить, что мы подразумеваем действия со стандартной двойной точностью, при которой машинное слово составляет 64 двоичных разряда (бита), а относительная ошибка представления числа составляет около 10^{-15} . Необходимо подчеркнуть, что использование простой (одинарной) арифметики с соответствующей погрешностью округлений порядка 10^{-7} при решении больших задач находится на грани допустимого риска: в некоторых случаях повезет, но для многих реальных ситуаций конечный результат может отличаться от истинного на сотни процентов. Вообще говоря, для очень “плохих задач” даже двойная точность может оказаться недостаточной. В этих случаях могут применяться существующие программные средства для выполнения арифметических операций с расширенной мантиссой, но мы на этих специальных вопросах останавливаться не будем.

Картина информационных процессов очень запутана из-за наличия многоуровневой памяти, а также в силу возможности совмещения обменов с выполнением арифметических вычислений. Например, может обнаружиться парадоксальный на первый взгляд факт, что решение задачи на 10 процессорах ускорится, в сравнении с одним, в 15 раз (это явление называется свехлинейным ускорением). Ларчик здесь открывается просто: на 10 процессорах все числовые данные задачи могут уместиться в сверхбыстрых кэшах, а на одном — нет, и в последнем случае будет затрачиваться значительное время на обращение к более медленной памяти. Предложенный читателю набор эмпирических фактов “из жизни ЭВМ” — это “верхушка айсберга” сложнейшей картины функционирования МВС. Трагедия заключается в том, что более или менее приличной доступной модели компьютерных вычислений не существует (сапожник без сапог), в результате чего исследователю, пытающемуся оптимизировать параллельную реализацию алгоритма или задачи, приходится действовать вслепую “методом тыка”.

Проблематика эффективного распараллеливания алгоритмов — это активнейшая область современной вычислительной математики. Одно из узких мест здесь — это решение алгебраических задач сверхвысоких порядков (до миллиардов уравнений), поскольку требуемые для них объемы памяти и арифметических операций растут нелинейно с увеличением числа степеней свободы.

Естественной тенденцией истории развития и оптимизации численных методов является усложнение их логической структуры. Если в “однопроцессорную” эру на это можно было не обращать серьезного внимания, то при реализации на МВС такие алгоритмы, как правило, значительно теряют в эффективности распараллеливания. Сложилась ходячая притча: “Плохие алгоритмы распараллеливаются хорошо, а хорошие плохо”. Яркой демонстрацией этого правила являются методы Монте-Карло, основанные на реализации независимых случайных блужданий. Другим историческим примером являются многолетние дискуссии о явных и неявных сеточных аппроксимациях нестационарных многомерных начально-краевых задач математической физики. Явные методы просты и почти идеально распараллеливаются, но являются условно устойчивыми и зачастую требуют слишком жестких ограничений на величину временного шага.

Неявные же алгоритмы абсолютно устойчивы, но более сложны в реализации, особенно на МВС.

В последние десятилетия главным орудием распараллеливания ресурсоемких многомерных краевых задач стали методы декомпозиции областей (Domain Decomposition), имеющие свою славную историю. Предложенный еще в 19-м веке альтернирующий метод Шварца долгие годы служил удобным средством теоретического исследования задач в сложных расчетных областях путем сведения их к последовательности решений более простых вспомогательных задач. При этом исходная область разбивалась на несколько пересекающихся подобластей (чаще всего на две), а искомое решение строилось с помощью специального итерационного процесса. С появлением МВС данный подход получил второе рождение: количество подобластей стало исчисляться сотнями и тысячами, вместо альтернирующего метода появился аддитивный, декомпозицию стали делать и без пересечения подобластей, а на внутренних границах смежных подобластей начали использовать различные интерфейсные условия. Идея распараллеливания оказалась прозрачной: на каждой итерации надо синхронно решать вспомогательные задачи в подобластях на “своих” процессорах, а затем между ними обмениваться необходимой информацией. Сейчас в этой области существует огромное сообщество с большим потоком алгоритмов, теорем, публикаций и научных конференций.

Компьютерный мир очень динамичен, и “моделлерам” приходится с трудом за ним поспевать. Уже входят в практику компьютерные многомашинные сетевые технологии (Grids), которые позволяют собрать в единое целое самые разношерстные вычислительные средства, в том числе удаленные на тысячи километров (многомашинный “зоопарк”). Здесь уже важно не столько само распараллеливание, сколько объединение распределенных ресурсов для решения одной суперзадачи. Другое грядущее революционное направление — программируемые логические интегральные схемы. Фактически это означает проектирование и создание вычислительной системы под конкретную задачу, что становится реальным благодаря разительным успехам кремниевых технологий в создании достаточно дешевых суперчипов. Хотя эти идеи и носились в воздухе более двух десятилетий, сейчас они делают вызов прикладным математикам и программистам, поскольку здесь возникает новая массовая про-

блема — трудоемкое микропрограммирование для сложнейшего конгломерата элементарных компьютерных устройств.

Ищите да обрящете

В последние десятилетия появляются работы общеметодического и даже философского характера, посвященных будущему математического моделирования, направлениям научных супервычислений и компьютерным технологиям пост-петафлопсных масштабов. Резюмируя их и предыдущие разделы введения, можно следующим образом констатировать ряд новых проблем, обусловленных достигнутым в 21-м веке прогрессом естественных наук и вычислительно-информационных систем.

- На повестку дня выходят сверхзадачи новых рубежей изучения окружающего мира, требующие для своего решения петавычислений и казавшиеся фантастическими еще десятилетие назад: молекулярная биология, геновая инженерия и медицина, нанотехнологии и наноматериалы, изменения климата, долгосрочный прогноз погоды и природные катастрофы, мониторинг окружающей среды, антропогенных воздействий и экологических процессов, геофизика, управление жизненным циклом производства, разработка новых видов вооружений и т.д. Путевой картой (road map) междисциплинарных исследований становятся разномасштабные явления микро- и макромира, обратные задачи идентификации параметрических моделей, оптимизации технологических процессов и проектирования. Крупномасштабные компьютерные эксперименты в тесном сочетании с теоретическими исследованиями и натурными экспериментами становятся основным орудием человеческого познания. А массовое внедрение моделирования призвано стать реальной производительной силой, определяющего рост эффективности труда и валового национального продукта. Можно сформулировать следующее прогностическое утверждение: если два социальных потрясения за последние десятилетия связаны с интернетом и мобильным телефоном, то третьим будет глобальное моделирование в облачных вычислениях.
- На фоне бурного увеличения мощностей суперкомпьютеров и развития их архитектур фактически происходит

резкое отставание производительности программистского труда. Глобализация прикладных проблем инициирует стратегические направления математического моделирования в формировании больших коллективных проектов с передовыми средствами объектно-ориентированного программирования и поддержки полного жизненного цикла программного продукта (UML – Unified Modeling Language, DCOM – Distributed Component Object Management, CCA – Common Component Architecture) – от постановки физико-математических моделей до бизнес-планирования и менеджмента. Важно подчеркнуть “две стороны медали”, обязательные для успеха таких разработок: их высокая наукоемкость невозможна без неукоснительной фундаментальной поддержки разносторонних специалистов, а практические требования высокой производительности обуславливают жесткую дисциплину и организацию труда, характерную отнюдь не для академических учреждений, а для профессиональных компаний, создающих промышленные программные продукты. Эффективное внедрение вычислительно-информационных технологий нереально без создания организационных инфраструктур и государственной политики поддержки научных инноваций, формирования целевых фондов, грантов и других форм необходимого финансирования передовых коллективов разработчиков средств наукоемкого моделирования.

- Стихийно или полустихийно, но в мире активно развивается рынок самого разного программного обеспечения для математического моделирования: свободно доступные или коммерческие библиотеки и пакеты прикладных программ (широко известные разработки NETLIB, MKL, ANSYS, FEMLAB, NASTRAN и множество других), инструментальные, графические и системные средства, – непрерывно модифицируемые и адаптируемые к новым компьютерным архитектурам и телекоммуникационным технологиям. Жизненный цикл крупных разработок составляет обычно несколько десятилетий, в течение которых они претерпевают развитие и модификации, насколько это позволяют изначальные рамки проекта. При этом неизбежно их

моральное старение и появление новых продуктов, в большей или меньшей степени использующих последние достижения активно прогрессирующей вычислительной математики. В этих условиях экстенсивного роста отнюдь не умаляется стратегическая роль фундаментальных исследований по всем аспектам методологии и практики вычислительного эксперимента, миссией которых является выявление новых тенденций и перспективных направлений математического моделирования.

На этом фоне возникают новые концепции ППО, прототипами которых являются уже известные системы OPEN FOAM и DUNE, представляющие собой интегрированные инструментальные окружения, непрерывно развиваемые различными группами разработчиков, ориентированные на массовых пользователей и адаптируемые к бурной эволюции вычислительно-информационных технологий и компьютерных архитектур.

Резюмируя общеметодологический анализ рассматриваемой предметной области, можно констатировать, что математическое моделирование (ММ) процессов и явлений — это междисциплинарная проблема с широчайшей сферой деятельности. По своей сути, ММ в широком смысле становится главным оружием познания окружающего мира, управления производствами и социальными отношениями. Известно, что успех развития любой отрасли определяется уровнем ее математизации (в современную эпоху сюда надо добавить — и компьютеризации). Изучаемые объекты — физические, химические, биологические или другой природы — должны в первую очередь быть описаны специалистом-экспертом в данной предметной области на своем научном языке, после чего задание на моделирование переформулируется в математической постановке, которая должна в общем случае пройти теоретическое исследование и обоснование. Понимание вопросов существования, единственности, корректности и других свойств решения задачи позволяет создать (а скорее всего — выбрать) вычислительную схему, для которой собирается программная модель для проведения вычислительного эксперимента на конкретной компьютерной конфигурации. Необходимо себе представить, что за обеспечением всех этих этапов, при жестком удовлетворении производственных требований, скрывается огромный объем работ по созданию

вычислительно-информационных технологий (ВИТ), обладающих не только современной скоростной производительностью, которая зачастую выдвигается на первый план, но и не менее важным высоким уровнем искусственного интеллекта, отвечающего за автоматизацию построения моделей и алгоритмов, а также за комфортность всевозможных интерфейсов для широкого круга конечных пользователей – инженеров, операторов, лаборантов и т.д. – с многообразной профессиональной подготовкой.

Массовое внедрение таких инноваций неизбежно приведет фактически к новым специальностям и дисциплинам, в которых неотъемлемой составляющей становятся вычислительно-информационные технологии. Безвозвратно в прошлое уходит образ физика-теоретика с ручкой, листом бумаги и логарифмической линейкой, бывшими главными атрибутами ученых до середины XX-го века. ВИТ и моделирование приносят с собой кардинальное изменение не только методов проведения фундаментальных и прикладных исследований, но также производственных и социальных отношений, основывающихся на новой парадигме компьютерной грамотности, провозглашенной академиком А.П.Ершовым около 50 лет назад.

Возрастание общепознавательной роли моделирования в наукоемких технологиях активно вызывает появление новых понятий: *simulation*, *data mining*, *deep knowledge*, *digital design*, – которые в русском языке пока не имеют своих сложившихся аналогов. Однако в целом имеющиеся тенденции свидетельствуют о формировании нового поколения науки о знаниях, разрабатывающей свои когнитивные технологии и онтологические принципы, и это развитие средств искусственного интеллекта имеет серьезное философско-методологическое значение. По одному образному выражению, мы находимся в переходном периоде от “палеоинформатики” к “неоинформатике”.

Реально уже наступившая эпоха реиндустриализации с интернетом вещей, цифровым проектированием, военными игровыми стратегиями, “умными” домами и городами, электронными правительствами, аддитивными технологиями, 3D-принтерами и грядущими технологическими прорывами требует своего глубокого осмысливания, но уже очевидно, что в мировом научно-техническом прогрессе математическому моделированию принадлежит ключевая роль.

Замечания по литературе к “Введению”

Представленное выше “Введение в предметную область” по своему смыслу охватывает все аспекты рассматриваемой в книге проблематики. В последующих главах, касающихся анализа непрерывных и дискретных математических моделей, обзор литературы дается непосредственно при изложении соответствующих разделов книги. А в данном пункте мы осветим кратко библиографические аспекты по темам, которые выходят за рамки настоящего тома.

Здесь в первую очередь отметим общеметодологическое направление, связанное с проблемами информатики в моделировании и с концепцией крупномасштабного вычислительного эксперимента как орудия получения фундаментальных и прикладных знаний. В алфавитном порядке, здесь можно перечислить работы Э. Дермана [63], А. А. Дородницына [65], А. П. Ершова и В. П. Ильина [70], [71], [80], Г. И. Марчука [126], А. А. Самарского и А. П. Михайлова [163], Н. Н. Яненко и А. Н. Коновалова [207], [208].

Большая самостоятельная тема — это вопросы геометрического моделирования, которые включают такие направления, как “вычислительная геометрия” и “геометрическая информатика”, на которых зиждутся все САПР-овские разработки, уже давно составляющие мировой востребованный промышленный сегмент. В данном случае следуют сначала назвать книги школы В. Л. Рвачева [95], [148], [170] по R -функциям, а также современные монографии М. Дельфора, Дж. Золессио [58] по дифференциальному исчислению геометрических форм и Ю. Базилевса с коллегами — по изogeометрическому анализу [94]. Среди изданий по САПР-овской тематике мы приводим следующие: Д. Я. Левин с соавторами [108], К. Ли [109], Ф. Премарата и М. Шеймос [147], Д. М. Ушаков [181], А. Фокс и М. Пратт [187] и П. Шенан с коллегами [200]. Здесь же мы упомянем книги В. П. Гергеля с соавторами [44] по методологии принятия решений и Г. С. Альтшуллера [4], которым была предпринята серьезная попытка создания “теории решения изобретательских задач”. Подчеркнем, что данный перечень не включает публикации по сеточной проблематике, библиография которой содержится в шестой главе.

Программистские аспекты моделирования (объектные технологии, средства распараллеливания алгоритмов, организация

проектов и т.д.) в библиографии представлены работами А. П. Ершова и В. П. Ильина [70], [71], [80], Дж. Малонея [117], В. Э. Мальпшкина и В. Д. Корнеева [118], Л. В. Массель с соавторами [129] и Н. Н. Яненко с А. Д. Рычковым [210]. Такая актуальная область, как искусственный интеллект, представлена нами интересной книгой А. Клеппе [89]. И наконец, по немаловажным проблемам наукоемкого менеджмента, или управления, мы рекомендуем книги Г. И. Марчука [126] и Г. Чесбро [197].

Глава 1

Операторные постановки краевых задач

Целью этой главы является краткое описание предварительных сведений о функциональных пространствах, линейных операторах, функционалах и классических операторных формулировках краевых задач, необходимых для последующего изложения материала.

Задачи математического моделирования делятся на две части — стационарные и нестационарные. И хотя формально все независимые переменные равноправны, с прикладной точки зрения их целесообразно сразу разделить на пространственные переменные (по числу которых мы разделяем задачи на одномерные, двумерные и трехмерные, обозначаемые соответственно как 1-D, 2-D и 3-D) и время t — если оно присутствует в рассматриваемой постановке. Как правило, в последнем случае задачи называются начально-краевыми, т.е. для искомых функций задаются начальные условия при $t = 0$. Хотя точкой отсчета может быть любой момент времени $t = t_0$, определение $t_0 = 0$ роли не играет, в силу принципа Галилея. Что касается пространственной размерности задач, которую далее обозначаем через d , то наиболее актуальным случаем является $d = 3$, что соответствует реальному пространству окружающего мира. Хотя следует заметить, что одно- и двумерные задачи зачастую представляют самостоятельный интерес, а уравнения с $d > 3$ тоже встречаются в насущных прикладных проблемах, например — в финансовой математике.

Если в исходных уравнениях присутствуют вторые или старшие производные по времени, то искомые функции могут быть заданы для разных моментов, но в подавляющем числе случаев рассматриваются *задачи Коши*, когда все временные условия даны при $t = 0$. Фактически всегда искомое решение представляет интерес в ограниченной области и по времени, и по пространству, т.е. $0 \leq t \leq t_e < \infty$, а для радиус-вектора точки имеем $|\vec{r}| \leq R < \infty$, хотя иногда условия могут рассматриваться и на бесконечности (как по времени, так и по пространству).

Хотя математический формализм допускает, а иногда делает даже более удобным, рассмотрение комплексных величин и

уравнений, мы ограничиваемся, за отдельными исключениями, только вещественными переменными, к которым при необходимости всегда можно перейти путем введения обозначений для вещественных и мнимых частей числа. Используемые в данной главе понятия и свойства математических объектов даются на языке функционального анализа, теории операторов, дифференциальных и интегральных уравнений. Однако изложение дается по возможности кратко и в замкнутой форме, что позволяет его рассматривать как необходимый экскурс в соответствующие научные направления, призванный никак не отпугнуть читателя, интересующегося математическим моделированием, а в какой-то степени войти в данную предметную область. Естественным моментом является то соображение, что для решения функциональных уравнений надо иметь “математический минимум” знаний об их характеристиках. Современная вычислительная математика бурно развивается и впитывает в себя новые достижения теоретической математики, примерами обзоров которых являются фундаментальные монографии В. И. Арнольда [5] и С. П. Новикова в соавторстве с И. А. Таймановым [135]. Во многих журнальных публикациях, да и некоторых книгах, изложение численных методов уже ведется на языке внешних алгебр, дифференциальных форм и теории групп, симплектических пространств и многообразий, однако этот уровень требует специальной подготовки читателя, и мы вынуждены следовать афоризму Козьмы Пруtkова “никто не обнимет необъятного”.

Данная глава построена следующим образом. В параграфе 1.1 вводятся функциональные пространства с основными определениями и теоремами (доказательства везде опускаются), без которых в строгом смысле невозможно говорить об операторах. В следующем параграфе даются классические операторные постановки абстрактных краевых задач, а в последнем разделе, наоборот, – приводятся примеры конкретных прикладных проблем, представляющих несомненный практический интерес для разнообразных приложений. Здесь присутствуют задачи тепло-массопереноса, фильтрации и гидрогазодинамики, упругости и электромагнетизма, интегральные уравнения теории потенциала и гамильтоновы системы уравнений. Более глубокое изучение присущих данным проблемам вопросов имеется в многочисленных книгах Г. И. Марчука [121]– [128] и А. А. Самарского с коллегами [156]– [165], а также А. Н. Тихонова [175]– [177],

О. М. Белоцерковского [20, 21], П. Н. Вабищевича [30]– [34], В. А. Галкина [41], С. К. Годунова с коллегами [46]– [49], В. М. Головизнина [50], Н. Н. Яненко [206]– [210], А. Н. Коновалова [90, 91], А. А. Петрова [144, 145], Р. П. Федоренко [182, 183], Д. тер Хаара [191], Б. Н. Четверушкина [198], Ю. И. Шокина с соавторами [195], [205], П. Монка [133] и других авторов, см. например, [1, 3, 29, 52, 53, 55, 67, 93]. Мы рассматриваем постановки только прямых задач математического моделирования и отсылаем заинтересованного читателя к содержательному обзору литературы по обратным задачам, имеющемуся в книге С. И. Кабанихина [86]. Актуальность данной проблематики в последние десятилетия чрезвычайно повышается, а основополагающие результаты здесь связаны со школами А. Н. Тихонова, В. В. Иванова и М. М. Лаврентьева.

В определенном смысле настоящая глава является кратким введением в курс математической физики, хотя здесь с равным основанием можно употреблять и такие распространенные наименования, как механика, математические методы в химии, вычислительная биология и др. А на языке такой новой дисциплины, как “математическое моделирование”, которой посвящена данная книга, первая глава содержит описание рассматриваемых далее непрерывных моделей в операторных формулировках функционального анализа. Среди многочисленных книг, посвященных изложению данного теоретического материала, можно выделить компактную монографию В. И. Лебедева [104] — хотя бы потому, что она написана математиком-вычислителем.

1.1. Основные функциональные пространства: определения и формулировки теорем

Пусть Ω — *ограниченная область* (открытое связное множество) в n -мерном вещественном евклидовом пространстве R^n , Γ и $\bar{\Omega} = \Omega \cup \Gamma$ — соответственно *граница* ($(n - 1)$ -мерное многообразие) и *замыкание области* Ω . Через $x = (x_1, \dots, x_n)$ будем обозначать элемент (точку) пространства R^n , где $x_i, i = 1, \dots, n$, — координаты точки (вещественные числа, множество которых обозначается как R) в какой-то координатной системе.

Функция $\chi_\Omega(x)$, равная единице при $x \in \Omega$ и нулю при $x \notin \Omega$, называется *характеристической функцией* множества Ω .

Функция $\varphi(x)$, $x \in R^n$, называется *финитной*, если она рав-

на нулю вне некоторого шара с радиусом r и центром x^0 , т. е. $|x - x^0| \equiv ((x_1 - x_1^0)^2 + \dots + (x_n - x_n^0)^2)^{1/2} < r$. Носителем (обозначение $\text{supp } \varphi$) финитной функции φ , непрерывной в R^n , называется замыкание множества тех точек, где $\varphi(x) \neq 0$.

Мы будем рассматривать только *измеримые функции*, т. е. совпадающие почти везде (за исключением, может быть, множества меры нуль) с пределом почти везде сходящейся последовательности кусочно-непрерывных функций. Соответственно, предполагаются *измеримыми* все рассматриваемые *множества*, т. е. имеющие измеримые характеристические функции.

Все используемые функции будем считать *интегрируемыми по Лебегу*, что предполагает возможность перехода к пределу под знаком интеграла (напомним, что для всякой кусочно-непрерывной финитной функции интегралы Римана и Лебега совпадают). Интеграл $\int_{\Omega} dx = \int \chi_{\Omega}(x) dx$ называется *мерой Лебега множества Ω* и обозначается через $\text{mes}(\Omega)$.

1.1.1. Метрические и нормированные пространства

Определение 1.1. Множество X называется метрическим пространством, если каждой паре его элементов x и y сопоставлено вещественное число $\rho_X(x, y)$, называемое расстоянием между элементами x, y , или метрикой пространства X , и удовлетворяющее следующим условиям:

- а) $\rho_X(x, y) \geq 0$, $\rho_X = 0$ тогда и только тогда, когда $x = y$ (аксиома тождества);
- б) $\rho_X(x, y) = \rho_X(y, x)$ (аксиома симметрии);
- в) $\rho_X(x, y) \leq \rho_X(x, z) + \rho_X(z, y) \quad \forall z \in X$ (аксиома треугольника).

Элементы метрического пространства называются также точками, а расстояние между ними обозначается через $\rho(x, y)$, если пространство X определяется по контексту. Множество $Y \subset X$ называется *ограниченным*, если найдутся такой элемент $x \in X$ и постоянная $c > 0$, что $\rho(x, y) < c$ для $\forall y \in Y$.

Если Y и Z — два множества метрического пространства X , то величина $\inf_{y \in Y, z \in Z} \rho(y, z)$ называется *расстоянием между Y, Z* .

Простейший пример метрического пространства есть R^1 — множество всех вещественных чисел, в котором расстояние определяется как $\rho(x, y) = |x - y|$ для всех $x, y \in R^1$.

Определение 1.2. Точка x метрического пространства X называется пределом последовательности точек $\{x_n\} \in X$ (пишется $x_n \rightarrow x$ или $x = \lim x_n$), если $\rho(x_n, x) \rightarrow 0$ при $n \rightarrow \infty$.

Если множество $Y \subset X$, то точка $x \in Y$ называется *внутренней* точкой множества Y , если $\exists \varepsilon > 0$ такое, что шар $D_\varepsilon(x) = \{y \in Y : \rho(y, x) < \varepsilon\} \subset Y$; точка $x \in X$ называется *предельной* точкой множества Y , если $\forall \varepsilon > 0 D_\varepsilon(x) \cap (Y/x) \neq \emptyset$; множество $\bar{Y}$, получаемое присоединением к Y всех его предельных точек, называется *замыканием* множества Y . Множество Y называется *замкнутым*, если $\bar{Y} = Y$, и *открытым*, если все его точки — внутренние.

Множество $Y \subset X$ называется *плотным* в множестве $Z \subset X$, если $\bar{Y} \supseteq Z$, и *всюду плотным* в пространстве X , если $\bar{Y} = X$.

Метрическое пространство называется *сепарабельным*, если в нем существует счетное всюду плотное множество.

Определение 1.3. Последовательность $\{x_n\} \in X$ называется сходящейся в себе, или фундаментальной (а также последовательностью Коши), если $\forall \varepsilon > 0 \exists n_0(\varepsilon)$ такое, что $\rho(x_n, x_m) < \varepsilon$ при $m, n \geq n_0(\varepsilon)$.

Нетрудно показать, что всякая фундаментальная последовательность ограничена и что если $x_n \rightarrow x$, то справедливо $\rho(x_n, x_m) \rightarrow 0$ при $m, n \rightarrow \infty$ (обратное утверждение, вообще говоря, неверно).

Определение 1.4. Метрическое пространство называется *полным*, если в нем всякая фундаментальная последовательность имеет предел, принадлежащий этому пространству.

Справедливо утверждение о том, что всякое неполное метрическое пространство X (в котором существуют фундаментальные последовательности, не имеющие предела в X) можно пополнить, добавив к нему пределы всех фундаментальных последовательностей из X . Это новое, расширенное, пространство

с той же метрикой $\rho(x, y)$ называется *пополнением пространства* X .

Например, пространство вещественных чисел R^1 есть пополнение множества рациональных чисел с той же метрикой (последнее не является полным, так как оно содержит последовательности, сходящиеся к иррациональному числу).

Второй пример: пространство X определенных на отрезке $[a, b]$ многочленов $P(x)$ с метрикой $\rho(P, P') = \max_{x \in [a, b]} |P(x) - P'(x)|$ не является полным. А так как X всюду плотно в пространстве непрерывных функций $C[a, b]$ (по теореме Вейерштрасса каждая непрерывная функция может быть сколь угодно точно приближена равномерно сходящейся последовательностью многочленов), то пополнение X есть пространство $C[a, b]$. Точнее говоря, пополнение X и $C[a, b]$ — это два изометрических пространства: два метрических пространства X и Y называются изометрическими, если для любой пары элементов $x, x' \in X$ найдется пара $y, y' \in Y$ такая, что между ними можно установить взаимно однозначное соответствие $\rho_X(x, x') = \rho_Y(y, y')$, в силу чего их можно считать метрически идентичными.

В математическом анализе имеется *теорема Больцано–Вейерштрасса* о том, что из всякой ограниченной бесконечной последовательности чисел можно выделить подпоследовательность, сходящуюся к конечному пределу. Это заключение справедливо для следующих множеств, лежащих в метрическом пространстве X .

Определение 1.5. Множество $M \subset X$ называется *бикompактным*, если из каждой последовательности $\{x_n\} \subset M$ можно выбрать сходящуюся подпоследовательность, предел которой принадлежит M . Это же множество называется *компактным*, если из каждой последовательности $\{x_n\} \subset M$ можно выделить фундаментальную последовательность.

Очевидно, что всякое *бикompактное множество* является замкнутым и компактным, а всякое *компактное множество* M ограничено (и является бикompактным тогда и только тогда, когда оно замкнуто).

Отметим, что метрические пространства являются частным случаем *топологических пространств*, в которых только из условий введенной топологии, т. е. системы окрестностей и откры-

тых множеств, исследуются такие рассмотренные выше топологические свойства, как плотность, полнота, сепарабельность и компактность.

Определение 1.6. Множество E элементов $(x, y, z, \dots)$ называется линейным пространством, если в нем определены операции сложения и умножения на число (скаляр линейного пространства) таким образом, чтобы выполнялись следующие свойства (аксиомы) для $\forall x, y, z \in E$ и любых скаляров λ, μ :

$$\begin{aligned} & x + y = y + x \in E \text{ (коммутативность сложения);} \\ & x + (y + z) = (x + y) + z \text{ (ассоциативность сложения);} \\ & \text{существует нулевой элемент } 0 \in E \text{ такой, что } x + 0 = x; \\ & \lambda(\mu x) = (\lambda\mu)x \text{ (ассоциативность умножения);} \\ & 1 \cdot x = x, 0 \cdot x = 0 \text{ (слева — 0-скаляр; справа — элемент } E); \\ & \left. \begin{aligned} & \lambda(x + y) = \lambda x + \lambda y \\ & (\lambda + \mu)x = \lambda x + \mu x \end{aligned} \right\} \text{ (дистрибутивность).} \end{aligned}$$

Линейные пространства иногда называются векторными, а их элементы — векторами.

Линейное пространство называется *вещественным* или *комплексным*, если в качестве скаляров $\lambda, \mu, \dots$ берутся вещественные или комплексные числа соответственно.

В линейном пространстве E нулевой элемент единственен и можно определить противоположный элемент $-x$, а также операцию вычитания по следующим правилам: $-x = (-1)x$, $x + (-x) = 0$, $x - y = x + (-y)$.

Множество $E_0 \subset E$ называется *подпространством* линейного пространства E , если при любых $x, y \in E_0$ для любых скаляров λ, μ справедливо $\lambda x + \mu y \in E_0$. Конечная система ненулевых элементов $x_1, \dots, x_n \in E_n$ называется *линейно независимой*, если их линейная комбинация $\lambda_1 x_1 + \dots + \lambda_n x_n$ может обращаться в нуль только при $\lambda_1 = \dots = \lambda_n = 0$. Бесконечная система элементов $\{x_k\}_1^\infty$ в линейном пространстве E называется *линейно независимой*, если любая ее конечная подсистема $x_{k_1}, \dots, x_{k_n}$ линейно независима. Совокупность всевозможных линейных комбинаций $\sum_{i=1}^n \lambda_i x_i$, где $x_i \in M \subset E$, а λ_i — произвольные скаляры, называется *линейной оболочкой* множества M и обозначается $\text{span}(M)$ или $\text{span}(x_1, \dots, x_n)$.

Если x, y — два элемента пространства E , то совокупность всех элементов $\alpha x + (1 - \alpha)y$, где $0 \leq \alpha \leq 1$, называется *отрезком*.

Множество $M \subset E$ называется *выпуклым*, если оно целиком содержит отрезок, определяемый любыми двумя его элементами.

Примером выпуклого множества является *шар* радиуса $r > 0$ с центром в точке x_0 , определяемый как совокупность элементов $x \in E$, удовлетворяющий строгому неравенству $\rho(x, x_0) < r$, или *замкнутый шар*, характеризуемый условием $\rho(x, x_0) \leq r$. Однако такое множество, как сфера ($\rho(x, x_0) = r$), уже не является выпуклым.

Если в линейном пространстве E существует конечная система n линейно независимых элементов, линейная оболочка которой совпадает с E , то данное пространство называется n -мерным (конечномерным), а всякая линейно независимая система из n элементов называется *базисом* пространства E . Линейное пространство, не являющееся конечномерным, называется бесконечномерным. Любой элемент x из пространства E единственным образом представляется в виде линейной комбинации базисных элементов x_i .

Определение 1.7. Линейное пространство X называется *нормированным*, если на множестве его элементов x определена вещественная функция, называемая *нормой* и обозначаемая $\|x\|_X$, или $\|x\|$, удовлетворяющая следующим четырем свойствам (аксиомам нормированного пространства) для $\forall x, y \in X$ и любого скаляра λ :

$$\begin{aligned} \|x\| &\geq 0, & \|\lambda x\| &= |\lambda| \|x\|, & \|x + y\| &\leq \|x\| + \|y\|, \\ & & \text{если } \|x\| &= 0, & \text{то } x &= 0. \end{aligned}$$

Здесь второе соотношение называется свойством однородности нормы, а третье — неравенством треугольника.

Если на элементах $x \in X$ определена вещественная функция $|x|$, удовлетворяющая только первым трем свойствам, то пространство X называется *полунормированным*, а функция $|x|$ — *полунормой*.

На одном и том же множестве линейного пространства могут быть определены различные нормы. Например, в конечномерном пространстве R^n n -мерных векторов $x = (x_1, \dots, x_n)$ употребляются векторные нормы

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}, \quad \|x\|_\infty = \max_i |x_i|,$$

называемые соответственно *октаэдрической*, *сферической* и *кубической*. Две нормы $\|x\|_1, \|x\|_2$ в нормированном пространстве X называются *эквивалентными*, если существуют такие положительные постоянные c_1 и c_2 , что для любого $x \in X$ выполняется равенство

$$c_1 \|x\|_1 \leq \|x\|_2 \leq c_2 \|x\|_1.$$

В конечномерном пространстве R^n все нормы, в том числе и приведенные выше, являются эквивалентными, хотя константы эквивалентности могут зависеть от размерности n . Если последовательность $\{x_n\} \in X$ сходится в какой-то норме, то она также сходится в любой другой эквивалентной норме.

Формальным обобщением конечномерного пространства R^n являются введенные Гильбертом пространства числовых последовательностей $x = (x_1, x_2, \dots, x_n, \dots)$, для которых $\|x\|_p = \sum_{i=1}^{\infty} |x_i|^p < \infty$, где $p = 1$ или $p = 2$. Соответствующие пространства обозначаются через l_1 или l_2 . Такие пространства являются сепарабельными. Однако множество m , состоящее из всевозможных ограниченных последовательностей $x = \{x_1, x_2, \dots, x_n, \dots\}$, после введения нормы $\|x\| = \sup_i \{|x_i|\}$ становится линейным нормированным пространством, не являющимся сепарабельным.

Определение 1.8. Линейное нормированное пространство называется банаховым, если оно полное.

Именно такие пространства, зачастую обозначаемые буквой B и называемые B -пространством, будут нами рассматриваться в дальнейшем.

Любое конечномерное линейное нормированное пространство полно и, следовательно, является банаховым.

В банаховом пространстве определены линейные комбинации конечного числа элементов $\{x_i\}_1^n$, а также бесконечные ряды

$$\sum_{k=1}^{\infty} x_k = x_1 + \dots + x_n + \dots$$

Этот ряд называется *сходящимся*, если сходится последовательность его частичных сумм $\{S_n = \sum_{k=1}^n x_k\}$, а предел этой по-

следовательности называется *суммой* ряда. Говорят, что ряд *абсолютно сходится*, если сходится ряд

$$\sum_{k=1}^n \|x_k\| = \|x_1\| + \dots + \|x_n\| + \dots$$

В пространстве B из абсолютной сходимости ряда вытекает обычная сходимость.

В конечномерном банаховом пространстве всякое ограниченное множество замкнуто и компактно. Верно в некотором смысле и обратное утверждение: линейное нормированное пространство, в котором какой-либо замкнутый шар компактен, является конечномерным.

Полезным является следующее утверждение (критерий компактности): для компактности множества M метрического пространства B необходимо, а в случае полноты B и достаточно, чтобы для любого $\varepsilon > 0$ существовала конечная ε -сеть, т. е. такое конечное множество элементов $x_i, \dots, x_n \in B$, что для каждого элемента $x \in M$ найдется элемент сети x_k , находящийся от x на расстоянии меньше ε : $\rho(x, x_k) < \varepsilon$.

Таким образом, в линейном нормированном пространстве компактность множества означает, что его можно с любой точностью приблизить множеством из конечномерного пространства.

Определение 1.9. Система элементов $\{x_\alpha\} \in B$, $\alpha \in U$ (U — некоторое счетное множество индексов) называется *полной*, если ее линейная оболочка образует всюду плотное в B множество. Последовательность элементов $\{e_k\} \in B$ называется *базисом*, если для любого $x \in B$ существует, и при том единственная, последовательность чисел $\{\lambda_k\}$, $k = 1, 2, \dots$, такая, что имеет место разложение x по базису, т. е. $x = \sum_{k=1}^{\infty} \lambda_k e_k$.

Определение 1.10. Вещественное линейное пространство E называется *евклидовым*, если каждой паре его элементов x, y поставлено в соответствие вещественное число (x, y) , называемое скалярным произведением и удовлетворяющее следующим четырем аксиомам:

- $(x, x) \geq 0$ и $(x, x) = 0$ тогда и только тогда, когда $x = 0$;
- $(x, y) = (y, x)$;

- $(\lambda x, y) = \lambda(x, y)$ для любого вещественного числа λ ;
- $(x + y, z) = (x, z) + (y, z)$.

Евклидовое пространство становится нормированным, если в нем определить норму (евклидовую) по формуле $\|x\|_E = (x, x)^{1/2}$. Для элементов $x, y \in E$ справедливо *неравенство Коши–Буняковского*

$$|(x, y)| \leq \|x\| \cdot \|y\|,$$

которое иногда также называется *неравенством Шварца*.

Если же скалярное произведение определить как комплекснозначное число (x, y) , удовлетворяющее первой и последней вышеприведенным аксиомам, а также третьей аксиоме для любого комплексного числа λ , но с заменой второй аксиомы на свойство $(x, y) = \overline{(y, x)}$ (черта означает переход к комплексно-сопряженному числу), то получаемое комплексное линейное пространство называется *унитарным*. Для него справедливы определение нормы и неравенство Коши–Буняковского, как для евклидова пространства.

Определение 1.11. Полное евклидовое или полное унитарное пространство называется *гильбертовым пространством*.

Иногда неполное пространство, в котором определены скалярное произведение и соответствующая норма, называют *предгильбертовым* или пространством с гильбертовой структурой. Такое пространство изометрично гильбертовому, причем предгильбертовое пространство всегда можно пополнить, превращая его в гильбертовое (полное) пространство.

Гильбертовое пространство принято обозначать буквой H . Для последовательности $\{x_n\} \in H$ сходимость по норме $x_n \rightarrow x$, т. е. $\|x_n - x\| \rightarrow 0$ при $n \rightarrow \infty$ называется *сильной сходимостью*. Наряду с ней в гильбертовом пространстве определяется еще один вид сходимости: последовательность $\{x_n\} \in H$ называется *слабо сходящейся* к элементу $x \in H$, если $\lim_{n \rightarrow \infty} (x_n, f) = (x, f)$ для $\forall f \in H$. Для гильбертовых пространств нетрудно показать следующие свойства:

- последовательность $\{x_n\} \in H$ не может слабо сходиться к двум различным элементам из H ;

- если последовательность сходится сильно к $x \in H$, то она слабо сходится к x ;
- если последовательность слабо сходится, то она ограничена;
- для $\forall x, y \in H$ имеет место равенства параллелограмма

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2);$$

- гильбертово пространство является *строго нормированным*, т. е. равенство $\|x + y\| = \|x\| + \|y\|$ достигается тогда и только тогда, когда $x = \lambda y$, $\lambda > 0$.

Важными понятиями гильбертовых пространств являются соотношения ортогональности и проектирования. Два ненулевых элемента гильбертового пространства $x, y \in H$ называются *ортогональными* ($x \perp y$) между собой, если $(x, y) = 0$. Если, кроме того, $\|x\| = \|y\| = 1$, то они называются *ортонормированными*. Два множества H' и H'' из H называются ортогональными ($H' \perp H''$), если $(h', h'') = 0$ для всех $h' \in H'$, $h'' \in H''$.

Теорема 1.1 (о минимальном расстоянии). Пусть M — замкнутое выпуклое множество в H и $x \notin M$. Тогда существует единственный элемент $y \in M$, реализующий расстояние от точки x до множества M :

$$\rho(x, M) \equiv \inf_{u \in M} \|x - u\| = \|x - y\|,$$

причем $(x - y) \perp M$, если M — подпространство H .

Указанный элемент $y \in M$ называют также *наилучшим приближением*, или *проекцией* на M элемента $x \notin M$. Справедливо следующее утверждение: точка y будет проекцией точки x на выпуклое множество M в том и только том случае, если для $\forall z \in M$ справедливо $(z - y, x - y) \leq 0$.

Теорема 1.2. Пусть L есть подпространство в H , тогда для $\forall x \in H$ и $\forall y \in L$, таких что $\rho(x, L) = \|x - y\|$, существует однозначное представление $x = y + z$, где $z \in L^\perp \perp L$, а $L^\perp$ является подпространством в H и называется ортогональным дополнением к L . Элементы y, z называются ортогональными проекциями x на подпространства $L, L^\perp$ соответственно, и для них справедлива теорема Пифагора

$$\|x\|^2 = \|y\|^2 + \|z\|^2.$$

Данный результат эквивалентен следующему: для $\forall x \in H$ справедливо единственное разложение $x = y + z$, где $y \in L$, $z \perp L$. Другими словами, гильбертово пространство раскладывается в сумму ортогональных подпространств L и $L^\perp$, причем для каждого L его ортогональное дополнение $L^\perp$ определяется единственным образом.

Можно также рассматривать ортогональные (прямые) суммы произвольных гильбертовых пространств, не являющихся подпространствами некоторого заранее заданного гильбертового пространства. Например, ортогональная сумма пространств H_1 и H_2 есть гильбертово пространство $H = H_1 \oplus H_2$, элементами которого являются всевозможные пары $\{u, v\}$, где $u \in H_1$, $v \in H_2$ причем

$$\alpha\{u, v\} = \{\alpha u, \alpha v\}, \quad \{u_1 + u_2, v_1 + v_2\} = \{u_1, v_1\} + \{u_2, v_2\},$$

а скалярное произведение определяется равенством

$$(\{u_1, v_1\}, \{u_2, v_2\}) = (u_1, u_2) + (v_1, v_2).$$

Отметим, что в гильбертовом пространстве конечная последовательность элементов $\{x_i\}_{i=1}^N$ является линейно независимой тогда и только тогда, когда отличен от нуля *определитель матрицы Грама*

$$|G| \equiv |G(x_1, x_2, \dots, x_n)| = \begin{vmatrix} (x_1, x_1) & (x_2, x_1) & \dots & (x_n, x_1) \\ (x_1, x_2) & (x_2, x_2) & \dots & (x_n, x_2) \\ \dots & \dots & \dots & \dots \\ (x_1, x_n) & (x_2, x_n) & \dots & (x_n, x_n) \end{vmatrix}.$$

Пусть L есть n -мерное подпространство гильбертового пространства H , содержащее какие-то n линейно независимых векторов $g_1, g_2, \dots, g_n$. Так как между любыми $n + 1$ элементами подпространства L имеется линейная зависимость, то всякий вектор $g \in L$ представим в виде линейной комбинации $g_1, g_2, \dots, g_n$, т. е. L является линейной оболочкой этих векторов.

Возьмем произвольный элемент $h \in H$ и обозначим через

$$g = \lambda_1 g_1 + \lambda_2 g_2 + \dots + \lambda_n g_n$$

его проекцию на L ($\lambda_1, \lambda_2, \dots, \lambda_n$ — некоторые числа). По свойству проекции разность $z = h - g$ должна быть ортогональна подпространству L , что означает выполнение системы равенств

$$(z, g_k) = (h, g_k) - \lambda_1(g_1, g_k) - \dots - \lambda_n(g_n, g_k) = 0, \quad k = 1, \dots, n,$$

детерминант которой есть определитель Грама $G(g_1, \dots, g_n)$.

С его помощью легко выражается расстояние $\rho(h, L)$ от точки h до подпространства L :

$$\begin{aligned} \rho^2(h, L) &= \|z\|^2 = \|h - g\|^2 = (z, h) \\ &= (h, h) - \lambda_1(g_1, h) - \lambda_2(g_2, h) - \dots - \lambda_n(g_n, h). \end{aligned}$$

Отсюда, после исключения с помощью предыдущей системы величин λ_k , приходим к формуле

$$\rho^2(h, L) = G(h, g_1, g_2, \dots, g_n) / G(g_1, g_2, \dots, g_n).$$

Очевидно, что величина $\rho(L)$ может интерпретироваться как норма вектора ошибки аппроксимации элемента $h \in H$ элементами из подпространства L .

Так как $G(g_1) = (g_1, g_1) > 0$ при $g_1 \neq 0$, то из последнего соотношения по индукции следует, что определитель Грама линейно независимых векторов всегда положителен. Данный факт можно рассматривать как обобщение неравенства Коши–Буняковского, которое в рассматриваемом контексте утверждает, что для линейно независимых векторов g_1, g_2 имеет место неравенство $G(g_1, g_2) > 0$.

Две последовательности $\{f_i\}_1^\infty, \{g_i\}_1^\infty$ из гильбертового пространства H называются *биортогональными*, если выполняются равенства

$$(f_i, g_j) = 0 \quad \text{для } i \neq j, \quad (f_i, g_i) \neq 0.$$

Не для всякой системы линейно независимых элементов найдется биортогональная последовательность. Справедливо следующее утверждение: последовательность элементов $\{g_i\}_1^\infty$ имеет биортогональную к ней последовательность в том и только в том случае, когда ни при каком натуральном j элемент g_j не принадлежит замкнутой линейной оболочке остальных элементов $g_1, \dots, g_{j-1}, g_{j+1}, \dots$. Отсюда, в частности, следует, что для конечномерной системы линейно независимых векторов всегда можно построить биортогональную последовательность.

1.1.2. Пространства функций $L_p(\Omega)$ и $C^p(\Omega)$

Рассмотрим теперь примеры функциональных пространств, на которых основан аппарат исследования и обоснования численных методов решения задач математического моделирования.

Через $L_p(\Omega)$ обозначаются *пространства Лебега* функций, имеющих ограниченные соответствующие нормы $\|\cdot\|_{L_p(\Omega)}$, которые определяются как

$$\|u\|_{L_p(\Omega)} = \left(\int_{\Omega} |u(x)|^p dx \right)^{1/p} \quad (1.1)$$

при $1 \leq p < \infty$ и

$$\|u\|_{L_{\infty}(\Omega)} = \operatorname{ess\,sup} |u(x)| : x \in \Omega. \quad (1.2)$$

Напомним, что $\operatorname{ess\,sup} |u(x)|$ (другое обозначение $\operatorname{vrai\,sup} |u(x)|$) есть инфимум чисел $\alpha \in \mathbb{R}$ таких, что неравенство $|u(x)| > \alpha$ для $x \in \Omega$ выполняется только на множестве меры нуль.

Важно отметить, что в (1.1) и везде далее используется обозначение интеграла Лебега (а не Римана), обладающего свойством «полноты»: предел интегрируемых функций есть интегрируемая функция. Именно в силу этого пространства Лебега $L_p(\Omega)$ при $1 \leq p \leq \infty$ оказываются банаховыми.

Частным случаем пространств Лебега является $L_2(\Omega)$ — гильбертово пространство измеримых функций, суммируемых с квадратом, в котором определены скалярное произведение

$$(u, v) = \int_{\Omega} u(x)v(x) dx$$

и норма

$$\|u\|_{L_2} = [(u, u)]^{1/2} = \left[\int_{\Omega} u^2(x) dx \right]^{1/2}.$$

Пусть $\alpha = (\alpha_1, \dots, \alpha_n)$ — целочисленный вектор (мульти-индекс) и $|\alpha| = \alpha_1 + \dots + \alpha_n$. Через $D^{\alpha}f(x)$ обозначаем производную функции $f(x)$ порядка $|\alpha|$ (в классическом «поточечном» смысле):

$$D^\alpha f = \frac{\partial^{|\alpha|} f(x_1, \dots, x_n)}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}, \quad D^0 f(x) = f(x).$$

Для любых функций $u \in L_p(\Omega)$ и $v \in L_q(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$, $p \geq 1$, справедливо *неравенство Гельдера*

$$\left| \int_{\Omega} uv \, dx \right| \leq \left(\int_{\Omega} |u|^p \, dx \right)^{1/p} \left(\int_{\Omega} |v|^q \, dx \right)^{1/q},$$

частным случаем которого при $p = q = 2$ является рассмотренное в п. 1.1.1 неравенство Коши–Буняковского.

Через $C^p(\Omega)$ обозначаем пространство функций, непрерывных вместе с производными порядка $|\alpha| \leq p$ ($0 \leq p < \infty$) в области Ω , и через $C^p(\bar{\Omega})$ — класс функции $u(x)$ из $C^p(\Omega)$, у которых все производные $D^\alpha u$, $|\alpha| \leq p$, допускают непрерывное продолжение на замыкание $\bar{\Omega}$ ($C(\Omega) = C^0(\Omega)$ и $C(\bar{\Omega}) = C^0(\bar{\Omega})$ — классы функций, непрерывных в Ω и $\bar{\Omega}$ соответственно). Данные пространства являются сепарабельными банаховыми, т. е. линейными, полными и нормированными в соответствии с определением

$$\|u\|_{C^p(\Omega)} = \max_{|\alpha| \leq p} \sup_{x \in \Omega} \{ |D^\alpha u(x)| \}. \quad (1.3)$$

1.1.3. Линейные функционалы и билинейные формы

Рассмотрим сначала следующие два определения.

Определение 1.12. Линейный функционал на линейном пространстве B есть линейная функция, действующая из B в пространство $R = R^1$, т. е. функция $L : B \rightarrow R$ такая, что

$$L(u + av) = L(u) + aL(v) \quad \forall u, v \in B, \quad a \in R.$$

Линейный функционал $L(u)$ называется *непрерывным* в точке $u_0 \in B$, если $L(u_n) \rightarrow L(u_0)$ при $u_n \rightarrow u_0$. Из непрерывности линейного функционала в одной точке следует его непрерывность всюду в линейном нормированном пространстве B .

Теорема 1.3. Линейный функционал L непрерывен на банаховом пространстве B , если и только если он ограничен, т. е. существует такая константа C , что $|L(v)| \leq C\|v\|_B \quad \forall v \in B$.

Определение 1.13. Множество непрерывных линейных функционалов на B называется пространством B' , двойственным к банаховому пространству B .

Для непрерывного линейного функционала определена конечная величина

$$\|L\|_{B'} = \sup_{0 \neq v \in B} \left\{ \frac{|L(v)|}{\|v\|_B} \right\}, \quad (1.4)$$

называемая *нормой* в двойственном пространстве B' . Пространство B' — полное относительно данной нормы, т. е. является банаховым.

Теорема 1.4 (Рисса о представлении функционала). *Любой непрерывный линейный функционал L на гильбертовом пространстве H единственным образом представим в виде скалярного произведения*

$$L(v) = (u, v)$$

для некоторого $u \in H$. Более того, в двойственном пространстве H' определена норма

$$\|L\|_{H'} = \|u\|_H.$$

Определение 1.14. Билинейная форма $a(\cdot, \cdot)$ на линейном пространстве H есть отображение $a : H \times H \rightarrow R$ такое, что отображения $u \rightarrow a(u, v)$ и $v \rightarrow a(u, v)$ суть линейные формы (функционалы) на H . Билинейная форма называется симметричной, если $a(u, v) = a(v, u)$ для всех $u, v \in H$.

В силу своего определения билинейная форма в гильбертовом пространстве удовлетворяет всем аксиомам скалярного произведения.

Определение 1.15. Билинейная форма $a(\cdot, \cdot)$ на нормированном линейном пространстве H называется ограниченной (непрерывной), если $\exists c_1 < \infty$ такое, что

$$|a(v, w)| \leq c_1 \|v\|_H \|w\|_H \quad \forall v, w \in H, \quad (1.5)$$

и V -эллиптической (коэрцитивной) на $V \cap H$, если $\exists c_2 > 0$ такое, что для соответствующей квадратичной формы справедливо неравенство

$$a(v, v) \geq c_2 \|v\|_H^2 \quad \forall v \in V. \quad (1.6)$$

Симметричная коэрцитивная билинейная форма называется *положительно определенной* (с.п.о.).

1.1.4. Обобщенные производные и пространства Соболева

Определение 1.16. *Пространством основных функций* $\mathcal{D}(\Omega)$ называется множество $C_0^\infty(\Omega)$ финитных, имеющих в Ω компактный носитель и бесконечно дифференцируемых в области Ω функций, снабженное топологией счетного числа норм

$$\|\varphi\|_{C^p(\bar{\Omega})} = \max_{|\alpha| \leq p} \sup_{x \in \bar{\Omega}} \{|D^\alpha \varphi(x)|\}, \quad p = 0, 1, \dots$$

Отметим, что введенное таким образом пространство не является, вообще говоря, банаховым.

Определение 1.17. *Обобщенной функцией* называется всякий линейный непрерывный функционал на пространстве основных функций $\mathcal{D}(\Omega)$. Линейное множество $\mathcal{D}'(\Omega)$ является полным пространством.

Определение 1.18. Функция $u(x)$ называется *локально интегрируемой* в области Ω , если $u \in L_1(K)$ для всякого компактного множества $K \subset \Omega$.

Определение 1.19. Пусть функция $u(x)$ локально интегрируема в области Ω . Пусть также существует локально интегрируемая в Ω функция $g(x)$ такая, что справедливо равенство

$$\int_{\Omega} g(x) \varphi(x) dx = (-1)^{|\alpha|} \int_{\Omega} u(x) D^\alpha \varphi(x) dx \quad \forall \varphi \in \mathcal{D}(\Omega).$$

Тогда $g(x)$ называется *обобщенной* (слабой) *производной* функции $u(x)$ и обозначается через $D^\alpha u(x)$.

Заметим, что мы не делаем отличия в обозначениях классических и обобщенных производных, поскольку будем иметь дело главным образом только с последними, а при необходимости их различие определяется по контексту.

Определение 1.20. Пространством Соболева $W_p^k(\Omega)$ называется множество локально интегрируемых в Ω функций, для которых существуют обобщенные производные D^α и при всех $|\alpha| \leq k$ ограничена норма

$$\|u\|_{W_p^k(\Omega)} = \left(\sum_{|\alpha| \leq k} \|D^\alpha u\|_{L_p(\Omega)}^p \right)^{1/p} \quad \text{для } 1 \leq p < \infty, \quad (1.7)$$

для $p = \infty$ определяемая как

$$\|u\|_{W_\infty^k(\Omega)} = \max_{|\alpha| \leq k} \{\|D^\alpha u\|_{L^\infty}\}. \quad (1.8)$$

Пространства Соболева являются банаховыми и включают различные известные пространства как частные случаи. Так, в одномерном случае функции из $W_1^1(\Omega)$ называются *абсолютно непрерывными на интервале Ω* . Для $k > 1$ можно также определить

$$W_\infty^k(\Omega) = \{u \in C^{k-1}(\Omega) : D^\alpha u \in \text{Lip}(\Omega) \forall |\alpha| \leq k-1\},$$

где $\text{Lip}(\Omega)$ — пространство *липшицевых функций*:

$$\text{Lip}(\Omega) = \{u \in L^\infty(\Omega) : \|u\|_{\text{Lip}(\Omega)} < \infty\}.$$

Здесь используется *липшицева норма*

$$\|u\|_{\text{Lip}(\Omega)} = \|u\|_{L^\infty(\Omega)} + \sup \left\{ \left| \frac{u(x) - u(y)}{x - y} \right| : x, y \in \Omega; x \neq y \right\}. \quad (1.9)$$

Существует другое определение соболевских пространств, которое может быть сделано следующим образом. Пусть $H_p^k(\Omega)$ означает замыкание $C^k(\Omega)$ по соболевской норме $\|\cdot\|_{W_p^k(\Omega)}$. В случае $p = \infty$ мы имеем $H_\infty^k = C^k$, и это не то же самое, что пространство $W_\infty^k(\Omega)$ (последнее относится к пространству липшицевых функций). Однако при $1 \leq p < \infty$ оказывается, что $H_p^k(\Omega) = W_p^k(\Omega)$, и этот факт может быть сформулирован следующим образом.

Теорема 1.5. Пусть Ω есть любое открытое множество. Тогда $C^\infty(\Omega) \cap W_p^k(\Omega)$ плотно в $W_p^k(\Omega)$.

Важным частным случаем является семейство пространств $H^m(\Omega)$ — пространства Соболева порядков $m \geq 0$, состоящие из функций $u \in L_2(\Omega)$, все частные обобщенные производные $D^\alpha u$ которых с $|\alpha| \leq m$ также принадлежат $L_2(\Omega)$; $H^m(\Omega)$ есть гильбертово пространство со скалярным произведением

$$(u, v)_{H^m} = \sum_{|\alpha| \leq m} \int_{\Omega} D^\alpha u D^\alpha v \, dx \quad (1.10)$$

и нормой

$$\|u\|_{H^m} = (u, u)_{H^m}^{1/2} = \left(\sum_{|\alpha| \leq m} \int_{\Omega} |D^\alpha u|^2 \, dx \right)^{1/2}. \quad (1.11)$$

Отметим, что в приведенных скалярных произведениях и нормах, строго говоря, надо еще указывать индекс « Ω », но мы его опускаем для краткости, так как это очевидно из контекста.

1.1.5. Липшицевы границы и теоремы вложения

Дадим несколько определений.

Определение 1.21. Будем говорить, что поверхность Γ принадлежит классу C^p , $p \geq 1$, если в некоторой окрестности каждой точки $x^0 \in \Gamma$ она представляется уравнением $\omega_{x^0}(x) = 0$, причем $\text{grad } \omega_{x^0}(x) \neq 0$, а функция $\omega_{x^0}(x)$ в этой окрестности непрерывна со всеми производными до порядка p включительно. Граница Γ области Ω называется *кусочно-гладкой*, если она состоит из конечного числа поверхностей класса C^1 .

Определение 1.22. Будем говорить, что область Ω имеет липшицеву границу Γ , если существуют $\varepsilon > 0$, целое N , конечная величина M и набор открытых множеств Ω_i таких, что для всех $x \in \Gamma$ шар радиуса ε с центром в точке x содержится в некотором Ω_i , причем не более N множеств Ω_i имеют ненулевое пересечение, а каждая определяемая соотношением $\Omega_i \cap \Omega = \Omega_i \cap D_i$ область D_i имеет границу, которая является графиком липшицевой функции ω_i (т. е. $D_i = \{(x, y) \in R^n\}; x \in R^{n-1}, y < \omega_i(x)\}$), удовлетворяющей неравенству $\|\omega_i\|_{\text{Lip}(R^{n-1})} \leq M$.

Нормы пространств Соболева различных порядков m связаны между собой важными неравенствами, устанавливаемыми в теореме вложения Соболева:

$$\|v\|_{H^k} \leq c\|v\|_{H^m}, \quad k < m,$$

где v — функция из пространства $H^m(\Omega)$, а c — не зависящая от x постоянная, но зависящая от значений k , m и от Ω , причем предполагается, что последняя имеет липшицеву границу.

Данное неравенство является одним из частных случаев многочисленных теорем вложения Соболева, среди которых мы сформулируем лишь некоторые результаты.

Теорема 1.6 (теорема вложения Соболева). *Имеют место следующие вложения при ограниченности области Ω и $1 \leq p \leq \infty$:*

$$W_p^m(\Omega) \subset W_p^k(\Omega) \text{ для любых целых } 0 \leq k \leq m;$$

$$W_q^k(\Omega) \subset W_p^k(\Omega) \text{ для любых целых неотрицательных } k, p \leq q.$$

Пусть теперь Ω — n -мерная область с липшицевой границей, k и m — целые, $0 \leq m < k$, а p — вещественное число в интервале $1 \leq p < \infty$, причем выполняются неравенства $k - m \leq n$ при $p = 1$ и $k - m > n/p$ при $p > 1$. Тогда существует постоянная c такая, что справедливо неравенство

$$\|u\|_{W_\infty^m(\Omega)} \leq c\|u\|_{W_p^k(\Omega)} \quad \forall u \in W_p^k(\Omega),$$

причем в классе эквивалентности $W_\infty^m(\Omega)$ существует функция из $C^m(\Omega)$.

Подчеркнем, что последнее утверждение не будет справедливым, если граница Γ не является липшицевой. Важность же этого результата заключается в том, что любая функция с достаточно регулярной обобщенной производной может рассматриваться как непрерывная и ограниченная.

Частным случаем теоремы вложения является следующее утверждение: при $m > n/2$ функции из $H^m(\Omega)$ непрерывны в $\bar{\Omega}$ и существует положительное число α такое, что $\forall u \in H^m(\Omega)$ справедливо неравенство

$$\max_{x \in \Omega} |u(x)| \leq c \|u\|_{H^m(\Omega)}.$$

Например, в одномерном случае ($n = 1$) непрерывными являются функции из $H^1(\Omega)$, а в двумерном и трехмерных случаях ($n = 2, 3$) — из $H^2(\Omega)$.

1.1.6. Теоремы о продолжении функции и о следах

Актуальность определения области с липшицевой границей наглядно проявляется в том, что именно такие области, и только такие, допускают продолжение функции во все пространство R^n с сохранением нормы.

Теорема 1.7 (о продолжении функции). Пусть Ω есть область с липшицевой границей. Тогда существует отображение продолжения $T : W_p^k(\Omega) \rightarrow W_p^k(R^n)$, определенное для всех целых неотрицательных k и вещественных чисел p в области $1 \leq p \leq \infty$, удовлетворяющее условию $Tv|_{\Omega} = v$ для всех $v \in W_p^k(\Omega)$ и неравенству

$$\|Tv\|_{W_p^k(R^n)} \leq c \|v\|_{W_p^k(\Omega)},$$

где c — не зависящая от v постоянная.

Другой важный вопрос при решении многомерных краевых задач: если в n -мерной области Ω известна функция какого-то класса, то что можно сказать о свойствах ее следа на границе Γ , т. е. о принадлежности его какому-то классу на этом $(n - 1)$ -м многообразии. Серьезность данной проблемы проистекает хотя бы из того факта, что при $n \geq 2$ пространство $W_2^1(\Omega)$ содержит неограниченные функции и говорить об их граничных значениях в классическом поточечном смысле неправомерно.

Если функция $f(x)$ определена в n -мерной области с липшицевой границей Γ , являющейся графиком некоторой функции $\omega(x)$, то естественным образом на Γ можно определить «поверхностный» интеграл как

$$\int_{\Gamma} f ds \equiv \int_{R^{n-1}} f(x, \omega(x)) \sqrt{1 + |\nabla \omega(x)|^2} dx.$$

Заметим, что если Γ есть липшицева граница, то весовая функция $\sqrt{1 + |\nabla \omega(x)|^2}$ принадлежит классу L_{∞} , и определение интеграла на границе является корректным, что позволяет, в

свою очередь, ввести соответствующие пространства лебеговых функций. В строгом виде данные рассуждения формулируются с помощью следующего утверждения.

Теорема 1.8 (о следах). Пусть Ω есть область с липшицевой границей Γ . Тогда для любого вещественного числа $1 \leq p \leq \infty$ существует константа c такая, что выполняется неравенство

$$\|v\|_{L_p(\Gamma)} \leq c \|v\|_{L_p(\Omega)}^{1-1/p} \|v\|_{W_p^1(\Omega)}^{1/p} \quad \forall v \in W_p^1(\Omega).$$

Если ввести операцию взятия следа функции таким образом, что $tr_\Gamma u$ означает функцию, определенную на многообразии Γ и совпадающую на нем почти всюду со значениями функции, заданной в $\bar{\Omega}$, то пространство $L_2(\Gamma)$, например, можно определить с нормой

$$\|u\|_{L_2(\Gamma)} = \|tr_\Gamma u\|_{L_2(\Gamma)} = \left(\int_\Gamma u^2 ds \right)^{1/2},$$

где ds — мера на многообразии Γ . В этом случае для $u \in H^1(\Omega)$ имеем $tr_\Gamma u \in L_2(\Gamma)$ и справедливо неравенство

$$\|tr_\Gamma u\|_{L_2(\Gamma)} \leq c \|u\|_{H^0(\Omega)},$$

где число c не зависит от u . Выполняется при этом также *обобщенное неравенство Фридрикса*

$$\|u\|_{L_2(\Omega)} \leq c(\gamma) [\|u\|_{H^1(\Omega)}^2 + \|tr_\Gamma u\|_{L_2(\Gamma_\gamma)}^2]^{1/2},$$

где $\Gamma_\gamma \subset \Gamma$, $\text{mes}(\Gamma_\gamma) \geq \gamma > 0$, а число $c(\gamma)$ не зависит от u . Из него следует, в частности, «классическое» *неравенство Фридрикса*, выполняемое для функций $u \in H^1(\Omega)$ с условием $u(x) = 0$, $x \in \Gamma(\gamma) \subset \Gamma$, $\text{mes} \Gamma(\gamma) \geq \gamma > 0$:

$$\|u\|_{L_2(\Omega)} \leq c \|u\|_{H^1(\Omega)}.$$

1.1.7. Пространства $H_0^m(\Omega)$ и $C_0^p(\Omega)$

Укажем еще на один тип пространств, обозначаемых через $H_0^m(\Omega)$ ($\overset{\circ}{W}_2^1(\Omega)$ при $m = 1$) и состоящих из функций, принадлежащих $H^m(\Omega)$ и обращающихся в нуль на границе Γ области Ω , т. е.

$$H_0^m(\Omega) = \{v \in H^m(\Omega); v = 0 \text{ на } \Gamma\}.$$

При условии ограниченности множества Ω для любой функции $v \in H_0^1(\Omega)$ выполняется следующее *неравенство Пуанкаре–Фридрихса*:

$$\|v\|_{H^m} \leq c(\Omega)|v|_1, \quad v \in H_0^1(\Omega), \quad (1.12)$$

где $c(\Omega)$ — не зависящая от v постоянная, а символ $|\cdot|_m$ означает *полунорму* в соответствии со следующим определением:

$$|v|_m = \left(\sum_{|\alpha|=m} \int_{\Omega} |\partial^\alpha u|^2 dx \right)^{1/2}. \quad (1.13)$$

Неравенство Пуанкаре–Фридрихса устанавливает, в частности, что при ограниченности множества Ω полунорма $|\cdot|_1$ является нормой на пространстве $H_0^1(\Omega)$, эквивалентной норме $\|\cdot\|_1$. Нетрудно показать также, что введенная полунорма $|\cdot|_0$ совпадает с нормой пространства $L_2(\Omega)$ и что для $u \in H^1(\Omega)$ справедливо следующее соотношение, иногда называемое *неравенством Пуанкаре*:

$$\|u\|_{L_2(\Omega)} \leq c \left[|u|_{H^1(\Omega)}^2 + \left(\int_{\Omega} u dx \right)^2 \right]^{1/2},$$

где постоянная c не зависит от u .

Из определения $H_0^m(\Omega)$ следует, что оно является подпространством пространства $H^m(\Omega)$. Введем еще «промежуточные» пространства, когда граница области $\Gamma = \Gamma_1 \cup \Gamma_2$ является объединением участков границы Γ_1 и Γ_2 , причем функции $v(x)$ обращаются в нуль только на Γ_1 :

$$H_0^m(\Omega, \Gamma_1) = \{v \in H^m(\Omega); v = 0 \text{ на } \Gamma_1\}.$$

Хотя для функций $v(x)$ из $H_0^m(\Omega, \Gamma_1)$ неравенство Пуанкаре–Фридрихса уже не выполняется, можно показать, что в этом подпространстве полунорма $|\cdot|_1$ есть норма, эквивалентная норме $\|\cdot\|_1$. Забегая вперед, отметим, что введение пространства $H_0^m(\Omega, \Gamma_1)$ (являющегося также подпространством $H^m(\Omega)$) естественным образом оправдывается при анализе смешанных краевых задач, когда на части границы Γ_1 задается условие Дирихле, а на остальной части Γ_2 — граничные условия второго или третьего рода.

Аналогично $H_0^m(\Omega)$ введем еще банаховые пространства функций $C_0^p(\Omega)$, непрерывных в Ω вместе со своими производными до p -го порядка включительно и обращающихся в нуль на границе Γ :

$$C_0^p(\Omega) = \{v \in C^p(\Omega); v = 0 \text{ на } \Gamma\}.$$

Формально, важным частным случаем $C_0^p(\Omega)$ является $C_0^\infty(\Omega)$ — множество бесконечно дифференцируемых функций, отличных от нуля только в строго внутреннем по отношению к $\bar{\Omega}$ множестве. Отсюда следует, что на границе Γ равны нулю не только сами функции $v \in C_0^\infty(\Omega)$, но и все их производные. Важно отметить также, что множество $C_0^\infty(\Omega)$ плотно в пространстве $L_2(\Omega)$. Отметим, что замыкание множества функции $C_0^\infty(\Omega)$ по норме пространства приводит к пространству $H_0^1(\Omega)$, которое является полным и плотным в $L_2(\Omega)$. При этом $H_0^1(\Omega)$ — замкнутое подпространство пространства $H^1(\Omega)$ (неплотное в $H^1(\Omega)$).

1.2. Операторные формулировки краевых задач

Целью данного пункта является рассмотрение основных определений и свойств операторов в абстрактных пространствах, а также изложение «классических» операторных формулировок краевых задач, условий существования и единственности их решений.

1.2.1. Линейные операторы и уравнения

Рассмотрим два метрических пространства X, Y и множество $D \subset X$. Если каждому элементу $x \in D$ по определенному правилу однозначно ставится в соответствие единственный элемент $y \in Y$, то говорят, что на множестве D задан *оператор* A с областью значений, принадлежащей пространству Y . Иногда говорят также, что задано *отображение* множества, называемого *областью определения* оператора A , $D(A) \subset X$ в множество Y , и пишут $A : D \rightarrow Y$ или $y = Ax$, или же $y = A(x)$.

Определение 1.23. Множество элементов вида $R(A) = \{y \in Y, y = Ax, x \in D(A)\}$ называется областью значений оператора A . Элемент y называется образом элемента x , а элемент x — прообразом элемента y .

Таким образом, $R(A)$ есть образ $D(A)$, что можно записать как $R(A) = AD(A)$. Если значения y оператора A являются вещественными или комплексными числами, то оператор называют функционалом.

Определение 1.24. Оператор $A(x)$ называется *непрерывным* в точке $x_0 \in D(A)$, если для $\forall \varepsilon > 0 \exists \delta > 0$ такое, что $\rho_Y(A(x), A(x_0)) < \varepsilon$ для $\forall x \in D$ из шара $\rho_X(x, x_0) < \delta$. Оператор $A(x)$ называется непрерывным в $D(A)$, если он непрерывен в любой точке $x \in D$.

Определение 1.25. Оператор $A : X \rightarrow Y$ с областью определения $D(A)$ называется *линейным*, если выполняются следующие условия:

- $D(A)$ — линейное подпространство пространства X ;
- $A(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 A x_1 + \lambda_2 A x_2$ для $\forall x_1, x_2 \in D(A)$ и любых скаляров λ_1, λ_2 .

Если $Ax = 0$ для $\forall x \in X$, то оператор A называется нулевым.

Область значений $R(A)$ всякого линейного оператора является линейным пространством. Линейный оператор A называется *расширением* линейного оператора $B : X \rightarrow Y$, если $D(A) \supset D(B)$ и $Ax = Bx$ для $\forall x \in D(B)$. Если $\overline{D(A)} = X$, то говорят, что оператор A задан на всюду плотном в X множестве $D(A)$.

Линейный оператор A , у которого область определения совпадает со всем банаховым пространством ($D(A) = X$), а область значений принадлежит банахову пространству Y , обладает следующим свойством: он непрерывен в любой точке $x \in X$, если он непрерывен в точке $0 \in X$. Такой оператор называется *ограниченным*, если конечна его *норма*, определяемая как

$$\|A\| = \sup_{x \in X} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{\|x\|_X=1} \|Ax\|_Y.$$

В силу своего определения норма оператора удовлетворяет неравенству

$$\|Ax\| \leq \|A\| \cdot \|x\|,$$

где значки у норм опущены, так как они понятны по контексту.

Справедливо следующее утверждение: для того чтобы линейный оператор $A : X \rightarrow Y$, $D(A) = X$, был непрерывен, необходимо и достаточно, чтобы он был ограниченным.

Линейный оператор называется *вполне непрерывным*, если он определен во всем пространстве X и отображает каждое ограниченное в X множество в компактное множество в Y . Очевидно, что вполне непрерывный оператор является ограниченным. Оператор A называется *замкнутым*, если из того, что $x_n \rightarrow x$, где $x_n \in D(A)$ и $Ax_n \rightarrow y$ следует, что $x \in D(A)$ и $y = Ax$. Ограниченный оператор является замкнутым. Если оператор A не замкнут, то для того, чтобы он имел замкнутое расширение (допускал замыкание), необходимо и достаточно, чтобы из $x_n \rightarrow 0$ ($x_n \in D(A)$) и $Ax_n \rightarrow y$ следовало, что $y = 0$. Наименьшее замкнутое расширение оператора A называется его *замыканием* и обозначается $\bar{A}$. Если оператор допускает замыкание, то из $x_n \rightarrow x$ при $x_n \in D(A)$ и $Ax_n \rightarrow y$ следует, что $y = \bar{A}x$.

Определение 1.26. Если $A : X \rightarrow Y$ есть линейный оператор, то множество $\ker A = \{x \in X : Ax = 0\}$ называется *ядром оператора* A . Множество $y \in Y$, для которых $y = Ax$ хотя бы для одного $x \in X$, называется *образом оператора* A и обозначается $\operatorname{im} A$. Другими словами, образ оператора есть образ области его определения, или область значений оператора.

Для того, чтобы линейное отображение $A : X \rightarrow Y$, $D(A) \subset X$, $R(A) \subset Y$, было взаимно однозначным, необходимо и достаточно, чтобы его ядро состояло только из нулевого элемента, т. е. $\ker A = 0$. В этом случае оператор A называется *невырожденным* и существует линейный *обратный оператор* $A^{-1} : Y \rightarrow X$, взаимно однозначно отображающий $R(A)$ на $D(A)$ и обладающий следующими свойствами:

$$A^{-1}Ax = Ix = x \text{ для } x \in D(A), \quad AA^{-1}y = y \text{ для } y \in R(A),$$

где оператор $I = A^{-1}A$ называется *единичным*. Отметим, что введенный единичный оператор имеет область определения $D(A)$ и может быть обозначен как $I_{D(A)}$. При этом, вообще говоря, он не равен единичному оператору $I_{R(A)} = AA^{-1}$, имеющему другую область определения $R(A)$.

Теорема 1.9. Обратный оператор A^{-1} существует и одновременно ограничен на $R(A)$ тогда и только тогда, когда существует некоторая постоянная $\alpha > 0$ такая, что для любого $x \in D(A)$ выполнено неравенство $\|Ax\| \geq \alpha\|x\|$, при этом $\|A^{-1}\| \leq \alpha^{-1}$.

Пусть $\mathcal{L}(X, Y)$ есть множество всех линейных непрерывных операторов, имеющих областями определения и значений банаховы пространства X, Y . На этом множестве можно определить операции сложения и умножения на скаляр операторов по формулам

$$(A + B)x = Ax + Bx, \quad (\lambda A)x = \lambda Ax, \quad x \in X, \quad \lambda \in R.$$

Нетрудно показать, что $A + B$ и λA являются линейными непрерывными операторами, а для множества $\mathcal{L}(X, Y)$ выполняются все аксиомы линейного пространства, в котором норма элементов-операторов введена выше.

Таким образом, множество $\mathcal{L}(X, Y)$ является банаховым пространством операторов, в котором сходимость по норме, или равномерная сходимость, последовательности операторов A_n к оператору A определяется как

$$\|A_n - A\| \rightarrow 0 \text{ при } n \rightarrow \infty.$$

Для операторов определяется еще один вид сходимости: говорят, что последовательность $\{A_n\} \in \mathcal{L}(X, Y)$ сильно сходится к оператору $A \in \mathcal{L}(X, Y)$, если для $\forall x \in X$ норма $\|A_n x - Ax\| \rightarrow 0$ при $n \rightarrow \infty$. Из сходимости по норме последовательности операторов следует их сильная сходимость, а обратное утверждение в общем случае не верно.

Операторный ряд $\sum_{k=1}^{\infty} A_k$ называется сходящимся по норме, если сходится по норме последовательность его частичных сумм $S_n = \sum_{k=1}^n A_k$. Говорят, что операторный ряд сходится абсолютно, если сходится числовой ряд $\sum_{k=1}^{\infty} \|A_k\|$. Справедливо следующее утверждение: если операторный ряд сходится абсолютно, то он сходится и по норме.

В приложениях зачастую встречаются пространства операторов $\mathcal{L}(X, X) = \mathcal{L}(X)$, для которых области определения и значений принадлежат одному банаховому пространству X . В пространстве $\mathcal{L}(X)$ можно ввести еще одну операцию — умножение операторов по формуле

$$BAx = B(Ax), \quad BAx \in X \quad \forall x \in X.$$

Для операторов $A, B \in \mathcal{L}(X)$ имеют место следующие оценки норм для их суммы и произведения:

$$\|A + B\| \leq \|A\| + \|B\|, \quad \|AB\| \leq \|A\| \cdot \|B\|.$$

Два оператора $A, B \in \mathcal{L}(X)$ называются *коммутативными* или *перестановочными*, если $BA = AB$ (в общем случае это равенство не верно, а операторы — не коммутативны).

Аналогично многочленам от скалярной переменной определяется операторный многочлен порядка N :

$$P_N(A) = \sum_{k=0}^N a_k A^k,$$

где $A^k = A \cdot A^{k-1}$, $A^0 = I$ и a_k — степени оператора и числовые коэффициенты. Вводится также и более общее понятие: если $\varphi(\lambda) = \sum_{k=0}^{\infty} a_k \lambda^k$ есть аналитическая в круге $|\lambda| < r$ функция комплексного переменного λ и $A \in \mathcal{L}(X)$, $\|A\| < r$, то по формуле

$$\varphi(A) = \sum_{k=0}^{\infty} a_k A^k$$

определяется *функция от оператора* (при $a_k \equiv 1$ данный ряд называется *рядом Неймана*).

Важным частным случаем является оператор $B \in \mathcal{L}(X)$, называемый *квадратным корнем* оператора $A \in \mathcal{L}(X)$, определяемый равенством $B^2 = A$ (если это уравнение имеет единственное решение) и обозначаемый как $A^{1/2} = B$.

В связи с введенными в п. 1.1.1 понятиями ортогональности и проектирования естественно ввести соответствующие операторы.

Определение 1.27. Пусть $X = Y = H$ — сепарабельное гильбертовое пространство, L — его подпространство и $\forall x \in H$ справедливо соотношение $x = y + z$, $y \in L$, $z \in L^\perp$. Тогда оператор отображения $y = P_L x$ называется оператором (ортогонального) *проектирования* пространства H на подпространство L , или *проектором* (ортопроектором) H на L .

Очевидно, что оператор проектирования (для простоты будем обозначать его также через P) является линейным и обладает свойствами $\|P\| = 1$, $P^2 = P$, т. е. $Pz = z$ для любого z из области значений P .

Рассмотренные в п. 1.1.3 линейные функционалы, понятие о которых было необходимо ввести в связи с определением пространств Соболева, являются частным случаем линейных операторов.

Если X — банаховое пространство, то линейные непрерывные функционалы $L(x)$, $x \in X$, принадлежат пространству $\mathcal{L}(X, R^1)$ линейных ограниченных функционалов, которое называется сопряженным к X и обозначается X^* . Так как X^* есть банаховое пространство, то на его элементах можно, в свою очередь, определить линейные функционалы, образовав «второе сопряженное» к X пространство X^{**} . Если $X^* = X$ с точностью до изометрии, то $X^{**} = X$ и такие пространства называют *рефлексивными*. Гильбертово пространство является рефлексивным и называется также самосопряженным.

В пространствах X и X^* можно определить новый тип сходимости — *слабую сходимость*. Говорят, что последовательность $\{x_n\} \in X$ сходится к $x \in X$ слабо в X , если для $\forall L(x) \in X^*$ $L(x_n) \rightarrow L(x)$ при $n \rightarrow \infty$. Введенное ранее понятие сильной сходимости к $A \in \mathcal{L}(X, Y)$ последовательности операторов $A_n \in \mathcal{L}(X, Y)$, примененное к функционалам, эквивалентно понятию слабой сходимости.

Справедливо следующее утверждение, называемое *теоремой Хана–Банаха* о продолжении линейного функционала с сохранением его нормы: если S — замкнутое подпространство банахового пространства X , а $L(x)$, $x \in S$, — определенный на S линейный непрерывный функционал, то существует непрерывный линейный функционал $\bar{L}(x)$, $x \in X$, определенный на X и являющийся расширением $L(x)$, с такой же нормой

$$L(x) = \bar{L}(x) \text{ для } \forall x \in S, \quad \|L\|_S = \|\bar{L}\|_X.$$

Многочисленные прикладные задачи могут быть записаны в виде линейного операторного уравнения первого рода относительно неизвестного $x \in X$:

$$Ax = y$$

с линейным оператором A и заданным элементом (правой частью) $y \in Y$. Если $y \in R(A)$ и оператор A имеет обратный A^{-1} , то решение данного уравнения единственно и имеет вид $x = A^{-1}y$.

Если A — непрерывно обратим, то операторное уравнение имеет единственное решение при любой правой части $y \in X$. Если при этом $\bar{x} = A^{-1}\bar{y}$ есть решение аналогичного уравнения с другой правой частью $\bar{y}$ и оператор A^{-1} ограничен, то

$$\|x - \bar{x}\| = \|A^{-1}(y - \bar{y})\| \leq \|A^{-1}\| \cdot \|y - \bar{y}\|.$$

Последнее означает, что малое изменение правой части в операторном уравнении влечет малое изменение решения, что называется *корректной разрешимостью* уравнения.

Определение 1.28. Если $A, A^{-1} \in \mathcal{L}(X)$ и X — банаховое пространство, то величина

$$\kappa(A) = \|A\| \cdot \|A^{-1}\| \geq 1$$

называется *числом* (мерой) *обусловленности оператора* A .

Пусть $\bar{x}$ — некоторое приближенное решение невырожденного операторного уравнения, а $r = y - A\bar{x}$ — его невязка. Тогда для ошибки $\delta = x - \bar{x}$ выполняются соотношения

$$A\delta = r, \quad \delta = A^{-1}r,$$

откуда получаем следующие неравенства для норм:

$$\|r\| \leq \|A\| \cdot \|\delta\|, \quad \|\delta\| \leq \|A^{-1}\| \cdot \|r\|.$$

Отсюда при $y \neq 0$ для относительной погрешности $\|\delta\|/\|x\|$ с помощью очевидных выражений

$$\|y\| \leq \|A\| \cdot \|x\|, \quad \|x\| \leq \|A^{-1}\| \cdot \|y\|$$

имеем двустороннюю оценку

$$\kappa^{-1}(A) \frac{\|r\|}{\|y\|} \leq \frac{\|\delta\|}{\|x\|} \leq \kappa(A) \frac{\|r\|}{\|y\|}$$

через относительную величину невязки $\|r\|/\|y\|$ и число обусловленности $\kappa(A)$.

Если же $\bar{x}$ — точное решение операторного уравнения с возмущенной правой частью $\bar{y} = y - \Delta$, то вследствие справедливости уравнения для ошибки

$$A(x - \bar{x}) = A\delta = \Delta$$

с помощью аналогичных преобразований получаем верхнюю оценку для погрешности решения через возмущение правой части

$$\frac{\|\delta\|}{\|x\|} \leq \kappa(A) \frac{\|\Delta\|}{\|y\|},$$

где опять же фигурирует число обусловленности. Отсюда видно, что число обусловленности $\kappa(A)$ играет ключевую роль при оценках погрешностей приближенных решений операторных уравнений.

Пусть $A \in \mathcal{L}(X, Y)$ и $f = f(x) \in X^*$, $x \in X$, — некоторый функционал, для которого также используется скобочное обозначение $\langle f, x \rangle$, а g — непрерывный линейный функционал в пространстве Y , т.е. элемент сопряженного пространства Y^* . Для произвольного $x \in X$ имеем $y = Ax \in Y$ и потребуем выполнения равенства

$$\langle f, x \rangle = \langle g, Ax \rangle.$$

Эта формула соотносит каждому функционалу $f \in X^*$ некоторый функционал $g \in Y^*$. Оператор, осуществляющий это отображение, называется *сопряженным* по отношению к данному оператору A и обозначается A^* , а соответствующее уравнение

$$A^*g = f$$

называется *сопряженным* к уравнению $Ax = y$. Таким образом, каждому $A \in \mathcal{L}(X, Y)$ соответствует сопряженный оператор $A^* \in \mathcal{L}(Y^*, X^*)$, переводящий Y^* в X^* . Можно показать, что A^* является линейным ограниченным оператором, причем $\|A^*\| = \|A\|$.

В случае гильбертовых пространств $X = Y = H$ имеем $X^* = Y^* = H$ и $A \in \mathcal{L}(H)$. В этом случае по теореме Рисса определение сопряженного (или эрмитового) оператора можно записать с помощью скалярного произведения как

$$(Ax, g) = (x, A^*g) \quad \forall x, g \in H.$$

При условии $A = A^*$, когда для $\forall x, g \in H$ выполняется равенство

$$(Ax, g) = (x, Ag),$$

оператор A называется *самосопряженным*.

Если A и B — самосопряженные операторы, то для них можно показать справедливость следующих утверждений:

- для любых скаляров λ_1 и λ_2 оператор $\lambda_1 A + \lambda_2 B$ — самосопряженный;
- оператор AB будет самосопряженным в том и только в том случае, когда операторы A и B коммутируют, т. е. $AB = BA$;
- если $A \in \mathcal{L}(H)$ — самосопряженный оператор, а $P_n(t)$ — многочлен от аргумента t , то $P_n(A)$ — самосопряженный оператор.

Для самосопряженного оператора в гильбертовом пространстве H имеем $D(A) = H$. Существует более широкий класс линейных операторов, называемых *симметричными*, которые обладают свойствами самосопряженных, но определены только на всюду плотном в H линейном многообразии, т. е. для них $\overline{D(A)} = H$. В отличие от самосопряженного, симметричный оператор может и не быть ограниченным.

Для некоторых пар симметричных (и, естественно, для самосопряженных) операторов можно установить важные отношения «больше» и «больше или равно».

Определение 1.29. Симметричный оператор A называется *положительно полуопределенным* (или неотрицательным, с обозначением $A \geq 0$), если $(Ax, x) \geq 0$ для $\forall x \in D(A)$, $x \neq 0$, и называется *положительно определенным* (или положительным, с обозначением $A > 0$), если существует такая постоянная $c > 0$, что $(Ax, x) \geq c(x, x)$ для $\forall x \in D(A)$.

Сделаем следующее терминологическое замечание: иногда разделяют понятия положительно определенного и положительного операторов, определяя последний по условию $(Ax, x) > 0$ $\forall x \in D(A)$, $x \neq 0$, но мы такого разграничения использовать не будем.

Если разность двух симметричных операторов $A - B$ есть положительный или неотрицательный оператор, то можно записать $A > B$ или $A \geq B$ соответственно.

Для симметричного положительно определенного оператора A справедливо *обобщенное неравенство Коши–Буняковского*

$$|(Ax, y)| \leq (Ax, x)^{1/2} (Ay, y)^{1/2}.$$

Для такого оператора на $D(A)$ можно определить новое скалярное произведение

$$(x, y)_A = (Ax, y),$$

с помощью которого обобщенное неравенство Коши–Буняковского формулируется как обычное:

$$|(x, y)_A| \leq (x, x)_A^{1/2} (y, y)_A^{1/2}.$$

Если операторы A, B самосопряжены и для всех $x \in H$ выполняется равенство их квадратичных форм $(Ax, x) = (Bx, x)$, то $A = B$. Для самосопряженного оператора норма определяется как

$$\|A\| = \sup_{\|x\|=1} |(Ax, x)|.$$

Определение 1.30. Скалярная величина λ называется собственным числом (значением) оператора A , если существует элемент $x \neq 0$, $x \in D(A)$, для которого

$$Ax = \lambda x,$$

где x называется соответствующим *собственным элементом* оператора A .

Очевидно, что собственный элемент определяется с точностью до скалярного множителя, т.е. все элементы вида αx при $\alpha \neq 0$ будут одновременно с x собственными элементами оператора A , соответствующими одному собственному числу λ . Максимальное число r ($1 \leq r \leq \infty$) линейно независимых элементов, соответствующих данному собственному числу λ , называется *кратностью* этого собственного значения. Если $r = 1$, то λ называется *простым собственным числом*. Если $x_1, \dots, x_r$ — линейно независимые элементы, соответствующие собственному числу λ кратности r , то любая их конечная линейная комбинация будет также соответствующим собственным элементом.

Кратность собственного числа r может быть бесконечной. Например, единичный оператор I имеет собственное число $\lambda = 1$, кратность которого равна размерности гильбертового пространства H .

Систему линейно независимых собственных элементов $x_1, \dots, x_n \dots$ (как и любую другую систему линейно независимых элементов, в том числе бесконечную) можно ортогонализировать, т. е. преобразовать в ортогональную систему $y_1, \dots, y_n \dots$, обладающую свойством

$$(y_i, y_j) = (y_i, y_i) \delta_{i,j},$$

где $\delta_{i,j}$ — символ Кронекера, с помощью следующего процесса ортогонализации Грама–Шмидта:

$$\begin{aligned} y_1 &= x_1, & y_k &= x_k - \sum_{i=1}^{k-1} \gamma_{k,i} y_i, \\ \gamma_{k,i} &= (y_i, x_k) / (y_i, y_i), & k &= 2, 3, \dots \end{aligned}$$

В свою очередь, ортогональная система $y_1, \dots, y_n, \dots$ может быть ортонормирована, т. е. преобразована в ортонормальную систему элементов $e_1, \dots, e_n, \dots$, $\|e_k\| = 1$, с помощью формулы $e_k = y_k / \|y_k\|$.

Собственные числа операторов в общем случае являются комплексными. Однако для симметричных операторов, даже для комплексных гильбертовых пространств, справедливо следующее утверждение.

Теорема 1.10. *Собственные числа симметричного оператора вещественны, а собственные элементы, соответствующие различным собственным значениям, ортогональны.*

Если система собственных элементов $x_1, \dots, x_k, \dots$ симметричного оператора A не более чем счетна, то ее можно выбрать ортонормированной:

$$\begin{aligned} (Ax_i, x_k) &= \lambda_k (x_i, x_k) = \lambda_i \delta_{i,k}, \\ Ax_i &= \lambda_i x_i, & i &= 1, 2, \dots \end{aligned}$$

Здесь предполагается, что все собственные значения пронумерованы, причем каждое собственное число повторяется столько раз, какова его кратность, а собственные элементы x_i нумеруются таким образом, чтобы каждому из них соответствовало

одно собственное число λ_i . Относительно таких операторов иногда говорят, что они имеют чисто точечный спектр и являются операторами скалярного типа.

Для $x \neq 0$ величина

$$\Phi_A(x) = (Ax, x)/(x, x)$$

называется *отношением Рэлея*.

Для самосопряженного оператора A справедливы следующие минимаксные свойства собственных значений:

$$m_A = \min_i \lambda_i = \inf_{x \neq 0} \frac{(Ax, x)}{(x, x)},$$

$$M_A = \max_i \lambda_i = \sup_{x \neq 0} \frac{(Ax, x)}{(x, x)}.$$

Величина M_A называется *спектральным радиусом* такого оператора, а его норма равна

$$\|A\| = \max\{|m_A|, |M_A|\}.$$

Если система собственных элементов $\{x_k\}$ полна и ортонормирована в H , то произвольный элемент $x \in H$ можно разложить в *ряд Фурье*

$$x = \sum_{k=1}^{\infty} a_k x_k, \quad a_k = (x, x_k),$$

а действие оператора A при этом выражается формулой

$$Ax = \sum_{k=1}^{\infty} \lambda_k a_k x_k.$$

Отсюда следует, что если функция $f(x)$ определена на некотором множестве Ω , включающем все собственные значения λ_k самосопряженного оператора A , то действие соответствующей функции от оператора описывается формулой

$$f(A)x = \sum_{k=1}^{\infty} f(\lambda_k) a_k x_k.$$

Знание базиса из собственных векторов $\{x_k\}$ оператора A позволяет формально просто выразить решение невырожденного операторного уравнения $Ax = y$, $y \in H$:

$$x = \sum_{k=1}^{\infty} \lambda_k^{-1} b_k x_k,$$

где b_k — коэффициенты разложения правой части в ряд Фурье

$$y = \sum_{k=1}^{\infty} b_k x_k, \quad b_k = (y, x_k).$$

1.2.2. Корректность краевой задачи и энергетические пространства

Пусть F, G — некоторые пространства функций, определенных в Ω и Γ соответственно, а U — пространство функций, определенных в замкнутой области $\bar{\Omega} = \Omega \cup \Gamma$. Введем в рассмотрение дифференциальные операторы L и l , заданные в области Ω и на границе Γ , определенные на функциях из U , и области значений которых находятся в F, G , т. е.

$$LU \subset F, \quad lU \subset G.$$

Оператор l называется оператором граничных (краевых) условий.

Абстрактная операторная *краевая задача* формулируется следующим образом: в пространстве U найти элемент u , являющийся решением уравнений

$$\begin{aligned} Lu &= f, & f &\in F, \\ lu &= g, & g &\in G, \end{aligned} \tag{1.14}$$

где f и g — заданные элементы своих пространств.

Определение 1.31. Краевая задача (1.14) называется корректной, если для любых $f \in F$ и $g \in G$ существует единственное решение $u \in U$, непрерывно зависящее от исходных данных.

Краевую задачу (1.14) можно записать в виде одного операторного уравнения

$$Au = \varphi, \quad u \in U, \quad \varphi \in \Phi, \tag{1.15}$$

если ввести обозначения

$$Au = \begin{cases} Lu \in F, \\ lu \in G, \end{cases} \quad \varphi = \begin{cases} f \in F, \\ g \in G, \end{cases} \quad \Phi = \begin{cases} F, \\ G. \end{cases}$$

Конкретные определения пространств и операторов формируют многообразие типов дифференциальных уравнений и краевых задач.

Пусть H — гильбертово пространство со скалярным произведением $(\cdot, \cdot)$ и нормой $\|v\| = (v, v)^{1/2}$, а $D(A) \subset H$ — всюду плотное в H линейное многообразие, являющееся областью определения некоторого линейного оператора A с областью значений $R(A) \subset H$.

Предположим, что оператор A является самосопряженным и положительно определенным в $D(A)$, т. е. для любых $u, v \in D(A)$ выполняются соотношения

$$(Au, v) = (u, Av), \quad (1.16)$$

$$(Av, v) \geq \alpha \|v\|^2, \quad \alpha > 0, \quad (1.17)$$

где α — не зависящая от v постоянная.

Запишем абстрактную краевую задачу в форме операторного уравнения

$$Au = \varphi, \quad u \in D(A), \quad \varphi \in H. \quad (1.18)$$

Отметим имеющееся существенное отличие последней постановки от (1.15): если ранее область определения оператора A рассматривалась как пространство, то в (1.18) $D(A)$ определяется только как множество. Напомним, что плотность множества $D(A)$ в пространстве H означает, что для любых $x \in H$ и $\varepsilon > 0$ существует элемент $y \in D(A)$ такой, что $\|x - y\| \leq \varepsilon$.

Заметим, что операторное уравнение (1.18) может вообще не иметь решений в $D(A)$. Это наглядно иллюстрируется следующей краевой задачей:

$$-\frac{d^2 u}{dx^2} = f(x), \quad 0 < x < 1, \quad u(0) = u(1) = 0.$$

Если положить $H = L_2[0, 1]$, $D(A) = C_0^2[0, 1]$, то в этом примере при определении правой части

$$f(x) = \begin{cases} 1, & 0 < x \leq 0.5, \\ 0, & 0.5 < x < 1, \end{cases}$$

решение в $D(A)$ не существует.

Очевидно, что в области определения самосопряженного положительно определенного оператора $D(A)$ можно ввести новое скалярное произведение и норму

$$(u, v)_A = (Au, v), \quad \|v\|_A = [(v, v)_A]^{1/2}, \quad (1.19)$$

которые удовлетворяют всем необходимым аксиомам и определяют некоторое гильбертово пространство $\tilde{H}_A$, состоящее из элементов $D(A)$. Если это пространство оказывается неполным, т. е. в нем имеется последовательность, сходящаяся в себе, но не имеющая предела в $\tilde{H}_A$, то существует полное гильбертово пространство H_A , получаемое из $\tilde{H}_A$ с помощью стандартной процедуры пополнения и называемое *энергетическим пространством*. Соответственно $(\cdot, \cdot)_A$ и $\|\cdot\|_A$ будем именовать *энергетическим скалярным произведением* и *энергетической нормой*.

Из неравенства (1.17) и самого построения энергетического пространства следуют соотношения между нормами

$$\|v\| \leq \|v\|_A / \alpha$$

и плотность пространства H_A в H (поскольку $D(A) \subset H_A \subset H$).

Приведем пример краевой задачи, которая порождает симметричный (но не самосопряженный) неограниченный оператор: пусть $H = L_2(0, \pi)$, $D(A) = \{u(x) : u(x) \in C^2[0, \pi], u(0) = u(\pi) = 0\}$,

$$Au = -\frac{d^2u}{dx^2} + u.$$

Симметричность оператора A показывается для $u, v \in D(A)$ с помощью интегрирования по частям, поскольку с учетом однородных граничных условий имеем

$$\int_0^\pi \left(-\frac{d^2u}{dx^2} + u\right) v dx = \int_0^\pi u \left(-\frac{d^2v}{dx^2} + v\right) dx.$$

Однако оператор $A : L_2(0, \pi) \rightarrow L_2(0, \pi)$ не ограничен. Действительно, последовательность функций $u_n(x) = \sqrt{2/\pi} \sin x$, $n = 1, 2, \dots$, удовлетворяющих краевым условиям $u_n(0) = u_n(\pi) = 0$, имеет норму, равную единице, но при этом

$$\|Au_n\|_{L_2} = (n^2 + 1) \rightarrow \infty \text{ при } n \rightarrow \infty.$$

1.2.3. Существование, устойчивость и единственность решений краевых задач

Теперь будет оправданным вернуться от (1.18) к операторному уравнению (1.15), в котором область определения и область значений оператора A можно считать гильбертовыми пространствами с нормами $\|\cdot\|_A$ и $\|\cdot\|$. Если оператор A является невырожденным, то решение уравнения (1.15) существует и формально записывается в виде $u = A^{-1}\varphi$.

Определение 1.32. Решение уравнения (1.15) называется *устойчивым*, если существует такая постоянная c , что для всех $\varphi \in H$ выполняется неравенство

$$\|u\|_A \leq c\|\varphi\|. \quad (1.20)$$

Смысл этого понятия, безусловно связанного с корректностью краевой задачи, представляется очевидным, хотя он и не тривиален. Пусть $\tilde{\varphi} = \varphi + \delta\varphi \in H$, $\tilde{u} = u + \delta u \in D(A)$ — правая часть и решение «возмущенной» операторной задачи $A\tilde{u} = \tilde{\varphi}$. Тогда в силу линейности A справедливо уравнение $A\delta u = \delta\varphi$, и при малых вариациях правой части $\delta\varphi$ соответствующие возмущения решения δu будут «достаточно малыми» (хотя это и не исключает того, что величина $\|\delta u\|_A$ будет намного больше $\|\delta\varphi\|$); главное следствие из (1.20) состоит в том, что $\|\delta u\|_A \rightarrow 0$ при $\|\delta\varphi\| \rightarrow 0$.

Теорема 1.11. *Решение уравнения (1.15), если оно существует, при условиях (1.16), (1.17) единственно и устойчиво.*

Рассмотрим в области $\Omega \subset R^n$ дифференциальное уравнение второго порядка *дивергентного вида*

$$Lu \equiv - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{i,j}(x) \frac{\partial u}{\partial x_j} \right) + a_0(x)u = f(x),$$

где дифференциальный оператор можно записать в форме $L = -\operatorname{div}(a \operatorname{grad} u) + a_0 I$, а коэффициенты предполагаются удовлетворяющими следующим условиям:

$$\begin{aligned} a_{i,j} \in C^1(\bar{\Omega}), \quad a_0 \in C(\bar{\Omega}), \quad a_{i,j}(x) = a_{j,i}(x), \\ x \in \bar{\Omega} = \Omega \cup \Gamma. \end{aligned} \quad (1.21)$$

Уравнение (1.20) называется *эллиптическим в области Ω* , если существует положительное число α_0 такое, что для любых вещественных чисел $\xi_1, \dots, \xi_n$ и для любого $x \in \Omega$ выполняется неравенство

$$\sum_{i,j=1}^n a_{i,j}(x) \xi_i \xi_j \geq \alpha_0 \sum_{i=1}^n \xi_i^2.$$

На искомое решение $u(x)$ наложим *смешанные краевые условия*, т.е. на разных частях границы $\Gamma = \Gamma_1 \cup \Gamma_2$ рассматриваются или граничные условия Дирихле (первого рода), или условия Ньютона (третьего рода, или условия Робена):

$$\begin{aligned} l_1 u|_{\Gamma_1} &\equiv u|_{\Gamma_1} = g_1(x), & x \in \Gamma_1; \\ l_2 u|_{\Gamma_2} &\equiv \left(\frac{\partial u}{\partial \bar{n}} + \sigma u \right) \Big|_{\Gamma_2} = g_2(x), & x \in \Gamma_2, \end{aligned} \quad (1.22)$$

где $\partial/\partial \bar{n}$ означает *конормальную производную*:

$$\frac{\partial u}{\partial \bar{n}} = \sum_{i,j=1}^n a_{i,j} \frac{\partial u}{\partial x_i} \cos(\bar{n}, \bar{x}^j),$$

а $\bar{n}(x)$, $\bar{x}^j$ — орты внешней нормали (определенной почти всюду на Γ_2) и j -й координатной оси, причем

$$\sigma \in C(\Gamma_2), \quad \sigma(x) \geq 0, \quad x \in \Gamma_2.$$

Если $\sigma(x) = 0$ для $x \in \Gamma_2$, то соответствующее условие на Γ_2 называется условием Неймана. В случае $x \in \Gamma_2 = 0$ или $\Gamma_1 = 0$ рассматриваемая задача (1.20)–(1.22) называется первой или третьей (второй — при $\sigma(x) \equiv 0$) краевой задачей соответственно.

Для исследования данной эллиптической краевой задачи важное значение имеет *формула Грина*

$$\int_{\Omega} (Lu)v \, dx = \int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + a_0 uv \right) dx + \int_{\Gamma} \frac{\partial u}{\partial \bar{n}} v \, ds, \quad (1.23)$$

справедливая для всех u, v из области определения оператора A краевой задачи

$$D(A) = \{u \in C^2(\Omega), l_k u|_{\Gamma_k} = g_k(x), \forall x \in \Gamma_k, k = 1, 2\}.$$

Множество $D(A)$ плотно в $L_2(\Omega)$, а в частном случае однородных граничных условий ($g_1(x)=0, g_2(x)=0$) для $u, v \in D(A)$ имеем равенство

$$\int_{\Gamma} \frac{\partial u}{\partial n} v ds = \int_{\Gamma} u \frac{\partial v}{\partial n} ds = 0.$$

При этом, в силу условий (1.21), билинейная форма

$$a(u, v) = \int_{\Omega} \left(a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + a_0 uv \right) dx \quad (1.24)$$

будет симметричной, и для нее с помощью свойства эллиптичности (1.23) и обобщенного неравенства Фридрихса (при условии $\text{mes}(\Gamma_1) > 0$) можно показать $\forall u \in D(A)$ справедливость соотношения

$$a(u, u) \geq \left(\alpha_0 c^{-1} + \min_{x \in \Omega} a_0(x) \right) \|u\|_{L_2(\Omega)}^2,$$

где константа c не зависит от u .

Таким образом, неравенство

$$\alpha = \alpha_0 c^{-1} + \min_{x \in \Omega} a_0(x) > 0 \quad (1.25)$$

является условием положительной определенности билинейной формы (1.24) и соответствующего оператора A краевой задачи, порождающего энергетическое пространство H_A со скалярным произведением $(Au, v) = a(u, v)$. В этом случае задача (1.21)–(1.24) корректна и имеет единственное решение.

Отметим, что данная краевая задача будет вырожденной при $a_0(x) = 0, \sigma(x) = 0$ и «чисто неймановских» граничных условиях, т. е. $\Gamma = \Gamma_2$. В этом случае «классическое» решение существует только при условии $\int_{\Omega} f(x) dx = 0$ и определено с точностью до постоянной.

1.3. Примеры и классификация задач моделирования

Назначение данного параграфа является двояким. С одной стороны, мы даем классификационные признаки задач математического моделирования, а с другой — приводим характерные

типы их постановок для различных приложений. Предлагаемый здесь материал традиционно излагается в монографиях и учебниках по уравнениям математической физики или по методам вычислений, а также в книгах, посвященным изучению конкретных предметных областей. Такой профессиональной литературы очень много, но мы отметим только немногих авторов, книги которых указаны в нашем списке литературы: В. И. Агошков [1], В. И. Арнольд [5], А. А. Арсеньев [7], Н. С. Бахвалов [17], Ю. А. Березин и Н. П. Жидков [24], Ю. Е. Бояринцев [25], П. Н. Вабищевич [30]–[34], В. А. Галкин [41], А. Д. Григорьев [52], А. М. Гришин [54], [55], Н. Т. Данаев [56], Б. П. Демидович [61], В. П. Дыминых [67], В. П. Ильин [79], [84], В. Е. Захаров [76], А. Н. Коновалов [90], [91], В. Г. Коробейников [93], Л. А. Крукиер [97], Г. И. Марчук [121], [122], Г. А. Михайлов [131], Н. Н. Моисеев [132], В. В. Пененко [143], В. М. Садовский [155], А. А. Самарский [163], А. Н. Тихонов [175]–[177], Л. Д. Фаддеев [174], Р. П. Федоренко [182], [183], Ю. И. Шокин [205], Н. Н. Яненко [206]–[210] – хотя этот список можно значительно расширить.

1.3.1. Уравнения, содержащие оператор Лапласа

Приведем несколько простейших примеров краевых задач для n -мерного евклидова пространства R^n , в котором определен дифференциальный оператор Лапласа

$$\Delta u \equiv \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \dots + \frac{\partial^2 u}{\partial x_n^2},$$

действующий на скалярную функцию $u(x_1, x_2, \dots, x_n)$.

Задача Дирихле для уравнения Пуассона

$$\begin{aligned} -\Delta u(\vec{x}) &= f(\vec{x}), & \vec{x} &\in \Omega \subset R^n, \\ u(\vec{x}) &= g(\vec{x}), & \vec{x} &= (x_1, x_2, \dots, x_n) \in \Gamma, \end{aligned} \quad (1.26)$$

соответствует определению $L = -\Delta$ и $l = I$, где I — тождественный оператор (если в (1.26) $f(\vec{x}) = 0$ для всех $\vec{x} \in \Omega$, мы получаем уравнение Лапласа). Данная относительно простая задача описывает распределения стационарных полей в самых различных приложениях: электро- и магнитостатика, тепло- и массоперенос, потенциальное течение жидкости и т.д.

Для функций $u(\vec{x})$ и $v(\vec{x})$, непрерывных вместе со своими первыми производными внутри $\bar{\Omega} = \Omega \cup \Gamma$ и имеющих непрерывные вторые производные внутри Ω , справедлива первая формула

Грина

$$\int_{\Omega} v \Delta u \, d\vec{x} = \int_{\Gamma} v \frac{\partial u}{\partial \vec{n}} \, ds - \int_{\Omega} \nabla v \nabla u \, dx.$$

Если в ней положить, в частности, $v \equiv 1$ и учесть соотношение $-\Delta u = f$, то для решения уравнения Пуассона (1.26) получаем следующий закон сохранения: количество появляющейся в объеме Ω субстанции равно ее потоку через границу Γ .

Кроме используемых в (1.26) декартовых координат, во многих приложениях употребляется оператор Лапласа в цилиндрической системе координат

$$\begin{aligned} x_1 &= r \cos \phi, & x_2 &= r \sin \phi, & x_3 &= z, \\ \Delta u &= \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 u}{\partial \phi^2} + \frac{\partial^2 u}{\partial z^2}, \end{aligned} \quad (1.27)$$

а также в сферических координатах:

$$\begin{aligned} x_1 &= r \sin \theta \cos \phi, & x_2 &= r \sin \theta \sin \phi, & x_3 &= r \cos \theta, \\ \Delta u &= \frac{1}{r^2} \left(\frac{\partial}{\partial r} r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 u}{\partial \phi^2}. \end{aligned} \quad (1.28)$$

Естественно, если функция u зависит не от всех координат, то определения оператора сводятся к одномерному или двумерному случаям.

Кроме дифференциальных уравнений, с оператором Лапласа связываются важные для многих приложений интегральные отношения. Пусть, например, $u = u(\vec{x})$ и $v = v(\vec{x})$ — функции, непрерывные вместе со своими первыми производными внутри замкнутой области $\bar{\Omega} = \Omega \cup \Gamma$ и имеющие непрерывные вторые производные внутри Ω . Если граница Γ является поверхностью с кусочно-непрерывной кривизной, то справедливы следующие первая и вторая формулы Грина:

$$\int_{\Omega} v \Delta u \, d\omega = \int_{\Gamma} v \frac{\partial u}{\partial \vec{n}} \, d\gamma - \int_{\Omega} \nabla v \nabla u \, d\omega, \quad (1.29)$$

$$\int_{\Omega} (u \Delta v - v \Delta u) \, d\omega = \int_{\Gamma} \left(u \frac{\partial v}{\partial \vec{n}} - v \frac{\partial u}{\partial \vec{n}} \right) \, d\gamma, \quad (1.30)$$

где $d\omega$ и $d\gamma$ представляют собой объемный и поверхностный дифференциалы соответственно. Эти и другие интегральные соотношения, как мы увидим позже, играют важную роль при

построении и обосновании численных методов (в частности, балансных, или консервативных, наследующих на дискретном уровне законы сохранения субстанции).

Начально-краевая задача для уравнения теплопроводности, решение которой $u(\vec{x}, t)$ удовлетворяет соотношениям

$$\begin{aligned}\frac{\partial u}{\partial t} - \Delta u &= f \text{ в } \Omega, \\ u(\vec{x}, 0) &= u_0(\vec{x}) \text{ в } D, \\ u(\vec{x}, t) &= g, \quad \vec{x} \in \partial D, \quad t \in (0, T),\end{aligned}$$

где в данном случае $L = \frac{\partial}{\partial t} - \Delta$, D — ограниченная в R^n область с границей ∂D , а Ω и Γ из (1.14) определяются как

$$\Omega = D \times (0, T) \subset R^{n+1}, \quad \Gamma = (\partial D \times (0, T)) \cup D, \quad 0 \leq T < \infty,$$

причем величины x_i и t здесь имеют физический смысл пространственных координат и времени, а $u_0(\vec{x})$ и $g(\vec{x}, t)$ называются функциями начальных данных и граничных условий.

С используемыми выше обозначениями аналогично определяется начально-краевая задача для *волнового уравнения*:

$$\begin{aligned}\frac{\partial^2 u}{\partial t^2} - \Delta u &= f \text{ в } \Omega, & u(\vec{x}, 0) &= u_0(\vec{x}), \\ \frac{\partial u}{\partial t}(\vec{x}, 0) &= u_1(\vec{x}) \text{ в } D, & u(\vec{x}, t) &= g, \quad \vec{x} \in \partial D, \quad t \in (0, T),\end{aligned}$$

отличительной особенностью которой является наличие двух начальных условий, что связано с присутствием второй производной по времени в дифференциальном операторе $L = \frac{\partial^2}{\partial t^2} - \Delta$. Название данного уравнения происходит из того, что оно описывает всевозможные волновые процессы в самых различных приложениях: электромагнетизма, механики сплошной среды, финансовых и социальных явлениях и т.д.

1.3.2. Классификация дифференциальных уравнений и операторов

Приведенные примеры являются характерными представителями задач для скалярных дифференциальных уравнений эллиптического, параболического и гиперболического типов.

Систематическую классификацию уравнений в частных производных можно наглядно провести для линейных уравнений второго порядка с переменными коэффициентами $a_{i,j}$, b_i , c :

$$Lu \equiv - \sum_{i,j=1}^n a_{i,j}(\vec{x}) \frac{\partial^2 u}{\partial x_i \partial x_j} + \sum_{i=1}^n b_i(\vec{x}) \frac{\partial u}{\partial x_i} + c(\vec{x}) u = f(\vec{x}). \quad (1.31)$$

Тип такого уравнения целиком определяется свойствами характеристической формы, выражаемой через коэффициенты главной части дифференциального оператора L :

$$Q(\vec{x}) = \sum_{i,j=1}^n a_{i,j}(\vec{x}) y_i y_j, \quad \|y\| \neq 0,$$

где $y_1, \dots, y_n$ — любые вещественные числа и $\|y\| = (\sum_{i=1}^n y_i^2)^{1/2}$. Здесь без ограничения общности квадратичную форму $Q(\vec{x})$ можно считать в любой точке $\vec{x}$ симметричной, т.е. $a_{ij}(\vec{x}) = a_{ji}(\vec{x})$. Определим теперь квадратную симметричную матрицу порядка n с зависящими от x элементами:

$$A(\vec{x}) = \{a_{i,j}(\vec{x})\}.$$

Если в точке $\vec{x}$ все собственные значения матрицы $A(\vec{x})$ одного знака (будем считать в этом случае их для определенности положительными), то уравнение (1.31) и дифференциальный оператор L называются *эллиптическими* в точке $\vec{x}$; если одно или несколько собственных чисел $A(\vec{x})$ отрицательны, а остальные — положительны, то уравнение называется *гиперболическим* или *ультрагиперболическим* в точке x ; если среди собственных чисел матрицы $A(\vec{x})$ имеются нулевые, то дифференциальное уравнение называется *параболическим* в точке $\vec{x}$.

Для характеристической формы эллиптического уравнения, очевидно, можно записать

$$Q(\vec{x}) = \sum_{i,j}^n a_{i,j} y_i y_j \geq \alpha(\vec{x}) \|y\|^2, \quad \alpha(\vec{x}) > 0. \quad (1.32)$$

Определение 1.33. Уравнение (1.31) называется эллиптическим в области Ω , если оно эллиплично в каждой точке $\vec{x} \in \Omega$.

Определение 1.34. Дифференциальное уравнение (1.31) называется *равномерно эллиптическим* в области Ω , если в условии (1.32)

$$\inf_{x \in \Omega} \alpha(\vec{x}) = \alpha_0 > 0, \quad (1.33)$$

где α_0 — не зависящая от $\vec{x}$ постоянная.

Рассмотрим теперь линейное дифференциальное уравнение в частных производных порядка $2m$ при $m > 1$:

$$\sum_{k=0}^{2m} \sum_{i_1, \dots, i_k=1}^n a_{i_1, \dots, i_k}(\vec{x}) \frac{\partial^k u}{\partial x_1 \dots \partial x_{i_k}} = f(\vec{x}). \quad (1.34)$$

Это уравнение называется эллиптическим в точке $\vec{x}$, если

$$Q(\vec{x}) = \sum_{i_1 \dots i_{2m}=1}^n a_{i_1 \dots i_{2m}}(\vec{x}) y_{i_1} \dots y_{i_{2m}} \neq 0 \quad (1.35)$$

при любых значениях y_i таких, что $\|y\| \neq 0$.

Так же, как и в случае уравнений второго порядка, можно определить понятие эллиптического в области и равномерно эллиптического в области дифференциального уравнения вида (1.34).

Характерной иллюстрацией эллиптического уравнения высокого порядка является *бигармоническое уравнение*

$$\Delta^2 u \equiv \frac{\partial^4 u}{\partial x_1^4} + 2 \frac{\partial^4 u}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4 u}{\partial x_2^4} = f(\vec{x}). \quad (1.36)$$

Аналогично определяются и *полигармонические уравнения* с целочисленными степенями оператора Лапласа $\Delta^\gamma, \gamma > 2$.

Перейдем далее к системе уравнений в частных производных, содержащей p неизвестных функций $u_1, \dots, u_p$. Обозначая через m_j наивысший порядок производных функции u_j , $1 \leq j \leq p$, такую систему можно записать в виде

$$\sum_{j=1}^p \sum_{k=1}^{m_j} \sum_{i_1 \dots i_k=1}^n a_{i_1 \dots i_k}^{(i,j)} \frac{\partial^k u_j}{\partial x_{i_1} \dots \partial x_{i_k}} = f_i(\vec{x}), \quad (1.37)$$

$$i = 1, \dots, p.$$

Выпишем определитель порядка p , у которого на пересечении i -й строки и j -го столбца находится элемент

$$A^{(i,j)}(\vec{x}) = \sum_{i_1 \dots i_{m_j}=1}^n a_{i_1 \dots i_{m_j}} y_{i_1} \dots y_{i_{m_j}}. \quad (1.38)$$

Определение 1.35. Система (1.37) называется эллиптической в точке $\vec{x} \in \Omega$, если

$$\det |A^{(i,j)}(\vec{x})| \neq 0 \text{ при } \|y\| \neq 0. \quad (1.39)$$

Характерным примером эллиптической системы служит система двумерных уравнений Ламе линейной теории упругости при $p = m_1 = m_2 = n = 2$, см. [90]:

$$\begin{aligned} (\lambda + 2\mu) \frac{\partial^2 u_1}{\partial x_1^2} + \mu \frac{\partial^2 u_1}{\partial x_2^2} + (\lambda + \mu) \frac{\partial^2 u_2}{\partial x_1 \partial x_2} &= f_1(x_1, x_2), \\ (\lambda + \mu) \frac{\partial^2 u_1}{\partial x_1 \partial x_2} + \mu \frac{\partial^2 u_2}{\partial x_1^2} + (\lambda + 2\mu) \frac{\partial^2 u_2}{\partial x_2^2} &= f_2(x_1, x_2), \end{aligned} \quad (1.40)$$

определитель (1.39) для которой равен

$$\begin{vmatrix} (\lambda + 2\mu)y_1^2 + \mu y_2^2 & (\lambda + \mu)y_1 y_2 \\ (\lambda + \mu)y_1 y_2 & \mu y_1^2 + (\lambda + 2\mu)y_2^2 \end{vmatrix} = \mu(\lambda + 2\mu)\|y\|^4.$$

Здесь u_1 и u_2 , f_1 и f_2 — компоненты вектора смещений, или перемещений, $\bar{u} = (u_1, u_2)^T$, $\bar{f} = (f_1, f_2)$, а λ и μ — коэффициенты Ламе, удовлетворяющие условиям

$$0 < \mu_0 \leq \mu \leq \mu_1, \quad 0 < \lambda < \infty.$$

В общем случае, включающем трехмерный, уравнения стационарной упругости (в терминах смещений) могут быть записаны в операторной форме [48]

$$-\operatorname{div} \bar{\sigma}(\bar{u}) = \bar{f}, \quad (1.41)$$

где $\bar{\sigma}(\bar{u})$ есть тензор напряжений, определяемый по закону Гука для изотропного тела следующей формулой:

$$\bar{\sigma}(\bar{u}) = 2\mu \bar{\varepsilon}(\bar{u}) + \operatorname{tr}(\bar{\varepsilon}(\bar{u})) \bar{\delta}. \quad (1.42)$$

Здесь, в свою очередь, $\bar{\varepsilon}(\bar{u}) = \{\varepsilon_{i,j}(\bar{u})\}$ — симметричный тензор деформаций с компонентами

$$\varepsilon_{i,j}(\bar{u}) = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right),$$

$\bar{\delta} = \{\delta_{i,j}\}$ есть единичный тензор ($\delta_{i,j}$ — символ Кронекера), а $\operatorname{tr}(\bar{\varepsilon}) = \sum_i \varepsilon_{i,i}$ — след матрицы. Отметим также, что дивергенция тензора — это вектор с компонентами

$$(\operatorname{div} \bar{\sigma})_i = \sum_j \frac{\partial \sigma_{i,j}}{\partial x_j}.$$

Характерные с физической точки зрения граничные условия будут указаны в следующем пункте.

1.3.3. Типы краевых условий

Выше мы уже коснулись вопроса классификации краевых условий. Так, в нестационарных задачах, описываемых уравнениями параболического и гиперболического типов, выделяются «пространственные граничные условия» и «начальные данные». Необходимо подчеркнуть, что только после их задания конкретизируются свойства оператора решаемой математической задачи.

Определение типов граничных условий, обеспечивающих корректность краевой задачи, связано, конечно, с видом рассматриваемого дифференциального уравнения, и на этих моментах мы подробно останавливаться не будем. Наиболее просто классификация краевых задач проводится для уравнения Пуассона, и нами уже был приведен один пример — задачи первого рода (Дирихле), когда на всей границе Γ расчетной области Ω , в которой определено решение, задано граничное условие соответствующего наименования.

Если для решения уравнения Пуассона на всей границе области Ω задано *условие второго рода (Неймана)*

$$\frac{\partial u}{\partial \nu} = g(\vec{x}), \quad \vec{x} \in \Gamma,$$

(здесь $\nu = (\nu_1, \dots, \nu_n)$ означает единичную внешнюю нормаль к Γ) или *условие третьего рода (Ньютона, или Робена)*

$$\alpha u + \frac{\partial u}{\partial \nu} = g(\vec{x}), \quad \alpha \neq 0,$$

то мы получаем, соответственно, вторую или третью краевую задачу. Если же граница области $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$ содержит участки Γ_k , на которых заданы условия различных типов, то такая краевая задача называется смешанной.

Отметим, что в задаче Неймана для уравнения Пуассона $\Delta u = f$, например, решение не является единственным и определено с точностью до константы, что легко проверяется непосредственно. А условием его разрешимости является выполнение равенства

$$\int_{\Omega} f dv = \int_{\Gamma} g dS. \quad (1.43)$$

Однако при наличии участков границы Γ_1 , и/или Γ_3 с условиями Дирихле или Робена (при $\alpha < 0$ на Γ_3) решение существует и является единственным.

В случае, когда решается дифференциальное уравнение второго порядка с главной частью так называемого дивергентного вида

$$Lu = - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{i,j}(\vec{x}) \frac{\partial u}{\partial x_j} \right) + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} + cu = f(\vec{x}), \quad (1.44)$$

в граничных условиях второго или третьего рода вместо производной по нормали фигурирует *конормальная производная*

$$\frac{\partial u}{\partial N} \equiv \sum_{i,j=1}^n a_{i,j} \frac{\partial u}{\partial x_j} \cos(\nu, x_j),$$

которая совпадает с $\frac{\partial u}{\partial \nu}$ при $a_{i,j}(\vec{x}) = 1$ для всех значений индексов i, j .

Укажем еще на один вид характерной краевой задачи: для бигармонического уравнения (1.21) первой краевой задачей называется такая, в которой на всей границе заданы условия

$$u(\vec{x}) = g_0(\vec{x}), \quad \frac{\partial u(\vec{x})}{\partial \nu}, \quad \vec{x} \in \Gamma. \quad (1.45)$$

Относительно самой границы Γ обычно предполагается, что она гладкая или кусочно-гладкая. Для нас, однако, наиболее общим и интересным с прикладной точки зрения является липшицева граница.

Для векторной задачи теории упругости (1.41), (1.42) типичные краевые условия заключаются в задании вектора смещений на Γ_1 :

$$\bar{u}|_{\Gamma_1} = \bar{g}, \quad (1.46)$$

а также нормальных компонент напряжений на Γ_2 :

$$(\bar{\sigma}(\bar{u})\bar{\nu})_{\Gamma_2} = \bar{t}. \quad (1.47)$$

Примерами достаточно необычных постановок для векторных краевых задач являются рассмотренные далее в п. 1.3.7 граничные условия для системы уравнений Максвелла.

В “хороших” краевых или начально-краевых задачах граничные условия определяют вопросы существования, единственности, а также гладкости искомого решения, необходимой для обоснования конструируемых и применяемых численных методов. Особый класс представляют *сингулярно-возмущенные задачи*, в которых решение имеет особенности, связанные с неограниченностью производных в некоторых окрестностях. Эти сингулярности обуславливаются или наличием входящих углов на граничных поверхностях, или вырождением коэффициентов при старших производных в дифференциальных уравнениях. Примерами могут служить так называемые *задачи с пограничными или внутренними слоями*. Построение и обоснование алгоритмов для таких задач требует изучения асимптотических свойств решений вблизи локализации сингулярностей. Эти специальные вопросы составляют актуальное направление теории дифференциальных и интегральных уравнений, и только при квалифицированном использовании имеющихся аналитических оценок можно адекватно решать такого типа задачи.

До сих пор мы подразумевали, что расчетная область Ω является ограниченной, что определяет постановку *внутренних краевых задач*. Однако во многих приложениях требуется решать *внешние краевые задачи*, связанные с заданием условий для решения на бесконечности.

Рассмотрим свойства внешних краевых задач для *гармонических функций*, удовлетворяющих уравнению Лапласа $\Delta u = 0$.

Определение 1.36. Гармоническая функция трех переменных $u(x_1, x_2, x_3)$ называется регулярной в бесконечности, если

$$|u| \leq A/r, \quad |\partial u / \partial x_i| \leq A/r^2, \quad i = 1, 2, 3,$$

при достаточно большом $r \geq r_0$, где $r^2 = \sum_{i=1}^3 (x_i - x_i^0)^2$, а x_i^0 – координаты начала декартовой системы координат.

Внешние краевые задачи для регулярных в бесконечности гармонических функций по-разному ставятся для трех и двух независимых переменных. Пусть в трехмерном случае область Ω является внешней к некоторой замкнутой поверхности Γ . Тогда первая *внешняя краевая задача*, или внешняя задача Дирихле,

состоит в следующем: найти функцию $u(x_1, x_2, x_3)$, непрерывную всюду в $\bar{\Omega} = \Omega \cup \Gamma$, равномерно стремящуюся к нулю на бесконечности и удовлетворяющую условиям

$$\begin{aligned}\Delta u &= 0, \quad (x_1, x_2, x_3) \in \Omega, \\ u|_{\Gamma} &= g(x_1, x_2, x_3),\end{aligned}$$

где g – заданная на поверхности Γ функция. Сформулированная таким образом задача имеет единственное решение, регулярное в бесконечности.

Первая краевая задача на плоскости, т.е. в двумерном случае, формулируется для гармонической функции аналогично, за одним существенным отличием: решение $u(x_1, x_2)$ является ограниченным всюду в Ω , в том числе на бесконечности. Именно такая постановка обеспечивает существование единственного решения.

Если граница области состоит из участков с разными типами краевых условий, т.е. $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, то по аналогии с внутренними краевыми задачами определяются внешние задачи Неймана или Робена, или же внешние смешанные краевые задачи. При этом опять же, независимо от размерности области, внешняя задача Неймана имеет решение только при выполнении условия разрешимости (1.43), и это решение определено с точностью до произвольной постоянной.

В качестве более общей задачи в неограниченных областях рассмотрим более общее уравнение

$$\Delta u = pu_{tt} + qu_t - \mathcal{F}(\vec{r}, t), \quad p, q > 0,$$

где p и q – заданные числа или функции, которое описывает процессы вынужденных колебаний с затуханием. Будем считать, что $\mathcal{F}(\vec{r}, t)$ является локальной, или финитной, т.е. отличной от нуля в некоторой ограниченной области, а также периодической по времени, т.е. $\mathcal{F}(\vec{r}, t) = f(\vec{r})e^{i\omega t}$, где i есть мнимая единица, ω – частота колебаний, а функция пространственных аргументов $f(\vec{r})$ называется амплитудой колебаний функции источника. В этом случае, очевидно, решение задачи представимо в виде

$$u(\vec{r}, t) = v(\vec{r})e^{i\omega t},$$

где функция $v(\vec{r})$ удовлетворяет уравнению Гельмгольца,

$$\Delta v + k^2 v = f(\vec{r}).$$

Здесь коэффициент

$$k^2 = \omega^2 p - iq\omega$$

в общем случае является комплексным, а искомое решение $v(\vec{r})$ зависит от частоты ω параметрическим образом. Сделанное преобразование функции $u(\vec{r}, t)$ в функцию $v(\vec{r}, \omega)$, которые удовлетворяют формально нестационарному и стационарному уравнениям соответственно, называется переходом из временной области в частотную область.

Решение трехмерного уравнения Гельмгольца в неограниченной области с локальной функцией $f(\vec{r})$ существует, является единственным и удовлетворяет на бесконечности *условиям излучения Зоммерфельда*

$$v = O(|\vec{r}|^{-1}), \quad \frac{\partial v}{\partial r} + ikv = O(|\vec{r}|^{-1}).$$

Отметим, что первое из приведенных здесь условий в действительности является следствием второго.

При численном решении внешних краевых задач для дифференциальных уравнений возникает проблема аппроксимации условий на бесконечности, поскольку реально расчеты можно проводить только в ограниченной области. Простейший применяемый в данном случае прием – на достаточно большом, но конечном, расстоянии R от интересующей нас окрестности положить решение равным нулю и решить полученную внутреннюю краевую задачу. Дилемма здесь заключается в том, что если значение R выбрано недостаточно большим, то решение внутренней задачи будет слишком грубым приближением, а в противном варианте (R очень велико) получаемая внутренняя задача потребует для расчетов слишком больших вычислительных ресурсов.

Для преодоления указанных трудностей различными авторами были предложены варианты так называемых *абсорбционных*, или радиационных, *граничных условий*, идея которых заключается в более точном приближении условий Зоммерфельда. Их реализация состоит в использовании асимптотических свойств решения внешней задачи на больших расстояниях. Мы на этих специальных вопросах останавливаться не будем и отсылаем заинтересованного читателя к монографиям А. Д. Григорьева [52]

и П. Монка [133] по электромагнетизму, где эти вопросы особенно актуальны и хорошо исследованы.

Дальнейшим развитием методов решения внешних краевых задач явилась концепция искусственного построения *идеально согласованного слоя* (Perfectly Matched Layer, PML). Его назначением, как и абсорбционных граничных условий, является, на физическом языке, поглощение всей падающей на него энергии без отражений. Здесь математическая технология заключается в подборе геометрических и материальных свойств формируемого слоя, что отображается в типе определяемых дополнительных уравнений для PML.

1.3.4. Задачи тепло-массопереноса и фильтрации

В данном пункте мы фактически проиллюстрируем большой класс задач механики сплошной среды, связанных с переносом различных субстанций (тепла, массы и др.) в природных условиях и технических устройствах. Такого типа математические модели используются и в других, самых неожиданных, сферах человеческой деятельности: экономика, социология, транспортные задачи и т.д. Классические и более специальные постановки здесь описываются уравнениями тепло-массопереноса и Дарси, при разных уровнях детализации материальных свойств среды на микро- и макроуровне. Фундаментальным исследованиям соответствующих процессов посвящены монографии Г.И. Баренблатта [15], П. Н. Кочиной [16] с соавторами, О. М. Белоцерковского [20], П. Н. Вабищевича [30], С. К. Годунова и Е.И.Роменского [48], А. Н. Коновалова [91], С. Патанкара [142], В. В. Пененко [143], А. М. Франка [189], Н. Н. Яненко [209] с коллегами и многих других авторов.

Скалярное диффузионно-конвективное уравнение

$$-\frac{\partial}{\partial x}\left(a\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(b\frac{\partial u}{\partial y}\right) + p\frac{\partial u}{\partial x} + q\frac{\partial u}{\partial y} + ru = f, \quad (1.48)$$

которое мы для простоты рассматриваем в стационарном двумерном варианте, содержит три типа членов: диффузионные, конвективные и источники, которые описываются соответственно слагаемыми, содержащими производные второго, первого и нулевого порядков. Диффузионные коэффициенты a и b , конвективные коэффициенты p и q (зачастую имеющие смысл скорости среды), источниковый коэффициент r и свободный член

(правая часть уравнения), f – все они могут быть или константами, или зависеть от независимых переменных x, y , или быть нелинейными, т.е. включать зависимость от искомого решения u . Отметим, что в (1.48) диффузионные члены записаны в так называемом дивергентном виде с анизотропными коэффициентами. Типичный физический смысл уравнения (1.48), (??) связан с описанием распределения температуры или концентрации какой-то примеси в непрерывной движущейся среде.

Для более общего, в сравнении с (1.48), стационарного диффузионно-конвективного уравнения

$$-\nabla \cdot (\varepsilon A \nabla u - \vec{b}u) + au = f(\vec{x}), \quad \vec{x} \in \Omega, \quad (1.49)$$

мы приведем описание и некоторые свойства смешанной краевой задачи в соответствии с работой Р.З.Даутова, Е.М.Федотова “Гибридизированные схемы разрывного метода Галеркина для стационарных задач конвекции-диффузии”, Дифференциальные уравнения, т. 52, N 7, 2016. 1-19.

В (1.49) расчетная область $\Omega \in R^d$, $d = 2, 3$, предполагается многогранником с границей $\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3$, на разных (непересекающихся) частях которой заданы краевые условия 1-го, 2-го или 3-го рода в следующей форме:

$$\begin{aligned} u|_{\Gamma_1} &= g_1(\vec{x}), \quad (\varepsilon A \nabla u) \cdot \vec{n} = g_2(\vec{x}), \quad \vec{x} \in \Gamma_2, \\ (\varepsilon A \nabla u - \vec{b}u) \cdot \vec{n} &= g_3(\vec{x}), \quad \vec{x} \in \Gamma_3, \end{aligned} \quad (1.50)$$

где $\vec{n}$ есть поле единичных нормалей на Γ , $g_1(\vec{x})$ – след некоторой функции $u_D(\vec{x}) \in H^1(\Omega)$, а функции $g_2(\vec{x})$, $g_3(\vec{x})$ принадлежат $L_2(\Gamma_2)$ и $L_2(\Gamma_3)$ соответственно, причем множество Γ_1 не пусто ($\text{meas}(\Gamma_1) > 0$).

Будем считать, что участвующие в (1.49), (1.50) матрица A , вектор $\vec{b}$ и другие величины удовлетворяют условиям

$$\begin{aligned} A &\in [L_\infty(\Omega)]^{d \times d}, \quad \vec{b} \in [L_\infty(\Omega)]^d, \quad \nabla \cdot \vec{b}, a \in L_\infty(\Omega), \\ f(\vec{x}) &\in L_2(\Omega), \quad m_A(\xi, \xi) \leq (A(\vec{x})\xi, \xi) \leq M_A(\xi, \xi), \\ \xi &\in R^d, \quad a(\vec{x}) + 0.5 \nabla \vec{b}(\vec{x}) \geq m_a, \quad \vec{x} \in \Omega; \quad \vec{b} \cdot \vec{n} \geq 0, \quad \vec{x} \in \Gamma_2, \end{aligned} \quad (1.51)$$

где $m_A > 0$, $m_a \geq 0$, а ε есть положительный параметр, введенный для сравнительной характеристики диффузионного и остальных членов исходного уравнения.

Задача (1.49), (1.50) при условиях (1.51) имеет, и притом единственное, обобщенное решение

$$u \in u_D + V_0, \quad V_0 = \{v \in H^1(\Omega) : v|_{\Gamma_1} = 0\}, \quad (1.52)$$

удовлетворяющее вариационному равенству

$$a(u, v) = f(v) \quad \forall v \in V_0 \quad (1.53)$$

где билинейная и линейная формы определяются следующим образом:

$$\begin{aligned} a(u, v) = & \int_{\Omega} (\varepsilon A \nabla u \cdot \nabla v - \vec{b}u \cdot \nabla v) + (auv) d\vec{x} + \\ & \int_{\Gamma_2} \vec{b} \cdot \vec{n} uv d\vec{x}, \quad f(v) = \int_{\Omega} f v d\vec{x} + \int_{\Gamma_2} g_2 v d\vec{x} + \int_{\Gamma_3} g_3 v d\vec{x}, \end{aligned} \quad (1.54)$$

причем для соответствующей квадратичной формы имеем

$$a(v, v) \geq \varepsilon m_A \|\nabla v\|_{L_2(\Omega)}^2, \quad \forall v \in V_0.$$

Классической постановке диффузионно-конвективной задаче можно сопоставить смешанную формулировку, описываемую системой дифференциальных уравнений первого порядка

$$-\nabla \cdot \vec{q} + \nabla \cdot (\vec{b}u) = f, \quad \vec{q} = \varepsilon A \vec{\sigma}, \quad \vec{\sigma} = \nabla u. \quad (1.55)$$

Введенный здесь вектор потока $\vec{q}$ принадлежит пространству

$$H_2(\text{div}; \Omega) = \{\vec{q} \in [L_2(\Omega)]^d : \text{div } \vec{q} \in L_2(\Omega)\}. \quad (1.56)$$

При этом для пространств Соболева векторных функций $\vec{p} = (p_1, \dots, p_d)^T$ с компонентами p_k из скалярных соболевских пространств $H^s(\Omega)$ порядка s используется естественное обозначение $[H^s(\Omega)]^d$, нормы и полунормы при целых $s \leq 0$ в них определяются аналогично скалярному случаю:

$$\|\vec{p}\|_{s, \Omega} = \left(\sum_{l=0}^s |\vec{p}|_{l, \Omega}^2 \right)^{1/2}, \quad |\vec{p}|_{s, \Omega} = \left(\sum_{i=1}^d \int_{\Omega} \sum_{|\alpha|} = l |D^\alpha p_i|^2 d\vec{x} \right)^{1/2}. \quad (1.57)$$

Обозначение для скалярного произведения $(\cdot, \cdot)_{\Omega}$ используется одинаково как для скалярных функций $u, v \in L_2(\Omega)$ так и для векторов $\vec{p}, \vec{q} \in [L_2(\Omega)]^d$:

$$(u, v)_\Omega = \int_\Omega uv d\vec{x}, \quad (\vec{p}, \vec{q}) = \int_\Omega \vec{p} \cdot \vec{q} d\vec{x}.$$

Очевидно, что уравнение (1.49) можно обобщить на нестационарный случай с произвольной пространственной размерностью d :

$$\begin{aligned} c \partial u / \partial t - \nabla \cdot (\varepsilon A \nabla u - \vec{b}u) + au &= f(\vec{x}, t), \\ (x, t) \in Q &= \Omega \times (0, t_e), \quad t_e > 0, \\ u(\vec{x}, 0) &= u^0(\vec{x}), \quad \vec{x} \in \Omega. \end{aligned} \quad (1.58)$$

В этом случае участвующие в (1.58) величины предполагаются удовлетворяющими условиям вида (1.51), в которых надо только заменить Ω на Q и добавить

$$c \in L_\infty(Q), \quad u^0 \in L_2(\Omega).$$

При этом, в силу появления новой независимой переменной t , надо, дополнительно к используемым в (1.52)–(1.57) пространствам и нормам, ввести пространства $L_p(a, b; [H^3(\Omega)]^d)$, $H^p(a, b; [H^s(\Omega)]^d)$ измеримых функций $u(t) \in \{(a, b) \rightarrow [H^s(\Omega)]^d\}$ с нормами

$$\begin{aligned} \|u\|_{L_2(a, b; [H^s(\Omega)]^d)} &= \left(\int_a^b \|u(t)\|_{s, \Omega}^2 dt \right)^{1/2}, \\ \|u\|_{L_\infty(a, b; [H^s(\Omega)]^d)} &= \sup_{t \in (a, b)} \|u(t)\|_{s, \Omega}, \\ \|u\|_{H^p(a, b; [H^s(\Omega)]^d)} &= \|(\|u(t)\|)\|_{p, (a, b)}. \end{aligned}$$

Если в (1.66) искомое решение есть температура ($u = T$), то коэффициент c имеет смысл теплоемкости среды, а соотношение (1.66) называется уравнением теплопроводности. Мы подробнее рассмотрим уравнение теплопроводности для актуальной задачи Стефана с фазовыми переходами, когда при $T \leq T_f$ среда является жидкой (ликвидус), а при $T > T_f$ – твердой (солидус), причем положим для простоты $\vec{b} = a = 0, A = k$ (скалярный коэффициент теплопроводности):

$$c \frac{\partial T}{\partial t} = \operatorname{div}(k \operatorname{grad} T) + f.$$

В данном случае при $T < T_f$ имеем $c = c_s, k = k_s$, а в противном случае полагаем $c = c_l, k = k_l$, где теплофизические параметры

“ликвидуса” и “солидуса” могут зависеть от температуры, а на границе раздела фаз, являющейся изотермой $T = T_f$, справедливо уравнение баланса энергии (*условие Стефана*):

$$k_s(\partial T/\partial n)_s - k_l(\partial T/\partial n)_l = \rho\lambda v_n,$$

где λ – удельная теплота плавления, v_n – нормальная к поверхности раздела фаз составляющая скорости движения данной поверхности, а нормаль $\vec{n}$ выбирается внешней по отношению к твердой фазе.

В более современной энтальпийной постановке распределение температуры определяется уравнением

$$\frac{\partial H}{\partial t} = \operatorname{div}(k \operatorname{grad} T) + f,$$

где коэффициент теплопроводности k в анизотропных средах представляется тензором, а H – объемная энтальпия, определяемая формулой

$$H(T) = \int_{T_0}^T c(T) dT + \theta L.$$

Здесь c есть объемная теплоемкость, L – объемная теплота фазового перехода, T_0 – произвольно выбранное значение температуры, а T_f – температура фазового перехода. При этом в “жидкой” фазе ($T \geq T_f$) мы имеем $\theta = 1, c = c_l, k = k_l$, а в твердой ($T < T_f$) $\theta = 0, c = c_s, k = k_s$. По известной энтальпии температура находится следующим образом:

$$T = \begin{cases} H \leq H_2; & T = \frac{H - H_2}{c_s} + T_f, \\ H_2 < H < H_1 : & T = T_f, \\ H \geq H_1 : & T = \frac{H - H_1}{c_l} + T_f. \end{cases}$$

В данной модели фактически энтальпийный интервал $H \in [H_1, H_2]$ определяет переходную двухфазную зону.

Уравнения Дарси для течений в пористой среде формально являются некоторым обобщением рассмотренных уравнений (1.49), (1.58), однако они связаны с изучением очень важных процессов, имеющих свои исключительные особенности, которые дают все основания вынести их в самостоятельный класс задач.

Мы приведем в достаточно общей форме трехмерные уравнения многофазной изотермической (при постоянной температуре) модели фильтрации вязкой сжимаемой газо-жидкостной смеси в анизотропной пористой среде (ниже используются декартовые системы координат x, y, z). Более конкретно, мы описываем постановку для трехфазной среды, актуальной при изучении процессов нефтедобычи, где $n_l = 3$, а $l = 1, 2, 3$ соответствуют воде, нефти и газу. В данной модели “черной (нелетучей) нефти” предполагается, что вода и нефть не смешиваются между собой, а газ растворяется в воде и нефти.

$$\begin{aligned} \frac{\partial}{\partial t}(\varphi N_c) &= \operatorname{div} \sum_{l=1}^{n_l} \eta_{c,l} \xi_l \left(\bar{k} \frac{k_{c,l}}{\mu_l} (\Delta p_l - \gamma_l \Delta D) \right) + q_c, \\ c &= 1, \dots, n_c; \quad p_2 - p_3 = P_{c,2,3}, \quad p_2 - p_1 = P_{c,2,1}, \\ S_1 + S_2 + S_3 &= 1, \quad N_1 = \frac{S_1}{B_1}, \quad N_2 = \frac{S_2}{B_2}, \quad N_3 = \frac{S_3}{B_3} + R_{3,2} \frac{S_2}{B_2} \end{aligned} \quad (1.59)$$

где n_l и n_c – количество фаз и число компонентов в смеси, $p_l = p_l(t, x, y, z)$ – давление в l -й фазе, $B_l = B_l(p_l)$ – коэффициент объемного расширения фазы l , $\vec{N} = (N_1, \dots, N_{n_c})$ – вектор молярных плотностей, $\eta_{c,l} = \eta_{c,l}(p_1, \dots, p_{n_l}, \vec{N})$ – молярная доля c -го компонента в l -й фазе, $\xi_l = \xi_l(p_l, \vec{N})$ – молярная плотность фазы l , $\bar{k} = \bar{k}(p_1, \dots, n_l)$ – тензор абсолютной проницаемости, $\mu_l = \mu_l(p_l)$ – вязкость фазы l , $S_l = S_l(t, x, y, z)$ – насыщенность l -й фазы, $q_c = q_c(p_1, \dots, p_{n_l}, \vec{N}, t, x, y, z)$ – источник компоненты c , $\varphi = \varphi(p_1, \dots, p_n, x, y, z)$ – пористость среды, $R_{3,2} = R_{3,2}(p_2)$ – растворимость газа в нефтяной фазе, $K_{c,l} = K_{c,l}(S_1, S_3)$ – относительная проницаемость фазы l , $\gamma_l = \rho_l g$ – вертикальный градиент гидростатического давления в фазе l , g – гравитационная постоянная, $\vec{D} = \vec{D}(x, y, z)$ – вектор глубины (сверху вниз), $\rho_l = \rho_l(p_l)$ – массовая плотность фазы l , $P_{c,2,3} = P_{c,2,3}(S_3)$ – капиллярное давление в системе нефть-газ, $P_{c,2,1} = P_{c,2,1}(S_1)$ – капиллярное давление в системе вода-нефть.

В системе (1.59) неизвестными являются функции N_c, S_l и p_l , а в самом распространенном случае количество компонент n_c равно трем. Для рассматриваемой проблемы задаются начальные значения искомого решения, а на внешней границе моделируемого резервуара ставятся граничные условия постоянного давления (условия Дирихле) или заданного потока (условия Неймана).

Как видно, описанная модель представляет нетривиальную математическую проблему, отягчаемую, как правило, сложной геометрией расчетной области (реальный пример – куст криволинейных скважин в сложной геологической структуре), с сильной разномасштабностью и контрастными свойствами материальных сред, которые на практике известны очень приближенно и по-хорошему требуют еще решения обратной задачи по идентификации параметров модели на основе сравнения расчетных данных с геофизическими измерениями.

1.3.5. Задачи гидро-газодинамики

Проблемы моделирования течений газа или жидкости являются, с одной стороны, одними из классических в механике сплошных сред, а с другой – экстремально сложными в силу наличия сингулярностей математических моделей и свойств их решений, без учета которых конструируемые численные методы не смогут обеспечить необходимого для практики высокого разрешения расчетных результатов. Данные задачи включают динамические нелинейные процессы с аттракторами, бифуркациями и катастрофами, с разнообразными типами разрывов решений и их режимов (ламинарные и турбулентные, дозвуковые и гиперзвуковые), с многогранностью материальных уравнений состояния, и т.д.

Изучение этих явлений составляет много ключевых направлений прикладной и вычислительной математики, которым посвящено огромное количество научных конференций, статей и монографий, среды которых мы назовем книги (в алфавитном порядке) Н.Н.Булеева [29], П.Н.Вабищевича [30], С.К.Годунова с соавторами [46], Б.Л.Рождественского и Н.Н.Яненко [151], П.Роуча [152], А.А.Самарского и Ю.П.Попова [164], А.И.Толстых [178], Ю.И.Шокина и Н.Н.Яненко с коллегами [195], [205].

Целью настоящего пункта является предварительное ознакомление читателя с тем богатством идей, которое заложено в постановки данных задач, актуальнейших для понимания естественно-природных и научно-технических приложений. Изучение естественно начать с простейших одномерных моделей, на которых уже можно понять своеобразие рассматриваемых процессов.

Так называемое *акустическое приближение* представляет со-

бой систему линейных дифференциальных уравнений

$$\begin{aligned}\frac{\partial u}{\partial t} + \frac{1}{\rho_0} \frac{\partial p}{\partial x} &= 0, \\ \frac{\partial p}{\partial t} + \rho_0 c_0^2 \frac{\partial u}{\partial x} &= 0,\end{aligned}\tag{1.60}$$

описывающее распространение плоских звуковых волн. В (1.60) функции u и p суть скорость и давление, а говоря точнее, отклонения скорости и давления от их значений в невозмущенной среде, вызванные распространением волн (в предположении малости самих возмущений). Постоянные ρ_0 и c_0 – это плотность среды и скорость звука, характеризующаяся сжимаемостью среды.

Если ввести *римановы инварианты*

$$Y = u + \frac{p}{\rho_0 c_0}, \quad Z = u - \frac{p}{\rho_0 c_0},\tag{1.61}$$

то исходная система приводится к виду

$$\frac{\partial Y}{\partial t} + c_0 \frac{\partial Y}{\partial x} = 0, \quad \frac{\partial Z}{\partial t} - c_0 \frac{\partial Z}{\partial x} = 0.\tag{1.62}$$

Отсюда очевидным образом следует общее решение

$$Y = f(x - c_0 t), \quad Z = g(x + c_0 t),$$

где f и g – произвольные дифференцируемые функции. Переходя с помощью (1.61) к исходным функциям u и p , получаем общее решение для уравнений акустики (1.60):

$$\begin{aligned}u &= [f(x - c_0 t) + g(x + c_0 t)]/2, \\ p &= \rho_0 c_0 [f(x - c_0 t) - g(x + c_0 t)]/2.\end{aligned}\tag{1.63}$$

Линии $x \pm c_0 t = \text{const}$, или же $\frac{dx}{dt} = \pm c_0$, называются *характеристиками* системы (1.60), и вдоль них остаются постоянными величины Z, Y соответственно. Если решение акустической задачи удовлетворяет начальным данным

$$u(x, 0) = u_0(x), \quad p(x, 0) = p_0(x),\tag{1.64}$$

то функции f и g в (1.63) определяются как

$$f(x) = u_0(x) + \frac{p_0(x)}{\rho_0 c_0}, \quad g(x) = u_0(x) - \frac{p_0(x)}{\rho_0 c_0}.\tag{1.65}$$

Несложно показать, что если начальные условия (1.64) заданы дифференцируемыми на отрезке $x \in [a, b]$, то система (1.60) будет обладать единственным решением внутри *характеристического треугольника*, образуемого прямыми

$$t = 0, \quad x - c_0 t = a, \quad x + c_0 t = b.$$

Если же начальные данные даны разрывными, т.е.

$$u_0(x) = \begin{cases} u_1, & x < x^*, \\ u_2 & x > x^*, \end{cases} \quad p_0(x) = \begin{cases} p_1, & x < x^*, \\ p_2, & x = x^*, \end{cases}$$

где постоянные u_1, u_2 и p_1, p_2 не равны попарно одновременно, то получим, что функции решения $u(x, t), p(x, t)$ имеют разрывы вдоль линий $x + c_0 t = x^*$ и $x - c_0 t = x^*$, в силу чего такая задача называется *задачей о распаде разрыва*.

Более сложная модель описывается квазилинейной системой *уравнений газодинамики*, представляющей собой уравнения непрерывности, движения и энергии, а также уравнение состояния:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} &= 0, \\ \frac{\partial(\rho u)}{\partial t} + \frac{\partial(p + \rho u^2)}{\partial x} &= 0, \\ \frac{\partial \rho(\varepsilon + u^2/2)}{\partial t} + \frac{\partial[\rho u(\varepsilon + u^2/2) + pu]}{\partial x} &= 0, \\ p &= p(\rho, \varepsilon) \quad \text{или} \quad \varepsilon = \varepsilon(p, \rho). \end{aligned} \quad (1.66)$$

Здесь ρ – плотность, p – давление, ε – внутренняя энергия единицы масса газа, u – скорость. В случае идеального газа уравнения состояния имеют следующий простой вид:

$$p = (\kappa - 1)\rho\varepsilon, \quad \varepsilon = \frac{p}{(\kappa - 1)\rho},$$

где величина κ называется показателем адиабаты.

Уравнение состояния может быть также задано в форме зависимости

$$\varepsilon = \varepsilon(V, S),$$

где $V = 1/\rho$ есть удельный объем, а S – энтропия. Эти величины связаны с давлением p и абсолютной температурой T термодинамическим тождеством

$$d\varepsilon + p dV = T dS.$$

При больших скоростях возможны такие движения, когда на некоторых перемещающихся в пространстве поверхностях возникают разрывы непрерывности в распределениях функций ρ, u, p, ε , которые называются *ударными волнами*. На поверхности разрыва (фронте ударной волны) должны выполняться условия непрерывности потока вещества, энергии и количества движения (условия Гюгонио), которые имеют следующий вид:

$$\begin{aligned}
[\rho]U - [\rho u] &= 0, \\
[\rho u]U - [p + \rho u^2] &= 0, \\
[\rho(\varepsilon + u^2/2)]U - [\rho u(\varepsilon + u^2/2) + pu] &= 0.
\end{aligned} \tag{1.67}$$

Здесь U означает скорость распространения ударной волны, а квадратными скобками обозначены разности значений стоящих внутри них величин по разные стороны разрыва.

Отметим, что если $\rho(x, t)$, $u(x, t)$, (x, t) – гладкие функции, то для любого контура Γ выполняется равенство

$$\oint_{\Gamma} \rho S dx - \rho u S dt = 0, \tag{1.68}$$

означающее закон сохранения энтропии (на решениях уравнений газовой динамики). Однако на разрывах этот закон не выполняется: при прохождении элемента среды через разрыв энтропия в нем возрастает. Если контур Γ является произвольным и может захватывать участки разрыва, то вместо (1.68) следует писать *закон возрастания* (или *неубывания*) энтропии

$$\oint_{\Gamma} \rho S dx - \rho u S dt \leq 0. \tag{1.69}$$

Уравнения газодинамики описывают и такие движения, в которых на некоторых поверхностях испытывают скачок первые производные функций ρ, p, u и других, в то время как сами эти величины остаются непрерывными. Такие ситуации называются *слабыми разрывами*, а примером здесь может быть *волна разрежения*, возникающая при истечении газа в вакуум.

Простейшая модель задач гидродинамики описывается линейным одномерным *уравнением переноса* с постоянной скоростью c :

$$\frac{\partial u}{\partial t} + c \frac{\partial u}{\partial x} = 0. \tag{1.70}$$

Эта задача обобщается, с одной стороны, на одномерные квазилинейные скалярные уравнения вида

$$\frac{\partial u}{\partial t} + \frac{\partial F(u)}{\partial x} = Q(x, t), \tag{1.71}$$

а с другой – на системы линейных гиперболических уравнений

$$\frac{\partial \vec{u}}{\partial t} + A \frac{\partial \vec{u}}{\partial x} = 0, \quad \vec{u} = (u_1, \dots, u_n),$$

где матрица $A \in \mathcal{R}^{n,n}$ предполагается имеющей различные вещественные собственные числа $\lambda_1, \dots, \lambda_n$ (точнее говоря, собственные значения могут быть и кратными, но сама матрица должна быть диагонализируемой).

Перейдем теперь к описанию более общих многомерных моделей.

Уравнение Навье–Стокса описывает установившееся движение вязкой несжимаемой жидкости и представляется в следующей векторной форме:

$$\begin{aligned} -\nu \Delta \vec{u} + (\vec{u} \cdot \nabla) \vec{u} + \nabla p &= \vec{f}(\vec{r}), \\ -\operatorname{div} \vec{u} &= 0, \quad \vec{r} \in \Omega, \quad \vec{u}|_{\Gamma} = 0. \end{aligned} \quad (1.72)$$

Здесь вектор $\vec{u}$ представляет соленоидальное (бездивергентное) поле скоростей с однородным условием Дирихле на границе Γ (условие прилипания). Число $\nu > 0$ – это коэффициент вязкости, от величины которого в значительной степени зависит характер течения жидкости (ламинарное или турбулентное). Скалярная функция $p(\vec{r})$ имеет смысл кинематического давления жидкости, и для нее никакого краевого условия не ставится, а для ее однозначного определения формулируется условие $\int_{\Omega} p d\Omega = 0$.

Функции $\vec{u}$ и p ищутся соответственно в векторном и скалярном пространствах

$$\vec{H}_0^1(\Omega) = (H_0^1(\Omega))^d, \quad L_2^0 = \{p \in L_2(\Omega) : \int_{\Omega} p d\Omega = 0\}.$$

Исследования решения нелинейной системы уравнений (1.72) ведутся уже в течение многих десятилетий, и в общем трехмерном случае ($d = 3$) существование единственного решения для нее не доказано. Более сложные постановки гидро-газодинамических задач связаны с введением нестационарности, сжимаемости и различных моделей турбулентности, которые связаны с тензорным определением вязкости, однако на этих вопросах мы не останавливаемся и отсылаем заинтересованного читателя к специальным монографиям [20, 29, 151, 152].

Задача Стокса. Мы рассмотрим линейную задачу, которая может быть рассмотрена как вспомогательная при решении нелинейной системы (1.72) (при этом мы формально освобождаемся

от требования бездивергентности скорости $\vec{u}$):

$$\begin{aligned} -\Delta \vec{u} + \nabla p &= f(\vec{r}), \\ -\operatorname{div} \vec{u} &= g(\vec{r}), \quad \vec{r} \in \Omega, \quad \vec{u}|_{\Gamma} = 0. \end{aligned} \quad (1.73)$$

Данная задача может также интерпретироваться как линеаризованное приближение к системе уравнений Навье–Стокса, соответствующее большим значениям ν , или “малым” скоростям $\vec{u}$. При этом фактически в (1.73) проведена перенормировка первого уравнения и введена новая скалярная функция p . Из интегрирования второго уравнения (1.73) по области Ω следует, что для существования решения задачи Стокса необходимо выполнение равенства $\int_{\Omega} g d\Omega = 0$.

1.3.6. Динамические задачи упругости

В данном пункте мы коротко остановимся на формулировках линейных задач теории упругости и вязкоупругости, следуя в основном работе А.Н.Коновалова и Ю.П.Попова “Оптимальные явно разрешимые дискретные модели с контролируемым бисбалансом полной механической энергии для динамических задач линейной теории упругости” (Сибирский математический журнал, т. 56, N 5, 2015, 1092-1099). Исследования задач упругости представляют собой классические разделы механики сплошных сред, всесторонне изученных в монографиях В.И.Арнольда [5], О.М.Белоцерковского [20], С.К.Годунова с коллегами [47], [48], А.Н.Коновалова [90], а также в [59], [155] и многих более специальных книгах.

Напомним, что стационарные задачи упругости описаны нами ранее в п. 1.3.2 в качестве примера постановки для уравнений эллиптического типа. Если же трехмерный вектор смещений $\vec{u} = (u_1, u_2, u_3)^T$ и вектор массовой силы $\vec{f} = (f_1, f_2, f_3)^T$ (например, силы тяжести) зависят от времени t , то в основу математических моделей для динамических задач линейной теории упругости полагается закон сохранения импульса в поле массовых сил:

$$\rho \frac{\partial^2 \vec{u}}{\partial t^2} = -R^T \bar{\sigma} = \rho \vec{f}. \quad (1.74)$$

Здесь $\rho = \rho(\vec{x})$ есть плотность материала, $\bar{\sigma}$ – тензор напряжений (см. (1.42)), а R^T – матричный дифференциальный оператор следующего вида (символ T означает транспонирование):

$$R^T \bar{\sigma} = \begin{bmatrix} \partial_1 & 0 & 0 & \partial_2 & \partial_3 & 0 \\ 0 & \partial_2 & 0 & \partial_1 & 0 & \partial_3 \\ 0 & 0 & \partial_3 & 0 & \partial_1 & \partial_2 \end{bmatrix} \begin{bmatrix} \bar{\sigma}_1 \\ \bar{\sigma}_2 \end{bmatrix} \equiv \operatorname{div} \bar{\sigma}, \quad (1.75)$$

где ∂_k , $k = 1, 2, 3$ обозначают производные по соответствующим направлениям, а $\bar{\sigma}_1$ и $\bar{\sigma}_2$ суть векторы, из которых составляется тензор $\bar{\sigma}$, т.е.

$$\bar{\sigma} = (\bar{\sigma}_1 \bar{\sigma}_2)^T, \quad \bar{\sigma}_1 = (\sigma_{1,1} \sigma_{2,2} \sigma_{3,3})^T, \quad \bar{\sigma}_2 = (\sigma_{1,2} \sigma_{1,3} \sigma_{2,3})^T.$$

К соотношению (1.74) необходимо добавить уравнение состояния (закон Гука), связывающее тензор напряжений с тензором деформаций

$$\bar{\varepsilon} = (\bar{\varepsilon}_1 \bar{\varepsilon}_2)^T, \quad \bar{\varepsilon}_1 = (\varepsilon_{1,1} \varepsilon_{2,2} \varepsilon_{3,3})^T, \quad \bar{\varepsilon}_2 = (2\varepsilon_{1,2} 2\varepsilon_{1,3} 2\varepsilon_{2,3})^T$$

следующим образом:

$$\bar{\sigma} = K \bar{\varepsilon} = \{\sigma_{i,j} = 2\mu \varepsilon_{i,j} + \lambda \delta_{i,j} (\varepsilon_{1,1} + \varepsilon_{2,2} + \varepsilon_{3,3})\}. \quad (1.76)$$

Здесь участвуют коэффициенты Ламе $\lambda > 0, \mu > 0$ и символ Кронекера, уже введенные в п. 1.3.2, а K есть симметричная положительно определенная матрица с двумя диагональными блоками:

$$K_1 = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda \\ \lambda & \lambda + 2\mu & \lambda \\ \lambda & \lambda & \lambda + 2\mu \end{bmatrix}, \quad K_2 = \begin{bmatrix} \mu & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \mu \end{bmatrix}, \quad K = \begin{bmatrix} K_1 & 0 \\ 0 & K_2 \end{bmatrix}.$$

Отметим, что с помощью матриц K_1 и K_2 уравнение состояния (1.76) можно переписать в форме системы

$$\bar{\sigma}_l = K_l \bar{\varepsilon}_l, \quad l = 1, 2.$$

Кроме того, для замыкания модели линейной динамической теории упругости необходимо добавить соотношение, определяющее связь между тензором деформации и вектором скорости, которое в матричной форме записывается как

$$\bar{\bar{\varepsilon}} = R\bar{u} = \begin{bmatrix} \partial_1 & 0 & 0 \\ 0 & \partial_2 & 0 \\ 0 & 0 & \partial_3 \\ \partial_2 & \partial_1 & 0 \\ \partial_3 & 0\partial_1 & 0 \\ 0 & \partial_3 & \partial_2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \{\varepsilon_{i,j} = \frac{1}{2}(\partial_j u_i + \partial_i u_j)\}. \quad (1.77)$$

Обозначая через H^1 и H^2 гильбертовы пространства векторов и тензоров соответственно, оператор связи между $\bar{u} \in H^1$ и $\bar{\bar{\sigma}} \in H^2$ определяется как $R : H^1 \rightarrow H^2$. Ему можно сопоставить сопряженный оператор $H^* : H^2 \rightarrow H^1$, в соответствии с определением скалярных произведений (понимаемых в смысле интегралов, или свертки) в соответствующих пространствах:

$$(\bar{u}, R^* \bar{\bar{\sigma}})_{H^1} = (R\bar{u}, \bar{\bar{\sigma}})_{H^2}, \quad \bar{u} \in H^1, \quad \bar{\bar{\sigma}} \in H^2.$$

Введенный в (1.74), (1.75) транспонированный оператор связан с сопряженным соотношением $R^* = -R^T$.

Если ввести в рассмотрение вектор скорости перемещений $\bar{v} = \frac{\partial \bar{u}}{\partial t}$, описываемую системой уравнений (1.74)–(1.77) модель можно переписать в следующей форме:

$$\rho \frac{\partial \bar{v}}{\partial t} - R^T \bar{\bar{\sigma}} = \rho \bar{f}, \quad \bar{\bar{\sigma}} = R\bar{u}, \quad \bar{\bar{\sigma}} = K\bar{\bar{\varepsilon}}, \quad \frac{\partial \bar{u}}{\partial t} = \bar{v}, \quad (1.78)$$

$$\bar{u}(\bar{x}, 0) = \varphi_1(\bar{x}), \quad \bar{v}(\bar{x}, 0) = \varphi_2(\bar{x}),$$

где функции φ_1 и φ_2 определяют начальные данные, а задание областей операторов R и R^T требует соответствующего уточнения краевых условий на Γ_1 и Γ_2 .

Наконец, для однозначной постановки задачи требуется конкретизировать области определения уравнений, а также поставить необходимые начальные и краевые условия. Мы предполагаем, что решение ищется в ограниченной области изменения временной и пространственной переменных

$$t \in [0, t_e < \infty], \quad \bar{x} \in \bar{\Omega} = \Omega \cup \Gamma, \quad \Gamma = \Gamma_1 \cup \Gamma_2,$$

причем на различных участках границы для всех t задаются условия следующего вида:

$$\bar{u}|_{\Gamma_1} = 0, \quad \sigma_{i,j} n_j|_{\Gamma_2} = 0, \quad (1.79)$$

где n_j – компоненты орта внешней нормали к Γ_2 . Начальные условия Коши для $\bar{x} \in \bar{\Omega}$ задаются в форме

$$\bar{u}(\bar{x}, 0) = \bar{\varphi}_1, \quad \frac{\partial \bar{u}}{\partial t}(\bar{x}, 0) = \bar{v}(\bar{x}, 0) = \varphi_2(\bar{x}). \quad (1.80)$$

Наряду с рассмотренными формулировками могут рассматриваться и другие замкнутые модели линейной теории динамической упругости. Например, с использованием свойств введенных в (1.78) матричных операторов легко получаем следующую постановку “в перемещениях”:

$$\rho \frac{\partial^2 \bar{u}}{\partial t^2} - R^T K R \bar{u} = \rho \bar{f}. \quad (1.81)$$

Переход к данной формулировке предполагает разрешимость операторного уравнения

$$R \bar{u}(\bar{x}, t) = \bar{\sigma}(x, t) \quad (1.82)$$

для любого $t > 0$. Условие разрешимости (1.82) можно записать в виде

$$(\bar{\sigma}, \bar{\psi})_{H^2} = 0, \quad R^* \bar{\psi} = 0, \quad \bar{\psi} \in H^2,$$

который является эквивалентным так называемому условию совместности (сплошности среды) Сан-Венана.

Вводя величины $\hat{w} = (\bar{v}, \bar{\sigma})^T$ и $\hat{f} = (\rho \bar{f}, 0)^T$, постановку (1.78) можно свести к так называемой базовой модели с кососимметричным оператором A :

$$D \frac{\partial \hat{w}}{\partial t} + A \hat{w} = \hat{f}, \quad \frac{\partial \bar{u}}{\partial t} = \bar{v}, \quad \bar{\sigma} = K^{-1} \bar{\sigma}, \quad D = \begin{bmatrix} \rho E_1 & 0 \\ 0 & K^{-1} E_2 \end{bmatrix},$$

$$E_1 \bar{v} = v \in H^1, \quad E_2 \bar{\sigma} = \bar{\sigma} \in H^2, \quad A = \begin{bmatrix} 0 & R^* \\ -R & 0 \end{bmatrix}. \quad (1.83)$$

Справедливо следующее важное утверждение: базовая модель (1.83) обладает дополнительным законом сохранения

$$\frac{\partial J(t)}{\partial t} = (\rho \bar{f}, \bar{v})_{H^1}, \quad J = J_1(t) + J_2(t), \quad (1.84)$$

где величины

$$J_1(t) = 0.5(\rho \bar{v}, \bar{v})_{H^1}, \quad J_2(t) = 0.5(\bar{\sigma}, \bar{\sigma})_{H^2},$$

соответственно характеризуют кинетическую энергии упругой среды и потенциальную энергию упругих деформаций. Если соотношения (1.84) переписать в форме

$$J(t + \Delta t) = J(t) + \int_t^{t+\Delta t} (\rho \bar{f}, \frac{\partial \bar{u}}{\partial t})_{H^1} dt = J(t) + Q(t + \Delta t), \quad (1.85)$$

то величина Q в (1.85) – это работа массовых сил на приращения $\Delta \bar{u}$ за время Δt . Таким образом, уравнение (1.84) представляет собой закон сохранения полной механической энергии упругой среды.

Рассмотренные динамические задачи могут быть обобщены на случай вязкоупругой изотропной среды (модель Кельвина–Фойгта). В отличие от (1.76), уравнения состояния при этом записываются следующим образом:

$$\bar{\sigma}'_1 = K_1 \bar{\sigma}_1 + \eta_1 \frac{\partial \bar{\sigma}}{\partial t}, \quad \bar{\sigma}'_2 = K_2 \varepsilon_2 + \eta_2 \frac{\partial \bar{\sigma}_2}{\partial t}, \quad (1.86)$$

где $\eta_1(\bar{x}) > 0$, $\eta_2(\bar{x}) > 0$ суть скалярные функции, определяющие объемную и сдвиговую вязкости рассматриваемой среды. Если ради упрощения положить $\eta_1 = \eta_2 = \eta$, то из (1.86) получаем

$$\bar{\sigma}' = K \bar{\varepsilon} + \eta \frac{\partial \bar{\varepsilon}}{\partial t} = \bar{\sigma} + \eta R \bar{v}.$$

Таким образом, для задачи динамики изотропной среды “типа Кельвина” приходим к базовой модели

$$\rho \frac{\partial \bar{v}}{\partial t} + R^* \bar{\sigma} + R^* \eta R \bar{v} = \rho \bar{f}, \quad K^{-1} \frac{\partial \bar{\sigma}}{\partial t} - R \bar{v} = 0, \quad (1.87)$$

для которой закон сохранения имеет вид

$$\frac{\partial J}{\partial t} + (R^* \eta R \bar{v}, \bar{v})_{H^1} = (\rho \bar{f}, \bar{v})_{H^1}. \quad (1.88)$$

Здесь $\Phi = (R^* \eta R \bar{v}, \bar{v})_{H^1}$ есть неотрицательная диссипативная функция, характеризующая изменение механической энергии $J(t)$ при переходе $t \rightarrow t + \Delta t$.

Принципиальным отличием модели (1.87) от предыдущих является то, что она описывает необратимый процесс. Мерой необратимости является энтропия $s(t)$, которая для изотропной вязкоупругой модели удовлетворяет уравнению

$$\rho T \frac{ds}{dt} = \Phi, \quad (1.89)$$

где T есть абсолютная температура среды. Из соотношений (1.88), (1.89) следуют неравенства

$$s(t + \Delta t) \geq s(t), \quad J(t + \Delta t) \in J(t),$$

означающие, в частности, неубывание энтропии.

1.3.7. Постановки задач электромагнетизма

Данный класс математических проблем очень широк и в общей форме описывается системой *уравнений Максвелла*

$$\begin{aligned} \operatorname{rot} \vec{H}(\vec{r}, t) &= \partial \vec{D}(\vec{r}, t) / \partial t + \vec{J}(\vec{r}, t), \\ \operatorname{rot} \vec{E}(\vec{r}, t) &= -\partial \vec{B}(\vec{r}, t) / \partial t, \\ \operatorname{div} \vec{D}(\vec{r}, t) &= \rho(\vec{r}, t), \quad \operatorname{div} \vec{B}(\vec{r}, t) = 0. \end{aligned} \quad (1.90)$$

Здесь $\vec{E}$ и $\vec{H}$ – векторы напряженности электрического и магнитного полей, $\vec{D}$ и $\vec{B}$ – векторы электрической и магнитной индукции, $\vec{J}$ и ρ – плотность электрического тока и электрического заряда, $\vec{r}$ – радиус-вектор точки наблюдения, в которой определяется поле, t – время.

Многочисленные научно-технические приложения рассматриваемых ниже постановок включают актуальные задачи электродинамики и электрохимии, микроэлектроники и СВЧ-электроники, физики плазмы и геомагнетизма и т.д. Среди огромного количества литературы по изучению данных проблем мы отметим книги Ю. А. Березина и М. А. Вшивкова [23, 53], А. Д. Григорьева [52], П. Монка [133], В. А. Сырового [172], Р. Хокни [196], а также автора [79], [84].

Из системы уравнений (1.90) можно вывести скалярное *уравнение неразрывности*

$$\frac{\partial \rho}{\partial t} + \operatorname{div} \vec{J} = 0, \quad (1.91)$$

выражающее фундаментальный закон сохранения в дифференциальной форме. Плотность электрического тока целесообразно представить в виде суммы двух слагаемых:

$$\vec{J} = \vec{J}_c + \vec{J}_i, \quad (1.92)$$

где $\vec{J}_c$ – плотность тока, наведенного электрическим полем в среде с конечной проводимостью, а $\vec{J}_i$ – плотность стороннего тока, внешнего по отношению к рассматриваемой задаче. Этот ток, наряду с зарядом, и возбуждает электромагнитное поле в среде.

Для того, чтобы система уравнений Максвелла была определенной, к ней необходимо добавить материальные уравнения — соотношения, связывающие напряженности полей, индукции и плотность тока в среде (далее независимые переменные $\vec{r}$ и t для краткости опускаем):

$$\begin{aligned}\vec{D} &= \varepsilon \vec{E} + \xi \vec{H}, \\ \vec{B} &= \xi \vec{E} + \mu \vec{H}, \\ \vec{J} &= \sigma \vec{E}.\end{aligned}\tag{1.93}$$

Здесь последнее уравнение выражает закон Ома для среды с электропроводностью σ , величины ε и μ обозначают диэлектрическую и магнитную проницаемости, а ξ и ζ — кросс-проницаемости соответственно. Все эти коэффициенты в (1.93) в общем случае являются тензорами второго ранга. Если все параметры среды скалярны, ее называют изотропной, а в противном случае — анизотропной. Величины проницаемостей обычно представляют в виде произведений

$$\varepsilon = \varepsilon_0 \varepsilon_r, \quad \mu = \mu_0 \mu_r,$$

где ε_0 и μ_0 — электрическая и магнитная постоянные вакуума, а ε_r и μ_r — соответствующие относительные проницаемости среды.

Система уравнений (1.90) содержит в качестве независимой переменной время. Как говорят, она определена во временной области и справедлива для полей, зависящих от t по произвольному закону. Однако в природе и технике часто встречаются поля, изменяющиеся во времени по гармоническому закону:

$$a(\vec{r}, t) = A_m(\vec{r}) \cos(\omega t + \varphi_0),\tag{1.94}$$

где a — одна из проекций векторов поля на оси координат, A_m — ее амплитуда, $\omega = 2\pi f$ — круговая частота, f — частота изменения поля, φ_0 — начальная фаза, а аргумент косинуса $\varphi = \omega t + \varphi_0$ называется фазой величины a . Выражение (1.94) удобно записывать в следующей комплексной форме:

$$a(\vec{r}, t) = \operatorname{Re}(\dot{a} e^{i\omega t}), \quad \dot{a} = a_m e^{i\varphi_0},\tag{1.95}$$

где $\dot{a}$ — есть комплексная амплитуда величины a .

Записав все составляющие электромагнитного поля в виде (1.95) и подставив их в уравнения системы (1.90), после несложных преобразований получим уравнения Максвелла в следующей комплексной форме:

$$\begin{aligned}
rot \vec{\dot{H}} &= i\omega \dot{\epsilon} \vec{E} + \dot{\vec{J}}, \\
rot \vec{\dot{E}} &= -i\omega \dot{\mu} \vec{H}, \\
div(\dot{\epsilon} \vec{E}) &= \dot{\rho}, \quad div(\dot{\mu} \vec{H}) = 0.
\end{aligned}
\tag{1.96}$$

Здесь $\dot{\epsilon} = \epsilon \dot{\epsilon}_r = \epsilon_0(\epsilon'_r - i\epsilon''_r)$ и $\dot{\mu} = \mu_0 \mu_r = \mu_0(\mu'_r - i\mu''_r)$ – комплексные диэлектрическая и магнитная проницаемости среды, а точками над буквами отмечены комплексные амплитуды величин в соответствии с их определением в (1.59).

Полученная система уравнений (1.96) содержит в качестве независимой переменной, или параметра, угловую частоту. Поэтому говорят, что она определена в частотной области. Отметим, что в линейной среде, в которой материальные свойства не зависят от поля, временное и частотное представление функции связаны *прямым и обратным преобразованием Фурье*:

$$\begin{aligned}
g(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(t) e^{-i\omega t} dt, \\
g(t) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(\omega) e^{i\omega t} d\omega.
\end{aligned}
\tag{1.97}$$

Целесообразность выбора временной или частотной области определяется конкретными особенностями решаемой задачи. Во многих случаях удастся существенно упростить рассматриваемую постановку путем введения *электродинамических потенциалов*. Этот прием особенно эффективен при расчете поля в однородной среде, а заключается он в представлении напряженностей электрического и магнитного полей через векторный потенциал $\vec{A}$ и скалярный потенциал φ :

$$\begin{aligned}
\vec{E} &= -\frac{\partial \vec{A}}{\partial t} - \nabla \varphi, \\
\vec{H} &= \mu^{-1} rot \vec{A},
\end{aligned}
\tag{1.98}$$

которые связываются между собой так называемой *калибровкой Лоренца*

$$div \vec{A} + \epsilon \mu \frac{\partial \varphi}{\partial t} = 0. \tag{1.99}$$

При этом электродинамические потенциалы в однородной изотропной среде удовлетворяют волновым уравнениям

$$\begin{aligned}\nabla^2 \vec{A} - \varepsilon \mu \frac{\partial^2 \vec{A}}{\partial t^2} &= -\mu \vec{J}, \\ \Delta \varphi - \varepsilon \mu \frac{\partial^2 \varphi}{\partial t^2} &= -\varepsilon^{-1} \rho.\end{aligned}\tag{1.100}$$

В частотной области уравнения для потенциалов приобретают следующий вид:

$$\begin{aligned}\vec{E} &= -i\omega \vec{A} - \nabla \varphi, \\ \vec{H} &= \mu^{-1} \text{rot} \vec{A}, \\ \nabla^2 \vec{A} + k^2 \vec{A} &= -\mu \vec{J}, \\ \Delta \varphi + k^2 \varphi &= -\varepsilon^{-1} \rho,\end{aligned}\tag{1.101}$$

где $k = \omega \sqrt{\varepsilon \mu} = k_0 \sqrt{\varepsilon_r \mu_r}$ – волновое число для данной среды, а $k_0 = \omega/c$ есть волновое число в свободном пространстве, или вакууме (c – это скорость света). Два последних соотношения в (1.101) представляют собой векторное и скалярное уравнения Гельмгольца, в которых общее число неизвестных функций сокращено до четырех (три проекции вектора $\vec{A}$ и скаляр φ).

Для однозначного определения решения уравнений Максвелла во временной области необходимо в начальный момент времени $t = t_0$ задать значения напряженности электрического и магнитного полей

$$\vec{E}^0 = \vec{E}(\vec{r}, t_0), \quad \vec{H}^0 = \vec{H}(\vec{r}, t_0), \quad \vec{r} \in \bar{\Omega} = \Omega \cup \Gamma,$$

во всей расчетной области Ω с границей Γ , на различных частях которой должны быть заданы краевые условия: касательная составляющая или электрического поля $\vec{E}_\tau$, или магнитного поля $\vec{H}_\tau$.

В частности, на поверхности раздела диэлектрика и идеального проводника (электрическая стенка Γ_E) касательная составляющая вектора $\vec{E}$ обращается в нуль, т.е.

$$\vec{n} \times \vec{E} = 0, \quad \vec{r} \in \Gamma_E,\tag{1.102}$$

где $\vec{n}$ – нормаль к поверхности раздела сред. Аналогично, на магнитной стенке Γ_H (поверхность раздела диэлектрика и идеального магнетика) равна нулю касательная направляющая напряженности магнитного поля:

$$\vec{n} \times \vec{H} = 0, \quad \vec{r} \in \Gamma_H. \quad (1.103)$$

Отметим, что в соотношениях (1.102) и (1.103) поля $\vec{E}$ и $\vec{H}$ относятся к диэлектрику, а нормальные составляющие векторов $\vec{H}$ и $\vec{E}$ на электрической и магнитной стенках соответственно равны нулю (в предположении изотропности каждой из сред).

На поверхности раздела двух диэлектриков Γ_i с разными материальными свойствами дифференциальные уравнения не определены и для полей должны выполняться условия сопряжения

$$\begin{aligned} \vec{n} \times (\vec{E} - \vec{E}^-) &= \vec{J}_{\Gamma_i}^m, \\ \vec{n} \times (\vec{H}^+ - \vec{H}^-) &= \vec{J}_{\Gamma_i}^e, \\ \vec{n} \cdot (\vec{D}^+ - \vec{D}^-) &= \rho_{\Gamma_i}^e, \\ \vec{n} \cdot (\vec{B}^+ - \vec{B}^-) &= \rho_{\Gamma_i}^m, \end{aligned} \quad (1.104)$$

где знаки “+,” “−” относятся к разным односторонним значениям полей на внутренней границе Γ_i . В соотношениях (1.104) величины $\vec{J}_{\Gamma_i}^e$ и $\vec{J}_{\Gamma_i}^m$ суть поверхностные плотности электрического и магнитного токов, а $\rho_{\Gamma_i}^e$ и $\rho_{\Gamma_i}^m$ – поверхностные плотности электрического и магнитного зарядов. В частности, если поверхностные токи отсутствуют, то касательные составляющие электрического и магнитного полей на Γ_i непрерывны, а если отсутствуют поверхностные заряды, то на Γ_i непрерывны нормальные составляющие электрической и магнитной индукции.

При рассмотрении краевых задач электродинамики необходимо разделять внутренние и внешние задачи. В последнем случае, когда расчетная область Ω неограничена, на бесконечности ставятся условия излучения Зоммерфельда, устанавливающие, что при $r \rightarrow \infty$ напряженности $\vec{E}$ и $\vec{H}$ должны убывать быстрее, чем $1/r$.

Для упрощения электродинамических краевых задач из системы уравнений Максвелла можно исключить, например, вектор $\vec{H}$, в результате чего получаем дифференциальное уравнение относительно $\vec{E}$:

$$\text{rot}(\mu_r^{-1} \text{rot} \vec{E}) + \varepsilon_r \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} + \sigma \mu_0 \frac{\partial \vec{E}}{\partial t} = -\mu_0 \frac{\partial \vec{J}}{\partial t}. \quad (1.105)$$

Если поля изменяются по гармоническому закону, то для комплексной амплитуды $\vec{E}$ в частотной области мы получаем из (1.105) следующее векторное уравнение Гельмгольца:

$$\operatorname{rot}(\mu_r^{-1} \operatorname{rot} \vec{E}) + \varepsilon_r \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} + \sigma \mu_0 \frac{\partial \vec{E}}{\partial t} = -\mu_0 \frac{\partial \vec{J}}{\partial t}. \quad (1.106)$$

В частном случае стационарных полей следует достаточно простая, но актуальная для многих приложений, *задача электростатики*: из (1.98) мы имеем $\vec{E} = -\nabla \varphi$, а из предпоследнего уравнения в (1.90), с учетом следующего из (1.93) соотношения $\vec{D} = \varepsilon \vec{E}$ получаем уравнение диффузионного типа для скалярного потенциала:

$$\operatorname{div} \varepsilon \operatorname{grad} \varphi = -\rho. \quad (1.107)$$

Отсюда для однородной среды выводится уравнение Пуассона, формально являющееся частным случаем скалярного уравнения Гельмгольца из (1.65) с нулевой частотой, т.е. $k = \omega = 0$.

Уравнения, аналогичные (1.105), (1.106), могут быть выведены и для вектора напряженности магнитного поля $\vec{H}$. Достаточно содержательным примером здесь может служить нелинейная в общем случае *задача магнитостатики*:

$$\operatorname{rot} \nu \operatorname{rot} \vec{A} = \vec{J}, \quad (1.108)$$

где величина $\nu = (\mu_0 \mu_r)^{-1}$ называется магнитным сопротивлением. Особый интерес такие задачи представляют при наличии в расчетной области магнитных материалов, для которых величина ν зависит от искомого решения и в случае анизотропных сред представляется тензором.

1.3.8. Интегральные уравнения теории потенциала

Интегральные представления, наряду с дифференциальными уравнениями, являются мощным аппаратом решения задач математического моделирования. Мы в данном пункте ограничимся только сравнительно узким классом задач теории потенциала, связанных с употреблением оператора Лапласа, примером чего являются задачи электростатики из предыдущего пункта. Справедливости ради следует сказать, что рассматриваемые ниже интегральные уравнения актуальны для многих приложений, в том числе далеко выходящих за рамки электромагнетизма. Данные вопросы достаточно хорошо изучены как в работах XX в. (например, классический учебник Л.Н. Тихонова и А.А. Самарского [177]), так и в более современных публикациях, среди которых мы назовем книги О.М. Белоцерковского и

О. М. Лифанова [21] и П. Бенерджи [22], имеющие достаточно представительный обзор литературы.

Пусть Ω – ограниченная трехмерная область с достаточно гладкой поверхностью Γ . Для произвольной функции $u(x, y, z)$, непрерывной в $\Omega + \Gamma$ и имеющей непрерывные производные внутри Ω , справедлива интегральная формула Грина

$$\omega u(P) = \int_{\Gamma} \left[\frac{1}{R(P, Q)} \frac{\partial u}{\partial n_Q}(Q) - u(Q) \frac{\partial}{\partial n_Q} \left(\frac{1}{R(P, Q)} \right) \right] dS_Q - \int_{\Omega} \frac{\Delta u(Q)}{R(P, Q)} dV_Q. \quad (1.109)$$

Здесь Δ – оператор Лапласа, n_Q – внешняя нормаль к границе Γ в точке Q , $R(P, Q)$ – расстояние между точками P и Q , а ω принимает следующие значения:

$$\omega = \begin{cases} 4\pi, & \text{если точка } P \text{ лежит внутри } \Omega, \\ 2\pi, & \text{если точка } P \text{ лежит на границе } \Gamma, \\ 0, & \text{если точка } P \text{ лежит вне } \Omega. \end{cases}$$

Для гармонической функции, удовлетворяющей уравнению Лапласа $\Delta u = 0$, в точках P , лежащих внутри ω , получаем интегральное представление

$$u(P) = \frac{1}{4\pi} \int_{\Gamma} \left[\frac{\partial u}{\partial n_Q}(Q) - u(Q) \frac{\partial}{\partial n_Q} \left(\frac{1}{R(P, Q)} \right) \right] dS_Q. \quad (1.110)$$

Функция $G(P, Q) = R^{-1}(P, Q) = |\vec{r}(P) - \vec{r}(Q)|^{-1}$ является гармонической и представляет собой поле единичного заряда, помещенного в точке Q свободного (однородного) пространства.

Если $\rho(Q)$ есть распределение объемной плотности электрических зарядов (в свободном пространстве), то обусловленный ими потенциал определяется интегралом

$$u(P) = \int_{\Omega} \frac{\rho(Q)}{R(P, Q)} dV_Q. \quad (1.111)$$

Как показывает основная формула Грина (1.110), любая гармоническая функция может быть выражена с помощью интегралов, представляющих *поверхностные потенциалы*.

Потенциал поверхностной плотности зарядов $\mu(Q)$, распределенных на поверхности Γ , описывается интегралом

$$v(P) = \int_{\Gamma} \frac{\mu(Q)}{R(P, Q)} dS_Q \quad (1.112)$$

и называется *потенциалом простого слоя*.

Если функция $\nu(Q)$ описывает поверхностную плотность диполей, то обусловленный ею потенциал выражается поверхностным интегралом

$$\omega(P) = - \int_{\Gamma} \nu(Q) \frac{\partial}{\partial n_Q} \left(\frac{1}{R(P, Q)} \right) dS \quad (1.113)$$

и называется *потенциалом двойного слоя*.

Потенциал двойного слоя в некоторой точке Q_0 на поверхности Γ является разрывной функцией:

$$\omega_+(Q_0) = \omega(Q_0) + 2\pi \nu(Q_0), \quad \omega_-(Q_0) = \omega(Q_0) - 2\pi \nu(Q_0). \quad (1.114)$$

Здесь $\omega_+(Q_0)$ и $\omega_-(Q_0)$ – предельные значения потенциала двойного слоя при подходе к точке Q_0 с разных сторон поверхности.

Потенциал простого слоя непрерывен в точках поверхности Γ , но терпит разрыв его нормальная производная:

$$\left(\frac{\partial v}{\partial n} \right)_+ = \frac{\partial v}{\partial n}(Q_0) + 2\pi \mu(Q_0), \quad \left(\frac{\partial v}{\partial n} \right)_- = \frac{\partial v}{\partial n}(Q_0) - 2\pi \mu(Q_0), \quad (1.115)$$

где знаками “+”, “−” обозначены предельные значения производных, а $\frac{\partial v}{\partial n}(Q_0)$ есть так называемое среднее значение нормальной производной.

Мы рассмотрим построение интегральных уравнений, являющихся в определенном смысле эквивалентными смешанной краевой задаче для уравнения Лапласа

$$\Delta u = 0 \quad (1.116)$$

в трехмерной области Ω с границей $\Gamma = \Gamma_1 \cup \Gamma_2$, на различных участках которой заданы условия 1-го, 2-го или 3-го рода:

$$u|_{\Gamma_1} = g, \quad \left(\frac{\partial u}{\partial n} + \alpha u \right)|_{\Gamma_2} = g_2. \quad (1.117)$$

$$u|_{\Gamma_3^+} = u|_{\Gamma_3^-}, \quad \varepsilon + \frac{\partial u}{\partial n}|_{\Gamma_3^+} = \varepsilon - \frac{\partial u}{\partial n}|_{\Gamma_3^-}. \quad (1.118)$$

Если для трехмерной смешанной краевой задачи (1.116)–(1.118) решение внутри Ω искать в виде потенциала простого слоя, то с учетом соотношений (1.79) получаем систему интегральных уравнений для поверхностной плотности зарядов $\mu(Q)$:

$$\begin{aligned} \int_{\Gamma} \mu(Q) G_1(P, Q) dS_Q &= g_1(P), \quad P \in \Gamma_1, \\ 2\pi\mu(P) + \int_{\Gamma} \mu(Q) \left(\frac{\partial G_1}{\partial n_Q} + \alpha G_1(P, Q) \right) dS_Q &= g_2(P), \quad P \in \Gamma_2, \\ 2\pi\mu(P) + \frac{\varepsilon_+ - \varepsilon_-}{\varepsilon_+ + \varepsilon_-} \int_{\Gamma} \mu(Q) \frac{\partial G_1}{\partial n_Q} dS_Q &= 0, \quad P \in \Gamma_3. \end{aligned} \quad (1.119)$$

Здесь ядро записывается в виде $G_1(P, Q) = 1/R(P, Q)$, а производные от ядра могут быть вычислены по формулам

$$\frac{\partial G_1(P, Q)}{\partial n_Q} = \frac{\partial G_1}{\partial x} \cos(n_Q, x) + \frac{\partial G_1}{\partial y} \cos(n_Q, y) + \frac{\partial G_1}{\partial z} \cos(n_Q, z).$$

Для внешней краевой задачи, решаемой в неограниченной области $\Omega' = R^3 \setminus \bar{\Omega}$, являющейся дополнением к $\Omega \cup \Gamma$, во втором и третьем уравнениях системы (1.119) перед внеинтегральными членами левых частей надо поменять знаки.

Решение краевых задач можно также искать с помощью интегрального уравнения для плотности потенциала двойного слоя, что мы проиллюстрируем ради простоты для граничных условий только 1-го рода, в предположении односвязности границы $\Gamma = \Gamma_1$. При этом из соотношений (1.109), (1.110) получаем уравнение

$$\begin{aligned} \pm 2\pi\nu(P) + \int_{\Gamma} G_2(P, Q) \nu(Q) \alpha dS_Q &= g_1(P), \quad P \in \Gamma, \\ G_2(P, Q) &= \frac{d}{dn_Q} (R^{-1}(P, Q)), \end{aligned}$$

где знаки “+” или “−” используются для внутренней и внешней задач соответственно.

1.3.9. Гамильтоновы системы уравнений

Предложенные В. Р. Гамильтоном (1805–1865 гг.) фундаментальные принципы описания механических систем произвели такое

впечатление на его последователя К. Г. Якоби, что он предложил рассматривать всю предыдущую историю исследований Галилея, Ньютона, Эйлера и Лагранжа только как введение в предмет.

Если проблема исследований общих механических систем имеет началом работу Галилея, опубликованную в 1638 г., то “Принципы” Ньютона (1687 г.) позволили свести ее к решению дифференциальных уравнений. Это дало возможность получить теоретические результаты для ряда сложных задач по движению свободных точечных масс (в том числе планет) или жестких тел, однако только благодаря публикациям Ж. Л. Лагранжа (1760, 1788 гг.) удалось построить элегантный и достаточно универсальный аппарат исследований. Дальнейшими вехами в развитии идей Лагранжа и Гамильтона послужили три монументальных тома А. Пуанкаре по лунной механике (1892, 1893, 1899 гг.), а в XX-ом веке — книги Сигеля и Мозера (1956, 1971 гг.), а также основополагающая монография В. И. Арнольда [5]. В последние десятилетия активное развитие получили исследования по распространению гамильтонова формализма на уравнения в частных производных, в том числе нелинейных. Здесь следует назвать книги В. Е. Захарова с соавторами [76], С. П. Новикова и И. А. Тайманова [135], С. Б. Куксина [98], Б. Лаймкюхлера, С. Райха [103], Л. Бругнано [28] К. Фенга [184], Д. Хаара [191] и Э. Хайрера [192]– [194], однако основной объем результатов опубликован пока в журнальной литературе. Следует сказать, что современное изложение данного материала опирается на аппарат внешних алгебр, дифференциальных форм и теории групп, однако мы ограничиваемся кратким описанием принципов гамильтонова подхода рамками “традиционного” функционального анализа.

Принципы лагранжевой и гамильтоновой механики. Основой лагранжеевого подхода является описание механической системы с d степенями свободы с помощью векторов *обобщенных координат* $q = (q_1, \dots, q_d)^T$, примером которых могут быть декартовы координаты, углы, длина дуги вдоль кривой и т.д. Ключевой момент здесь заключается во введении двух функционалов:

$$T = T(q, \dot{q}), \quad U = U(q), \quad (1.120)$$

первый из которых представляет *кинетическую энергию*, обыч-

но представляемую скалярным произведением $\frac{1}{2}\dot{q}^T M(q)\dot{q}$ (точка означает производную по времени) с симметричной положительно определенной матрицей $M(q)$, а второй — это зависящая только от координат *потенциальная энергия*. Для определенного нового функционала

$$L = T - U, \quad (1.121)$$

называемого *лагранжианом*, устанавливаются соотношения

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_k} \right) - \frac{\partial L}{\partial q_k} = 0, \quad k = 1, \dots, d, \quad (1.122)$$

которые известны как *уравнения Лагранжа* и при заданных начальных условиях полностью определяет поведение механической системы.

Обобщенные координаты взаимнооднозначно связаны с декартовыми координатами $\vec{r}_k$ (отметим, что в трехмерном случае число степеней свободы равно $d = 3N$), где N — число частиц механической системы, которую мы для простоты рассматриваем без дополнительных голономных связей). Что касается введенных в (1.122) *обобщенных скоростей* $\dot{q}_k$, то они связаны с “обычными” скоростями $\dot{\vec{r}}$ соотношениями

$$\dot{\vec{r}}_k = \sum_{i=1}^d \frac{\partial \vec{r}_k}{\partial q_i} \dot{q}_i, \quad k = 1, \dots, d. \quad (1.123)$$

Принципиальные нововведения Гамильтона заключаются в определении обобщенных импульсов

$$p_k = \frac{\partial L}{\partial \dot{q}_k}(q, \dot{q}), \quad k = 1, \dots, d, \quad (1.124)$$

как независимых переменных и в рассмотрении функционала

$$H = H(p, q) = p^T \dot{q} - L(q, \dot{q}), \quad (1.125)$$

называемым *гамильтонианом*, причем в (1.125) полагается, что величина $\dot{q}$ с помощью (1.124) выражается через p и q . Рассмотренные величины удовлетворяют системе обыкновенных дифференциальных уравнений первого порядка

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, \quad \dot{p}_k = -\frac{\partial H}{\partial q_k}, \quad k = 1, \dots, d, \quad (1.126)$$

называемыми *уравнениями Гамильтона*, или каноническими уравнениями движения.

Преобразование переменных (1.124) должно являться непрерывным дифференцируемым взаимнооднозначным отображением $\dot{q} \longleftrightarrow p$ и называется *преобразованием Лежандра*.

Если кинетическая энергия представляется квадратичной формой $T = \frac{1}{2} \dot{q}^T M(q) \dot{q}$ с симметричной положительно определенной матрицей $M(q)$, для фиксированного q мы имеем $p = M(q) \dot{q}$, так что преобразование Лежандра в данном случае существует. Далее, в силу соотношения $\dot{q} = M^{-1} p$, из определения гамильтониана в (1.125) мы получаем

$$H(p, q) = p^T M^{-1} p - L(q, M^{-1} p) = p^T M^{-1} p - \frac{1}{2} p^T M^{-1} p + U(q) = T(p, q) + U(q), \quad (1.127)$$

т.е. гамильтониан представляет полную энергию механической системы.

Отметим важное классификационное отличие между уравнениями Лагранжа и Гамильтона. Поскольку в лагранжиане $L(q, \dot{q})$ переменные не являются независимыми ($\dot{q}$ – производная по времени от q), то система (1.122) – это d обыкновенных дифференциальных уравнений второго порядка. Система же (1.126) включает $2d$ уравнений первого порядка, связывающих между собой такое же количество независимых переменных p_k, q_k . Очевидное структурное достоинство уравнений Гамильтона – их простая симметричная форма.

Если H не зависит от времени явно, то из (1.126) непосредственно получаем, что гамильтониан является первым интегралом уравнений движения, поскольку

$$\frac{dH}{dt} = \frac{\partial H}{\partial t} + \sum_k \frac{\partial H}{\partial p_k} \dot{p}_k + \sum_k \frac{\partial H}{\partial q_k} \dot{q}_k = \frac{\partial H}{\partial t} = 0. \quad (1.128)$$

Канонические преобразования. Очевидно, что при описании или решении уравнений движения можно осуществить их преобразование от совокупности переменных p_k и q_k к другой совокупности переменных α_k и β_k ($k = 1, \dots, d$).

Преобразование переменных

$$p_k = p_k(\alpha, \beta), \quad q_k = q_k(\alpha, \beta), \quad k = 1, \dots, d, \quad (1.129)$$

называется *каноническим*, если уравнения Гамильтона для функционала полной энергии механической системы $\bar{H}(\alpha, \beta)$ сохраняют каноническую форму (1.126), т.е.

$$\dot{\alpha}_k = -\frac{\partial \bar{H}}{\partial \beta_k}, \quad \dot{\beta}_k = \frac{\partial \bar{H}}{\partial \alpha_k}, \quad k = 1, \dots, d. \quad (1.130)$$

Необходимым и достаточным условием того, чтобы преобразование от p_k и q_k к переменным α_k и β_k было каноническим, является существование *производящей* (порождающей) *функции* $W(q_k, \beta_k)$ такой, что выполняются соотношения

$$p_k = \frac{\partial W}{\partial q_k}, \quad \alpha_k = -\frac{\partial W}{\partial \beta_k}. \quad (1.131)$$

Важно отметить, что для любого канонического преобразования производящая функция W всегда существует, фактически для двух наборов α переменных существуют четыре различные комбинации пар: q_k и β_k , q_k и α_k , p_k и β_k , p_k и α_k , – которым соответствуют свои производящие функции и уравнения движения. В частности, для производящей функции

$$W = \sum_k q_k \beta_k \quad (1.132)$$

каноническое преобразование имеет вид

$$p_k = \beta_k, \quad q_k = -\alpha_k. \quad (1.133)$$

При этом “старые” импульсы превращаются в “новые” координаты, а старые координаты переходят в новые импульсы. Отсюда видно, что можно отказаться от терминов “импульсы” и “координаты” для пар величин α_k и β_k , а лучше назвать их *сопряженными переменными*.

Скобки Пуассона. Из канонических уравнений движения, описывающих поведение величин p_k и q_k , можно найти и уравнения для любой функции $\mathcal{F}(p_k, q_k)$ этих переменных. Нетрудно убедиться, что если функция $\mathcal{F}$ явно не зависит от времени, т.е. $\partial \mathcal{F} / \partial t = 0$, а гамильтониан $H(p_k, q_k)$ удовлетворяет системе (1.126), то справедливы соотношения

$$\begin{aligned}\mathcal{F} &= \sum_k \left(\frac{\partial \mathcal{F}}{\partial q_k} \dot{q}_k + \frac{\partial \mathcal{F}}{\partial p_k} \dot{p}_k \right) = \sum_k \left(\frac{\partial \mathcal{F}}{\partial q_k} \frac{\partial H}{\partial p_k} - \frac{\partial \mathcal{F}}{\partial p_k} \frac{\partial H}{\partial q_k} \right) = \\ &= \{\mathcal{F}, H\} = -\{H, \mathcal{F}\},\end{aligned}\quad (1.134)$$

где обозначенная комбинация $\{\mathcal{F}, H\}$ носит наименование *скобок Пуассона*. Приведенные рассуждения обозначают, в частности, следующее важное утверждение: необходимым и достаточным условием того, чтобы функция $\mathcal{F}$ была интегралом движения, является равенство

$$\{\mathcal{F}, H\} = 0.$$

Для скобок Пуассона, примененных к независимым переменным p_k и q_k , которые называются *фундаментальными скобками*, выполняются следующие соотношения:

$$\{q_k, p_l\} = \delta_{k,l}, \quad \{q_k, q_l\} = 0, \quad \{p_k, p_l\} = 0. \quad (1.135)$$

Данные равенства определяют инвариантное свойство канонических переменных в следующем смысле: если переменные α_k и β_k определены из канонических преобразований (1.129), т.е. выполняются уравнения (1.130), то справедливы аналогичные (1.135) равенства для фундаментальных скобок:

$$\{\beta_j, \alpha_i\} = \delta_{i,j}, \quad \{\alpha_i, \alpha_j\} = 0, \quad \{\beta_i, \beta_j\} = 0. \quad (1.136)$$

С помощью несложно проверяемых выражение для функции $\mathcal{F}(p_k, q_k)$

$$\{p_k, \mathcal{F}\} = -\frac{\partial \mathcal{F}}{\partial q_k}, \quad \{q_k, \mathcal{F}\} = \frac{\partial \mathcal{F}}{\partial p_k} \quad (1.137)$$

устанавливается такой принципиальный факт, как инвариантность скобок Пуассона относительно канонических преобразований, означающее выполнение для пары функций равенства

$$\{\mathcal{F}, G\}' = \{\mathcal{F}, G\}, \quad (1.138)$$

где штрихом обозначена скобка Пуассона, записанная в канонических переменных α_i, β_i , $i = 1, \dots, d$.

Скобки Пуассона обладают таким замечательным свойством, как *тождество Якоби* для тройки функций f, g, h :

$$\{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0. \quad (1.139)$$

Справедливо также следующее полезное соотношение:

$$\{f, gh\} = g\{f, h\} + \{f, g\}h. \quad (1.140)$$

Если H есть гамильтониан механической системы, а $\mathcal{F}$ и G – интегралы движения, то выполняется равенство

$$\{h, \{\mathcal{F}, G\}\} = -\{\mathcal{F}, \{G, H\}\} + \{G, \{\mathcal{F}, H\}\} = 0, \quad (1.141)$$

т.е. их скобка Пуассона также является интегралом движения.

Примерами интегралов движения системы частиц с массами m_k , описываемых декартовыми координатами $\vec{r}_k$ и “обычными” импульсами $\dot{\vec{p}}_k = m_k \dot{\vec{r}}_k$, являются полный момент импульса

$$\vec{M} = (M_x, M_y, M_z) = \sum_k [\vec{r}_k, \vec{p}_k]$$

и полный импульс

$$\vec{P} = (P_x, P_y, P_z) = \sum_k \vec{p}_k.$$

В частности, для компонентов этих векторов справедливы следующие соотношения:

$$\{M_x, M_y\} = M_z, \quad \{M_y, M_z\} = M_x, \quad \{M_z, M_x\} = M_y,$$

$$\{P_x, M_x\} = 0, \quad \{P_y, M_y\} = 0, \quad \{P_z, M_z\} = 0.$$

Пример 1: гамильтонова электродинамика. В качестве актуального примера гамильтонова формализма рассмотрим движение точечной заряженной частицы с координатой $\vec{r}(t)$, зарядом e и массой m в электромагнитном поле, описываемом скалярным и векторным потенциалами $\varphi, \vec{A}$. Из соотношений (1.126) и (1.137) следует, что уравнения движения можно записать в форме

$$\dot{q}_k = \{q_k, H\}, \quad \dot{p}_k = \{p_k, H\}, \quad k = 1, \dots, d. \quad (1.142)$$

Импульс частицы выражается формулой

$$\vec{p} = m\dot{\vec{r}} + e\vec{A}, \quad (1.143)$$

а гамильтониан (сумма потенциальной и кинетической энергии) имеет вид

$$H = e\varphi + |\vec{p} - e\vec{A}|^2/2m. \quad (1.144)$$

Уравнения движения вида (1.142) в переменных $\vec{r}, \vec{p}$ принимают вид

$$\dot{\vec{r}} = \{\vec{r}, H\}, \quad \dot{\vec{p}} = \{\vec{p}, H\}, \quad (1.145)$$

откуда следуют привычные для электродинамики соотношения

$$\begin{aligned} \dot{\vec{r}} &= (\vec{p} - e\vec{A})/m, \\ \dot{\vec{p}} &= (e/m)[(\vec{p} - e\vec{A}), \text{rot}\vec{A}] - e \text{grad}\varphi. \end{aligned} \quad (1.146)$$

Пример 2: задача Кеплера. В историческом плане развитие гамильтонова формализма инициировалось постановками задач о решении систем обыкновенных дифференциальных уравнений, описывающих движение небесных объектов, и в первую очередь планет. Одной из главных целей исследований здесь является изучение условий интегрируемости задачи Коши, в смысле построения первых интегралов дифференциальных уравнений.

Классическим примером в данном случае является *задача Кеплера* о движении двух тел в трехмерном пространстве $\mathbb{R}^3$, описываемая уравнениями

$$\dot{q} = p, \quad \dot{p} = -\frac{K}{q^3}q, \quad q(t_0) = q_0, \quad p(t_0) = p_0,$$

где $(q, p) \in \mathbb{R}^3 \times \mathbb{R}$, а $K > 0$ – некоторая постоянная. Данный пример замечателен тем, что для него существует численный метод, который для достаточного малого временного шага h позволяет получить сколь угодно точное решение. Этот алгоритм состоит из трех этапов.

- Сначала вычисляются скалярные величины

$$r = |q_0|, \quad P = |p_0|^2/2, \quad M = q_0^T p_0, \quad A = K/r - P > 0, \quad \lambda = \sqrt{2A}$$

.

- Затем находится решение x нелинейного уравнения

$$\begin{aligned} f(x) &= K \arctan(x) + \frac{a_1 x + a_2 x^2 - a_1 x^3}{(1+x^2)^2} + a_3 = 0 \\ a_1 &= r(A - P), \quad a_2 = 2\lambda M, \quad a_3 = -A\lambda h/2. \end{aligned}$$

Это может быть сделано численно с помощью итерационного метода Ньютона

$$x_{n+1} = x_n - f(x_n)/f'(x_n).$$

- Решение исходных уравнений для $t_1 = t_0 + h$ выражается формулами

$$q_1 = q_0 + d_1 q_0 + d_2 p_0, \quad p_1 = p_0 + d_3 q_0 + d_4 p_0,$$

где используемые коэффициенты имеют следующий вид:

$$\begin{aligned} d_1 &= s^2(1 + P/A), \quad d_2 = (cs\lambda r + s^2 M)/A, \\ d_3 &= -cs\lambda(1 + P/A)/|q_1|, \quad d_4 = rd_1/|q_1|, \\ |q_1| &= r + rs^2(P/A - 1) + cs\lambda M/A, \\ c &= (1 - x^2)/(1 + x^2), \quad s = 2x/(1 + x^2). \end{aligned}$$

Доказано, что если шаг h меньше половины периода эллиптического движения тел (аналогичные выражения имеются для параболических и гиперболических траекторий), итерационный метод Ньютона сходится и приведенные формулы дают численное решение задачи Кеплера со сколь угодно малой ошибкой. Обзор литературы по данной проблеме приведен в диссертации Г. Вилмарта, доступной в сети Интернет (G. Vilmart. Etude d'integrateurs geometriques pour des equations differentielles/ These N4038, l'Universite de Geneve, 2008).

Симплектические отображения и уравнения. Вводя обозначение $z = (q, p)^T \in \mathcal{R}^{2d}$, гамильтонову систему (1.126) можно записать в следующей форме

$$\frac{dz}{dt} = J\nabla_z H(z). \quad (1.147)$$

Здесь введена кососимметричная *каноническая структурная матрица*

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} \in \mathcal{R}^{2d, 2d}, \quad (1.148)$$

где $I \in \mathcal{R}^d$ есть единичная матрица.

Если уравнения движения линейны и $H = \frac{1}{2}z^T Mz$, где $M \in \mathcal{R}^{2d, 2d}$, система (1.147) приводится к виду

$$\frac{dz}{dt} = JMz. \quad (1.149)$$

Матрица $A = JM$ называется *матрицей Гамильтона*. Если $M = M^T$, то решение (1.149) может быть выражено через собственные числа λ и собственные векторы v матрицы A . Так как $A^T = M^T J^T = -MJ$, собственная проблема представляется уравнением

$$-MJv = \lambda v. \quad (1.150)$$

Полагая здесь $w = Jv$, мы получаем соотношение

$$-Mw = \lambda J^{-1}w,$$

из которого после умножения на J имеем

$$-JMw = \lambda w.$$

Таким образом, $-\lambda$ также является собственным числом матрицы A , а соответствующий собственный вектор есть Jv . Пространство размерности $2d$ с координатами $p_1, \dots, p_d; q_1, \dots, q_d$ называется *фазовым пространством*, а правые части уравнений Гамильтона задают векторное поле, т.е. в каждой точке (p, q) фазового пространства определен $2d$ -мерный вектор $(-\partial H/\partial q, \partial H/\partial p)$.

Каждая гамильтонова система определяет однопараметрическое семейство *фазовых потоков*, представляющих собой группу преобразований, или отображений, фазового пространства

$$\varphi_t : (p(0), q(0)) \rightarrow (p(t), q(t)), \quad (1.151)$$

где $p(t), q(t)$ – решение уравнений Гамильтона. Другими словами, отображение φ_t сдвигает решение во времени, т.е. $\varphi_t(p^0, q^0) = (p(t, p^0, q^0), q(t, p^0, q^0))$ представляет собой решение в момент времени t , соответствующее начальным данным $p(0) = p^0, q(0) = q^0$.

Гладкое отображение $\Psi : \mathcal{R}^{2d} \rightarrow \mathcal{R}^{2d}$ на фазовом пространстве $\mathcal{R}^{2d}$ называется *симплектическим отображением* (относительно структурной матрицы J), если его якобиан $\Psi_z(z)$, $z = (p, q)$ удовлетворяет соотношению

$$[\Psi_z(z)]^T J^{-1} \psi_z(z) = J^{-1} \quad (1.152)$$

для всех z из области определения Ψ . Для линейного отображения $\Psi(z) = Bz$ условие симплектичности принимает вид

$$B^T J^{-1} B = J^{-1}, \quad (1.153)$$

а B при этом называется *симплектической матрицей*.

Справедливо следующее ключевое утверждение: фазовый поток $\varphi_t(z)$ гамильтоновой системы является симплектическим, т.е. якобиан этого отображения

$$\Psi_z(t) = \frac{\partial}{\partial z} \varphi_t(z) \quad (1.154)$$

удовлетворяет условию симплектичности (1.152).

Гамильтоновы уравнения в частных производных. Гамильтоновы формализм, рассмотренный выше для дискретных конечно-мерных систем, достаточно естественным образом переносится на функции непрерывных переменных, участвующих в уравнениях в частных производных, которые описывают процессы и явления, например, в механике сплошных сред или в теории полей.

Принципиальное отличие получаемых гамильтоновых уравнений заключается в том, что функции кинетической, потенциальной и полной энергий теперь относятся не к системе дискретных объектов, а представляют собой плотности соответствующих субстанций, отнесенных к единице объема. Точно также импульсы частиц заменяются на плотность импульса, а употребляемый по контексту термин гамильтониан в непрерывном случае является плотностью гамильтониана. Что касается фазовых пространств, то они становятся бесконечномерными полями, параметрически зависящими от времени и пространственных физических координат. Поскольку при этом исследуемые функции зависят не только от времени, но и от пространственных независимых переменных, для сохранения симметричной формы гамильтонового формализма необходимо ввести *функциональные, или вариационные, производные*. Например, если в случае одной пространственной переменной x функция $f(\xi)$ зависит от величины $\xi(t, x)$, то функциональная производная δ/ξ определяется формулой

$$\delta f / \delta \xi \equiv \partial f / \partial t - \frac{\partial}{\partial x} \frac{\partial t}{\partial (\partial \xi / \partial x)}. \quad (1.155)$$

Для иллюстрации рассмотрим одномерное квазилинейное волновое уравнение для скалярной функции $u(x, t)$:

$$u_{tt} = u_{xx} - f(u), \quad 0 < x < L, \quad (1.156)$$

где $f(u)$ – заданная достаточно гладкая функция, а $u(x, t)$ – L – периодическая гладкая функция. Более конкретно, будем считать решение удовлетворяющим граничному условию периодичности $u(x, t) = u(x + L, t)$, $L > 0$, и принадлежащим пространству $S = C^\infty[0, L]$. Полная энергия, или гамильтониан, для данной задачи имеет вид

$$H(\vec{u}) = \int_0^L \left[\frac{1}{2}v^2 + \frac{1}{2}u_x^2 + \mathcal{F}(u) \right] dx, \quad (1.157)$$

$$\mathcal{F}'(u) = f(u), \quad \vec{u} = (u, v)^T, \quad v = u_t,$$

а гамильтонова система записывается с помощью структурной матрицы J в следующей векторной форме:

$$J\vec{u}_t = \frac{\delta H}{\delta \vec{u}} = \left(\frac{\delta H}{\delta u}, \frac{\delta H}{\delta v} \right)^T, \quad J = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}. \quad (1.158)$$

Заметим, что если функция $f(u)$ явно не зависит от аргумента x , то для задачи (1.156), помимо сохранения во времени гамильтониана $H(\vec{u})$, существует еще и другой интеграл движения – полный момент

$$M(\vec{u}) = \int_0^L v u_x dx. \quad (1.159)$$

Справедливость этого утверждения следует из следующих несложно проверяемых соотношений:

$$\frac{d}{dt}M(\vec{u}) = \int_0^L (v_t u_x + v u_{xt}) dx = 0. \quad (1.160)$$

В качестве второго примера рассмотрим нелинейные уравнения типа Кортевега де Вриза [76]

$$\dot{u}(t, x) = \frac{1}{4}u''' + \frac{\partial}{\partial x}f'(u), \quad (1.161)$$

где $f(u)$ есть некоторая аналитическая функция (если $f(u) = \frac{1}{4}u^3$, то (1.161) представляет классическое уравнение Кортевега

де Вриза, обозначаемое как KdV), а на искомое решение $u(x, t)$ обычно накладываются нулевые граничные условия на бесконечности $x \rightarrow \pm\infty$, однако рассматриваются также и L -периодические решения. Соответствующий данному уравнению гамильтониан имеет вид

$$H(u) = \int \left[-\frac{1}{8}(u'(x))^2 + f(u) \right] dx, \quad (1.162)$$

где интеграл берется по расчетной области Ω .

Уравнение KdV возникло первоначально в теории мелкой воды, но впоследствии оказалось, что оно же встречается в целом ряде задач математической физики. В результате вычислительных экспериментов многими авторами было обнаружено существование решений типа уединенных бегущих волн, или *солитонов*, описываемых формулой

$$u(x, t) = \varphi(x - ct), \quad (1.163)$$

где $c \neq 0$ есть постоянная скорость волны, а φ – некоторая гладкая функция ее формы, сохраняющаяся при движении.

Более того, решение уравнения KdV может представляться и несколькими солитонами, движущимися с разными скоростями и проникающими друг сквозь друга без всякого взаимодействия. Это обстоятельство навело на мысль о существовании специальных законов сохранения. И действительно, исследованиями Крускала, Забусского, Лакса и других авторов было подтверждено существованием целой серии первых интегралов данной задачи.

Другим актуальным примером является комплексное *нелинейное уравнение Шредингера* (NLS), описывающее квантовую систему с переменным числом частиц, а также имеющее многочисленные приложения в нелинейной оптике, см. [174]:

$$\dot{u} = i \left[-\Delta u + V(x)u + \frac{\partial}{\partial \bar{u}} g(x, u, \bar{u}) \right], \quad (1.164)$$

где i – мнимая единица, V есть гладкая вещественная функция, а $g(x, u, v)$ – гладкая функция, являющаяся вещественной при $v = \bar{u}$ (черта означает комплексное сопряжение). Соответствующий NLS гамильтониан записывается в форме

$$H(u) = \frac{1}{2} \int [|\Delta u|^2 + V(x)|u|^2 + g(x, u, \bar{u})] dx. \quad (1.165)$$

В уравнении (1.164) пространственная переменная пробегает всю вещественную ось $-\infty < x < \infty$, а на бесконечности обычно ставятся асимптотические краевые условия решения. Типичными являются также квазипериодические граничные условия вида $u(x + 2L, t) = e^{i\theta} u(x, t)$, где величина $\theta \in [0, 2\pi)$ не зависит от времени t , а решение при этом ищется на интервале $-L \leq x < L$.

В случае $V(x) = \text{const}$ и $g = \gamma|u|^4$ из NLS получается *уравнение Захарова–Шабата*, см. [76].

В заключение данного пункта отметим появление такого нового направления, как гамильтонов *мульти-симплектический формализм*. Он основывается на рассмотрении времени не пассивно, а как равноправной независимой переменной, что приводит к определению дополнительных переменных и структурных матриц. Для иллюстрации мы приводим мульти-симплектическое представление для волнового уравнения (1.156):

$$Kz_t + Lz_x = \nabla S_z(z), \quad z = (u, v, w, \varphi)^T \in \mathcal{R}^4. \quad (1.166)$$

Здесь $K, L \in \mathcal{R}^{4,4}$ суть кососимметричные матрицы

$$K = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad L = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad (1.167)$$

а $\nabla S_z(z)$ обозначает обычный градиент в $\mathcal{R}^4$ от следующей алгебраической функции $S(z)$:

$$S(z) = \frac{1}{2}v^2 - \frac{1}{2}w^2 + \mathcal{F}(u). \quad (1.168)$$

Мульти-симплектическая форма гамильтонова уравнения зачастую имеет преимущества при построении и исследовании численных методов. Например, при использовании так называемых “поглощающих” граничных условий $u_t = \pm u_x$ при $x = 0, L$ краевая задача не описывается классическим гамильтоновым формализмом, поскольку в данном случае вариационная производная от $H(u)$ не обращается в нуль на границе области.

Глава 2

Обобщенные решения и вариационные формулировки краевых задач

Содержание данной главы – это расширение понятий классических постановок: определение обобщенных решений краевых задач для дифференциальных уравнений и их связь с классическими решениями, соответствующие вариационные формулировки для дифференциальных и абстрактных операторных постановок в общем несимметричном и симметричном случаях, а также вопросы корректности, существования и единственности обобщенных решений.

Исторически вариационные методы восходят еще к использованному Риманом так называемому принципу Дирихле, согласно которому решение краевой задачи 1-го рода сообщает функционалу энергии (“интегралу Дирихле”) минимальное значение. Принцип Дирихле был надолго заброшен после того, как Вейерштрасс в 1895 году показал на примере, что наименьшее значение интеграла Дирихле может и не достигаться.

Реабилитация вариационных подходов произошла в 1908 году, когда В.Ритц опубликовал свой метод решения минимизационных задач. После этого появилось много работ как по применению, так и по обоснованию метода Ритца, получившего еще название энергетического метода.

В 1913 году И.Г. Бубнов указал в небольшой заметке, что те же самые уравнения приближенного метода Ритца, получаемые из условия минимума функционала, могут быть выведены из совершенно другого принципа: говоря современным языком, им построены вариационные равенства, в которых базисные функции предполагались ортогональными, что обеспечивало простую формулу для решения. В 1915 году Б.Г. Галеркин развил этот подход (который еще можно назвать проекционным), отказавшись от условий ортогональности и вообще отойдя от изначальности минимизационной задачи, что позволило его применить формально для решения различных дифференциальных и интегральных уравнений.

Дальнейшее развитие вариационных методов в плане теоре-

тического обоснования и распространения на решение разнообразных задач математической физики связано с именами Р. Куранта, К. Фридрихса, Г.И. Петрова, М.В. Келдыша, М.А. Красносельского, М.И.Вишика, С.Г. Михлина и многих других известных математиков.

Математический аппарат вариационных постановок краевых задач базируется на понятиях обобщенных производных и на свойствах соболевских пространств, которые кратко описаны в предыдущей главе, а фундаментальное исследование данных проблем излагается во многих специальных монографиях: И. Бабушка и Б. Сабо [154], С.К. Бреннер и Л.Р. Скотта [26], Ф. Бреци и М. Фортина [27], О. Зенкевича и К. Моргана [77], В. Г. Корнеева [92], Г. И. Марчука и В. И. Агошкова [126], Ж.-П. Обэна [137], Л.А. Оганесяна, В.Я. Ривкинда и Л.А. Руховца [138, 139], Ф. Сьярле [173], В.В. Шайдурова [199] и других. Среди активно развиваемых в последние десятилетия можно отметить такие направления, как разрывные методы Галеркина [149], алгоритмы наименьших квадратов [62] и многомасштабные подходы [203], [204].

Огромная практическая роль и повсеместное распространение вариационных принципов обязаны рождению методов конечных элементов с финитными базисными функциями, история которых начинается с работы Р. Куранта 1943 г. Настоящая глава содержит описание основных принципов построения приближенных вариационных решений, а алгоритмам и технологиям МКЭ посвящена пятая глава.

2.1. Обобщенные решения дифференциальных уравнений

Пусть в области Ω n -мерного евклидова пространства задано эллиптическое уравнение второго порядка с так называемой дивергентной главной частью дифференциального оператора, имеющей вид $\operatorname{div}(a(\vec{x})\operatorname{grad}u)$:

$$Lu \equiv - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{i,j}(\vec{x}) \frac{\partial u}{\partial x_j} \right) + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} + cu = f(\vec{x}). \quad (2.1)$$

Требования на гладкость участвующих в (2.1) функций пока не конкретизируем, за исключением того, что данное дифференциальное уравнение должно иметь смысл.

Умножая это уравнение на некоторую функцию $v(\vec{x})$ и производя интегрирование обеих его частей по области Ω , с помощью справедливой для любых двух функций $u, v \in H^1(\Omega)$ (наличие этих свойств в (2.1) и ниже мы предполагаем) *формулы Грина*

$$\int_{\Omega} \sum_{i,j=1}^n \frac{\partial}{\partial x_i} a_{i,j}(\vec{x}) \frac{\partial u}{\partial x_j} v d\vec{x} = - \int_{\Omega} \sum_{i,j=1}^n a_{i,j}(\vec{x}) \frac{\partial u}{\partial x_j} \frac{\partial v}{\partial x_i} d\vec{x} + \int_{\Gamma} \frac{\partial u}{\partial N} v d\gamma \quad (2.2)$$

получаем интегральное соотношение

$$\begin{aligned} \int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} v + cuv \right) d\vec{x} \\ = \int_{\Omega} f v d\vec{x} + \int_{\Gamma} \frac{\partial u}{\partial N} v d\gamma. \end{aligned} \quad (2.3)$$

Очевидно, что функция $u(\vec{x})$, удовлетворяющая дифференциальному уравнению (2.1) и неоднородному граничному условию Неймана $\frac{\partial u}{\partial N}|_{\Gamma} = g$, при любой функции $v(\vec{x})$ удовлетворяет равенству, получаемому из (2.3) коррекцией последнего члена с учетом краевого условия:

$$\begin{aligned} \int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} v + cuv \right) d\vec{x} \\ = \int_{\Omega} f v d\vec{x} + \int_{\Gamma} g v d\gamma, \end{aligned} \quad (2.4)$$

если только все операции здесь имеют смысл. Обратное утверждение, вообще говоря, не справедливо, и именно поэтому важное значение имеет следующее понятие.

Определение 2.1. Функция $u \in H^1(\Omega)$, удовлетворяющая интегральному равенству (2.4) при любой функции $v(\vec{x}) \in H^1(\Omega)$, называется обобщенным (слабым) решением неоднородной задачи Неймана $\frac{\partial u}{\partial N}|_{\Gamma} = g$ для дифференциального уравнения (2.1).

Заметим, что к обобщенному решению предъявляются меньшие требования на гладкость, чем к классическому решению, которое должно обеспечивать существование вторых производных в (2.1).

Функции $v(\vec{x})$, участвующие в формулах (2.2)–(2.4), а также в определении обобщенного решения, будем называть *допустимыми функциями*.

Если конормальная производная на границе неизвестна (например, рассматривается первая краевая задача), то для перехода от (2.1) к (2.3) можно дополнительно потребовать равенства нулю функции $v(\vec{x})$ на границе Γ . В связи с этим имеет смысл новое определение.

Определение 2.2. Функция $u(\vec{x}) \in H_0^1(\Omega)$, удовлетворяющая равенству (2.4) при любой функции $v(\vec{x}) \in H_0^1(\Omega)$, называется обобщенным решением однородной задачи Дирихле ($u|_\Gamma = 0$) для дифференциального уравнения (2.1).

Напомним, что функции $v(\vec{x}) \in H_0^1(\Omega)$ удовлетворяют условию $v|_\Gamma = 0$, в силу чего поверхностный интеграл в (2.4) равен нулю при любом значении конормальной производной на границе.

В случае, когда решение $u(\vec{x})$ удовлетворяет неоднородному условию третьего рода

$$\alpha u + \frac{\partial u}{\partial N} = \beta, \quad (2.5)$$

из (2.1) вместо (2.4) следует равенство, содержащее дополнительный член с интегралом по границе:

$$\begin{aligned} \int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} v + cuv \right) d\vec{x} \\ = \int_{\Omega} f v d\vec{x} - \int_{\Gamma} \alpha u v d\gamma + \int_{\Gamma} \beta v d\gamma. \end{aligned} \quad (2.6)$$

Отсюда для однородной третьей краевой задачи естественно вводится определение, аналогичное предыдущим.

Определение 2.3. Функция $u \in H^1(\Omega)$, удовлетворяющая интегральному соотношению (2.6) при любой $v(\vec{x}) \in H^1(\Omega)$, называется обобщенным решением краевой задачи (2.1), (2.5).

Дальнейшее обобщение рассмотренных случаев заключается в переходе к смешанным краевым задачам, когда на различных участках границы $\Gamma = \Gamma_1 \cup \Gamma_2$ задаются граничные условия различных типов:

$$u|_{\Gamma_1} = 0, \quad \left(\alpha u + \frac{\partial u}{\partial N} \right) \Big|_{\Gamma_2} = \beta(\vec{x}). \quad (2.7)$$

При этом условия Неймана мы здесь специально не выделяем, так как формально они могут считаться частным случаем (2.5) при $\alpha = 0$.

Очевидно, что при краевых условиях (2.7) обобщенное решение определяется также интегральным равенством (2.6) с двумя существенными исключениями: поверхностные интегралы надо брать не по всей границе Γ , а по Γ_2 , а искомое решение $u(\vec{x})$ и допустимые функции $v(\vec{x})$ берутся из класса $H_0^1(\Omega, \Gamma_1)$ функций, обращающихся в нуль на Γ_1 .

Как видно из рассмотренных обобщенных формулировок, имеется принципиальная разница между граничными условиями первого рода, с одной стороны, и второго или третьего рода — с другой. А именно, допустимые функции должны удовлетворять однородным условиям Дирихле на Γ_1 ($v|_{\Gamma_1} = 0$) и не подчиняться никаким условиям на Γ_2 . В силу именно этих особенностей сложилась следующая классификация краевых условий: граничные условия Дирихле называются *главными*, а условия второго и третьего рода — *естественными граничными условиями*. Подчеркнем, что такое разделение никак не следует из классической постановки краевых задач.

Отметим еще такой практически важный случай, когда коэффициенты $a_{i,j}(\vec{x})$ имеют разрыв первого рода на некоторой «внутренней» поверхности Γ_i , которая может быть многосвязной и делит расчетную область Ω на m непересекающихся под областей Ω_k , $\Omega = \cup_{k=1}^m \Omega_k$. Тогда применение формулы Грина вместо (2.3) дает следующее соотношение:

$$\begin{aligned} \sum_{k=1}^m \int_{\Omega_k} \left(\sum_{i,j} a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} v + cuv - fv \right) d\vec{x} \\ = \int_{\Gamma} \frac{\partial u}{\partial N} v d\gamma + \int_{\Gamma_i} \left(\frac{\partial u}{\partial N} \Big|_{\Gamma_i^+} - \frac{\partial u}{\partial N} \Big|_{\Gamma_i^-} \right) v d\gamma, \end{aligned}$$

где в последнем члене участвуют односторонние конормальные производные по разные стороны внутренней границы Γ_i^+ , Γ_i^- . В прикладном отношении такие задачи означают, что расчетная область Ω состоит из под областей Ω_k с различными материальными свойствами.

Аналогичным образом определяются обобщенные решения для дифференциальных уравнений высокого порядка и для систем уравнений.

Например, для приведенной выше векторной системы уравнений Ламе (1.41), (1.42) это можно сделать, если ввести гильбертовы пространства векторных функций $\bar{L}_2(\Omega)$ и $\bar{H}^1(\Omega)$ со следующими определениями скалярных произведений:

$$\begin{aligned} (\bar{\sigma}, \bar{\tau})_{\bar{L}_2(\Omega)} &= \int_{\Omega} \bar{\sigma} : \bar{\tau} d\bar{x}, \\ (\bar{u}, \bar{v})_{\bar{H}^1(\Omega)} &= (\overline{\text{grad}} \bar{u}, \overline{\text{grad}} \bar{v})_{\bar{L}_2(\Omega)} + (u, v)_{\bar{L}_2(\Omega)}, \end{aligned}$$

где для тензоров используются обозначения

$$\overline{\text{grad}} \bar{u} = \left\{ \frac{\partial u_i}{\partial x_j} \right\}, \quad \bar{\sigma} : \bar{\tau} = \sum_{i,j} \sigma_{i,j} \tau_{i,j}.$$

Если векторная функция $\bar{u} \in \bar{H}^1(\Omega)$ удовлетворяет системе уравнений теории упругости, то, умножая обе ее части на некоторую функцию $\bar{v} \in \bar{H}^1(\Omega)$ с однородным граничным условием $\bar{v}|_{\Gamma_1} = 0$, после интегрирования по частям получаем

$$\int_{\Omega} [2\mu \bar{\varepsilon}(\bar{u}) + \lambda \text{tr}(\bar{\sigma}(\bar{u})) \bar{\delta}] : \overline{\text{grad}} \bar{v} d\bar{x} - \int_{\Gamma_2} \bar{\sigma}(u) \bar{\nu} \cdot \bar{v} ds = \int_{\Omega} \bar{f} \cdot \bar{v} d\bar{x}.$$

Взаимосвязь между обобщенным и классическим решениями есть вопрос для специального изучения, и далее мы продолжим рассмотрение постановок задач для обобщенных решений.

2.2. Вариационная постановка краевых задач

Очевидно, что интегральное соотношение (2.3) порождает билинейную форму $a(\cdot, \cdot) : H^1 \times H^1 \rightarrow R$,

$$a(u, v) = \int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + \sum_{i=1}^n b_i \frac{\partial u}{\partial x_i} v + cuv \right) d\bar{x}, \quad (2.8)$$

и линейную форму $f : H^1 \rightarrow R$,

$$f(v) = \int_{\Omega} f v d\vec{x}. \quad (2.9)$$

С помощью введенных обозначений задача нахождения обобщенного решения из интегрального равенства вида (2.4) или (2.6) представляет собой частный случай так называемой *абстрактной вариационной задачи*, записываемой в форме *вариационного равенства*: найти такой элемент $u \in U$, что для любого v из гильбертового пространства V выполняется равенство

$$a(u, v) = f(v), \quad (2.10)$$

где $U \subset V$ — подпространство, $a(\cdot, \cdot) : V \times V \rightarrow R$ — билинейная форма и $f : V \rightarrow R$ — линейная форма.

Для рассмотренной краевой задачи линейной теории упругости (1.41), (1.45), (1.46) билинейная форма определяется следующим образом:

$$a(\bar{u}, \bar{v}) = \int_{\Omega} [2\mu \bar{\varepsilon}(\bar{u}) : \bar{\varepsilon}(\bar{v}) + \lambda \operatorname{div} \bar{u} \operatorname{div} \bar{v}] d\vec{x}.$$

При этом вариационная задача формулируется как нахождение векторной функции $\bar{u} \in \bar{H}^1(\Omega)$, удовлетворяющей краевому условию $\bar{u}|_{\Gamma_1} = \bar{g}$ и уравнению

$$a(\bar{u}, v) = \int_{\Omega} \bar{f} \cdot \bar{v} d\vec{x} + \int_{\Gamma_2} \bar{t} \cdot \bar{v} ds$$

для любой функции $\bar{v} \in \bar{V} = \{\bar{v} \in \bar{H}^1(\Omega) : \bar{v}|_{\Gamma_1} = 0\}$.

По аналогии с классической операторной постановкой естественно ввести следующее понятие.

Определение 2.4. Абстрактная вариационная задача называется корректной, если ее решение существует, единственно и непрерывно зависит от исходных данных.

Если билинейная форма порождается дифференциальным оператором дивергентного вида, получаемым из (2.1) при $a_{i,j}(\vec{x}) = a_{j,i}(\vec{x})$ и $b_i = 0$, то она является симметричной:

$$a(u, v) = a(v, u) = \int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + cuv \right) d\vec{x}. \quad (2.11)$$

Очевидно, что если $c \geq 0$ и выполняются условия (1.17), (1.18) равномерной эллиптичности дифференциального оператора в области Ω , то билинейная форма (2.11) будет $H^1(\Omega)$ -эллиптической.

Рассмотрим теперь конкретизацию абстрактной вариационной постановки, связанную с характерной краевой задачей.

Пусть Ω — ограниченная в R^n область с липшицевой границей Γ , являющейся объединением своих частей Γ_1, Γ_2 . Определим пространство

$$U = V = H_0^1(\Omega, \Gamma_1) = \{v(\vec{x}) \in H^1(\Omega); v = 0 \text{ на } \Gamma_1\}, \quad (2.12)$$

билинейную форму — в соответствии с выражением (2.11), а также линейную форму

$$f(v) = \int_{\Omega} f v d\vec{x} + \int_{\Gamma_2} g v d\gamma. \quad (2.13)$$

Относительно функций $a_{ij}(\vec{x})$, $c(\vec{x})$, $f(\vec{x})$ и $g(\vec{x})$ предполагаем, что они удовлетворяют следующим условиям:

$$\begin{aligned} a_{i,j} &\in L_{\infty}(\Omega), \quad 1 \leq i, j \leq n, \quad c(\vec{x}) \in L_{\infty}(\Omega), \\ f &\in L_2(\Omega), \quad g \in L_2(\Gamma_2), \quad c(\vec{x}) \geq 0, \\ \sum_{i,j=1}^n a_{i,j}(\vec{x}) y_i y_j &\geq \alpha \|y\|^2, \quad \alpha > 0 \text{ при } x \in \Omega. \end{aligned} \quad (2.14)$$

Покажем теперь, что вариационная задача вида (2.9), определяемая соотношениями (2.11)–(2.14), порождает непрерывную (ограниченную) линейную форму $f(v)$, а также билинейную форму $a(u, v)$, которая удовлетворяет условиям непрерывности и V -эллиптичности.

Во-первых, из условий (2.14) следует, что билинейная форма (2.11) является $H_0^1(\Omega, \Gamma_1)$ -эллиптической, так как для всех $v \in H^1(\Omega)$ справедливо неравенство

$$a(v, v) = \int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \frac{\partial v}{\partial x_i} \frac{\partial v}{\partial x_j} + cv^2 \right) d\vec{x} \geq \alpha |v|_1^2, \quad (2.15)$$

и полунорма $|\cdot|_1$ есть норма в пространстве $H_0^1(\Omega, \Gamma_1)$.

Во-вторых, рассматриваемая билинейная форма непрерывна, так как для всех $u, v \in H^1(\Omega)$ при условиях (2.14) имеем

$$\begin{aligned}
|a(u, v)| &\leq \sum_{i=1}^n a_i |\partial_i u|_0 |\partial_i v|_0 + |c|_\infty |u|_0 |v|_0 \\
&\leq \max\{|a|, |c|_\infty\} \|u\|_1 \|v\|_1,
\end{aligned} \tag{2.16}$$

где $|a| = \max_i \{a_i = |a_{i,1}|_\infty + |a_{i,2}|_\infty\}$, а символы $|\cdot|_0$, $|\cdot|_\infty$ обозначают, соответственно, нормы в пространствах $L_2(\Omega)$ и $L_\infty(\Omega)$.

В третьих, линейная форма $f(v)$ из (2.13) непрерывна, так как для всех $v \in H^1$, $f(\vec{x}) \in L_2(\Omega)$ и $g(\vec{x}) \in L_2(\Gamma_2)$ выполняется неравенство

$$\begin{aligned}
|f(v)| &\leq |f|_0 \cdot |v|_0 + \|g\|_{L_2(\Gamma_2)} \cdot \|v\|_{L_2(\Gamma_2)} \\
&\leq (\|f\|_{L_2(\Omega)} + C_1(\Omega) \|g\|_{L_2(\Gamma_2)}) \|v\|_1.
\end{aligned} \tag{2.17}$$

При выводе последнего выражения использовано соотношение между нормами в пространствах $L_2(\Gamma)$ и $H^1(\Omega)$, которое можно сформулировать следующим образом:

$$\|v\|_{L_2(\Gamma)} \leq C_1(\Omega) \|v\|_{H^1(\Omega)}, \tag{2.18}$$

где $C_1(\Omega)$ — некоторая постоянная.

Возвращаясь к неравенствам (2.16) и (2.17), отмечаем, что они устанавливают ограниченность форм $a(u, v)$, $f(v)$, что и означает их непрерывность.

В качестве примера построения несимметричной билинейной формы, обладающей свойствами непрерывности и коэрцитивности, для простоты рассмотрим одномерную краевую задачу Дирихле

$$-u'' + u' + u = f(x), \quad x \in (a, b), \quad u(a) = u(b) = 0.$$

Сформулируем вариационную задачу в соответствии со следующими определениями:

$$\begin{aligned}
V &= H_0^1(a, b), \quad a(u, v) = \int_a^b (u'v' + u'v + uv) dx, \\
f(v) &= (f, v) = \int_a^b f v dx.
\end{aligned}$$

Доказательство непрерывности билинейной формы следует из соотношений

$$\begin{aligned}
|a(u, v)| &\leq |(u, v)_{H^1}| + \left| \int_a^b u'v \, dx \right| \\
&\leq \|u\|_{H^1} \|v\|_{H^1} + \|u'\|_{L_2} \cdot \|v\|_{L_2} \leq 2\|u\|_{H^1} \|v\|_{H^1}.
\end{aligned}$$

Для анализа коэрцитивности достаточно рассмотреть неравенство

$$\begin{aligned}
a(v, v) &= \int_a^b [(v')^2 + v'v + v^2] \, dx \\
&= \frac{1}{2} \int_a^b (v' + v)^2 \, dx + \frac{1}{2} \int_a^b [(v')^2 + v^2] \, dx \geq \frac{1}{2} \|v\|_{H^1}^2.
\end{aligned}$$

Первой проблемой при исследовании абстрактной вариационной задачи является вопрос о существовании и единственности ее решения. Здесь исчерпывающий ответ дается теоремой Вишика–Лакса–Мильграма, основанной на существовании неподвижной точки сжимающего отображения.

Теорема 2.1 (о сжимающем отображении). Пусть даны банаховое пространство V и отображение $T : V \rightarrow V$, удовлетворяющее неравенству

$$\|Tv_1 - Tv_2\| \leq M \|v_1 - v_2\|$$

для всех $v_1, v_2 \in V$ и фиксированной константы $0 < M < 1$. Тогда существует единственное $u \in V$ такое, что

$$u = Tu,$$

т. е. сжимающее отображение T имеет единственную неподвижную точку u .

Теорема 2.2 (Вишика–Лакса–Мильграма). Пусть V — гильбертово пространство, $a(\cdot, \cdot)$ — непрерывная, коэрцитивная, билинейная форма на V , а $f \in V'$ — непрерывный линейный функционал. Тогда существует единственное $u \in V$ такое, что

$$a(u, v) = f(v) \quad \forall v \in V.$$

Данную вариационную задачу можно сформулировать как нахождение единственного u , обеспечивающего соотношение

$$Au(v) = f(v) \quad \forall v \in V$$

или, для любого $f \in V'$, равенство функционалов

$$Au = f. \quad (2.19)$$

Фактически в условия последней теоремы справедлива цепочка неравенств

$$c_2 \|v\|^2 \leq a(v, v) = Av(v) \leq \|Av\|_{V'} \|v\| \quad \forall v \in V,$$

которая устанавливает существование непрерывного обратного оператора $A^{-1}: V' \rightarrow V$ и оценку

$$\|A^{-1}\| \leq \frac{1}{c_2}.$$

Таким образом, вариационная задача (2.10)–(2.14) имеет единственное решение $u \in H_0^1(\Omega, \Gamma_1)$, непрерывно зависящее от f и g . Само вариационное уравнение в данном случае имеет вид

$$\int_{\Omega} \left(\sum_{i,j=1}^n a_{i,j} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_j} + cuv \right) d\vec{x} = \int_{\Omega} f v d\vec{x} + \int_{\Gamma_2} g v d\gamma.$$

С помощью формулы Грина легко показывается, что его решение формально соответствует решению следующей смешанной краевой задачи:

$$\begin{aligned} - \sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{i,j} \frac{\partial u}{\partial x_j} \right) + cu &= f(\vec{x}), \quad \vec{x} \in \Omega, \\ u(\vec{x}) &= 0, \quad \vec{x} \in \Gamma_1, \quad \frac{\partial u}{\partial N} = g(\vec{x}), \quad \vec{x} \in \Gamma_2. \end{aligned}$$

Рассмотренную вариационную постановку можно переформулировать как задачу нахождения обобщенного решения для абстрактного операторного уравнения (2.19) в соответствующих функциональных пространствах.

Пусть A — линейный оператор с областью определения $D(A)$ из гильбертового пространства H . Тогда вариационная задача, соответствующая уравнению (2.19), в терминах скалярных произведений формулируется как нахождение $u \in H$, удовлетворяющего вариационному равенству

$$(Au, v) = (f, v) \quad \forall v \in H. \quad (2.20)$$

Отметим, что исследование вопросов существования и единственности обобщенных решений для систем дифференциальных уравнений, например, для рассмотренных в начале данного пункта вариационных задач теории упругости, остается принципиально таким же, но в математическом плане включает изучение гораздо более сложных проблем, которым посвящаются специальные работы или монографии. Прекрасной иллюстрацией здесь является книга П. Монка [133] с фундаментальными исследованиями свойств системы уравнений Максвелла для векторных электромагнитных полей.

2.3. Вариационная задача минимизации в симметричном случае

Рассмотрим вариационную задачу вида (2.9), в которой билинейная форма $a(u, v)$ является симметричной, непрерывной и положительно определенной. Если u — решение (единственное) такой задачи, то можно записать равенство

$$a(u - v, u - v) = a(u, u) - 2a(u, v) + a(v, v) = a(v, v) - 2(f, v) + (f, u) = J(v) + (f, u) \geq 0. \quad (2.21)$$

Здесь мы воспользовались *теоремой Рисса* для представления линейного функционала $f(v)$ в виде скалярного произведения (f, v) , а также ввели обозначение

$$J(v) \equiv a(v, v) - 2(f, v)$$

для так называемого *функционала энергии*. Очевидно, что данный функционал ограничен снизу, т.е. $J(v) \geq -(f, u)$, и достигает своего единственного минимума при $v = u$, так как в этом и только в этом случае квадратичная форма $a(u - v, u - v)$ в (2.21) равна нулю.

Таким образом, в симметричном случае абстрактная вариационная задача (2.9) сводится к нахождению минимума функционала энергии и может быть сформулирована следующим образом: пусть V — нормированное векторное пространство, $a(\cdot, \cdot) : V \times V \rightarrow R$ — симметричная, коэрцитивная и непрерывная билинейная форма, $f : V \rightarrow R$ — непрерывная линейная форма, U — непустое подмножество пространства V ; найти такой элемент $u \in U$, что

$$J(u) = \inf_{v \in U} \{J(v)\}, \quad (2.22)$$

где J — функционал, определяемый как

$$J : v \in V \rightarrow J(v) = a(v, v) - 2f(v). \quad (2.23)$$

Очевидно, что в рассматриваемом случае билинейная форма $a(\cdot, \cdot)$ является скалярным произведением на пространстве V . Это позволяет определить *энергетическое* (гильбертовое) *пространство* с определением нормы

$$\|v\|_E = [a(v, v)]^{1/2}.$$

Заметим, что данная *абстрактная задача минимизации*, сформулированная нами как следствие более общей несимметричной вариационной задачи (2.9), исторически появилась и исследована раньше. Более того, название *вариационная задача*, или вариационный метод, возникло как подход к решению дифференциального уравнения, заключающийся в сведении к равносильной задаче об отыскании функции, сообщающей наименьшее значение некоторому интегралу. Если последний представляет собой квадратичный положительно определенный функционал, что особо важно во многих приложениях и связано обычно с понятием энергии, то такой метод еще носит название *энергетического*.

Хотя вопросы существования и единственности решения симметричной вариационной задачи полностью определяются теоремой Вишика–Лакса–Мильграма, в методическом плане полезно рассмотреть следующее утверждение.

Теорема 2.3. Пусть V — полное пространство, U — замкнутое выпуклое подмножество из V , $a(\cdot, \cdot)$ — симметричная, непрерывная и V -эллиптическая форма, а $f(v)$, $v \in V$, — непрерывная линейная форма. Тогда абстрактная задача минимизации (2.22), (2.23) имеет единственное решение.

Дадим теперь эквивалентные формулировки рассматриваемой задачи.

Теорема 2.4. Элемент $u \in U$ является решением абстрактной задачи минимизации (2.22), (2.23) тогда и только тогда, когда он удовлетворяет соотношениям

$$a(u, v - u) \geq f(v - u) \quad \forall v \in U, \quad (2.24)$$

если U — замкнутое выпуклое множество из V , и

$$a(u, v) = f(v) \quad \forall v \in U, \quad (2.25)$$

если U — замкнутое подпространство.

Соотношения вида (2.24) называются *вариационными неравенствами*.

Рассмотрим несколько простых примеров вариационных задач минимизации. Пусть V — конечномерное пространство, $u, f, v = \{v_i : i = 1, \dots, N\}$ — вещественные векторы, а $A = \{a_{ij}\}$ — симметричная положительно определенная матрица. Тогда условие минимума функционала

$$J(v) = (Av, v) - 2(f, v), \quad (2.26)$$

где $(\cdot, \cdot)$ — обычное скалярное произведение в вещественном евклидовом пространстве, определяется из системы уравнений

$$\frac{\partial J}{\partial v_i} = 2 \left(\sum_{j=1}^n a_{ij} v_j - f_i \right) = -2r_i = 0. \quad (2.27)$$

Точнее говоря, эти соотношения являются условиями экстремума, но поскольку

$$\frac{\partial^2 J}{\partial v_i^2} = 2a_{ii} > 0,$$

в силу положительной определенности матрицы A , то в данном случае действительно достигается минимум.

Отметим еще, что введенный в (2.27) вектор невязки $r = \{r_i = (f - Av)_i\}$ представляет антиградиент убывания функционала $J(v)$ в точке v , а вектор u , обеспечивающий его минимум и определяемый с помощью (2.25), является решением алгебраической системы

$$Au = f,$$

которое называется *уравнением Эйлера* для функционала (2.26).

В более общем случае проблема нахождения функции, доставляющей минимум (или максимум) некоторому функционалу, представляет собой задачу *вариационного исчисления* и связана с построением соответствующего дифференциального уравнения Эйлера при необходимых краевых условиях. Принципы таких постановок можно проиллюстрировать на минимизации функционала

$$J(u) = \int_0^1 \varphi(x, u, u') dx, \quad (2.28)$$

где $\varphi = \varphi(x, y, z)$ — заданная функция, непрерывная вместе со своими первыми производными до второго порядка включительно относительно переменных x, y, z в некоторой области трехмерного евклидова пространства.

Предположим, что искомая функция $u(x)$ непрерывна, имеет непрерывную на отрезке $(0, 1)$ производную $u'(x)$ и удовлетворяет граничным условиям

$$u(0) = g_0, \quad u(1) = g_1.$$

Введем новую функцию $u_\alpha(x) = u(x) + \alpha v(x)$, где α — малый параметр, а v — некоторая достаточно гладкая функция с условиями на концах отрезка

$$v(0) = v(1) = 0.$$

Подставляя эту функцию в функционал J , получаем

$$J(u_\alpha) = \int_0^1 \varphi(x, u + \alpha v, u' + \alpha v') dx.$$

Данный функционал можно рассматривать как функцию от параметра α : $J(u_\alpha) = w(\alpha)$. При этом производную функции $w(\alpha)$ в точке $\alpha = 0$ называют *первой вариацией* функционала и обозначают обычно символом δJ :

$$\delta J = \frac{dw}{d\alpha} \Big|_{\alpha=0} = \int_0^1 (\varphi_u v + \varphi_{u'} v') dx = \int_0^1 v(x) \left(\varphi_u - \frac{d}{dx} \varphi_{u'} \right) dx, \quad (2.29)$$

где $\varphi_u = \partial\varphi/\partial u$, а последнее выражение получено с помощью интегрирования по частям.

Вторая вариация функционала J определяется как вторая производная функции $w(\alpha)$ в точке $\alpha = 0$ и обозначается через $\delta^2 J$:

$$\delta^2 J = \frac{d^2 w}{d\alpha^2} \Big|_{\alpha=0} = \int_0^1 [\varphi_{u'u'} (v')^2 + 2\varphi_{uu'} v v' + \varphi_{uu} v^2] dx. \quad (2.30)$$

Поскольку необходимым условием экстремума функции $w(\alpha)$ при $\alpha = 0$ является равенство нулю первой вариации функционала δJ , то в силу произвольности $v(x)$ получаем, что функция $u(x)$, доставляющая минимум функционалу $J(u_\alpha)$, должна удовлетворять дифференциальному уравнению

$$\varphi_u - \frac{d}{dx} \varphi_{u'} = 0, \quad (2.31)$$

которое является уравнением Эйлера для вариационной задачи, связанной с функционалом (2.28). Очевидно, что если при выполнении уравнения (2.31) вторая вариация функционала (2.30) положительна или отрицательна, то функция $u(x)$ доставляет функционалу J , соответственно, минимум или максимум.

Проиллюстрированный на простейшем примере принцип вариационного исчисления является достаточно общим и самостоятельным направлением, применимым и к нелинейным, и к многомерным, и к динамическим задачам.

Рассмотрим еще вариационную задачу минимизации, порождаемую одномерной смешанной однородной краевой задачей

$$Lu \equiv -(au')' + cu = f(x), \quad x \in (0, 1), \quad u(0) = u'(1) = 0. \quad (2.32)$$

В данном случае линейную форму определим через скалярное произведение функций

$$f(v) = (f, v) = \int_0^1 f v dx,$$

а билинейную форму запишем с помощью интегрирования по частям:

$$a(u, v) = \int_0^1 (-(au')' + cu)v \, dx = \int_0^1 (au'v' + cuv) \, dx + au'v|_0^1. \quad (2.33)$$

Определим пространство V допустимых функций $v(x)$ как

$$V = \{v \in H^1(0, 1), v(0) = 0\}. \quad (2.34)$$

Тогда с учетом граничного условия $u'(1) = 0$ из (2.29) определяется функционал

$$J(v) = \int_0^1 [a(v')^2 + cv^2 - 2fv] \, dx, \quad (2.35)$$

минимизация которого и характеризует симметричную вариационную задачу.

На этом примере легко провести и обратные рассуждения, т. е. что минимум функционала (2.35) достигается именно на решении исходной краевой задачи, или другими словами, (2.32) является уравнением Эйлера для функционала $J(v)$.

Действительно, если $u \in V$ — точка минимума $J(v)$, то для любого числа ε и любой функции $v \in V$ имеем

$$\begin{aligned} J(u) \leq J(u + \varepsilon v) &= J(u) + 2\varepsilon \int_0^1 (au'v' + cuv - fv) \, dx + \\ &\quad \varepsilon^2 \int_0^1 [a(v')^2 + cv^2] \, dx. \end{aligned} \quad (2.36)$$

Так как величина ε может быть любого знака, линейный член (первая вариация) должен отсутствовать, откуда с помощью формального интегрирования по частям (которое предполагаем выполнимым) получаем

$$0 = \int_0^1 (au'v' + cuv - fv) \, dx = \int_0^1 [-(au')' + cu - f]v \, dx + au'v|_0^1. \quad (2.37)$$

Поскольку данное равенство должно выполняться для любой функции $v(x) \in V$ из (2.34), функция $u(x)$ должна удовлетворять дифференциальному уравнению $Lu = f$ и, кроме того, граничному условию $u'(1) = 0$. А так как минимизация функционала осуществляется по функциям $v(x)$, удовлетворяющим

в нуле однородному условию Дирихле, то получаемое решение также будет удовлетворять условию $u(0) = 0$.

Таким образом, если участвующие в (2.37) вторые производные имеют смысл, то обобщенное решение вариационной задачи является классическим решением краевой задачи (2.32).

Подчеркнем еще раз отличие между главными и естественными граничными условиями: в данной вариационной задаче от допустимых, или пробных, функций $v(x)$ мы требовали удовлетворения условию Дирихле в нуле, но не ставили никакого условия на другом конце отрезка, — и условие Неймана $u'(1) = 0$ появилось автоматически как следствие задачи минимизации.

Рассмотрим теперь случай, когда коэффициент $a(x)$ в (2.35) терпит разрыв первого рода и интегрирование по частям в (2.36) неправомерно. Пусть такой скачок (внутренняя граница) будет в одной точке $\xi \in (0, 1)$, и тогда вместо (2.35) мы должны записать

$$\int_0^{\xi} [-(au')' + cu - f]v dx + au'v|_0^{\xi-} + \int_{\xi}^1 [-(au')' + cu - f]v dx + au'v|_{\xi+}^1 = 0.$$

Таким образом, решение вариационной задачи $u(x)$ должно удовлетворять (помимо уравнений (2.32)) дополнительно так называемым *условиям сопряжения*

$$au'|_{\xi-} = au'|_{\xi+},$$

где знаки «-», «+» означают односторонние предельные значения функций и производных.

Остановимся теперь на особенностях наличия неоднородного краевого условия третьего рода и запишем граничные условия в виде

$$u(0) = 0, \quad \alpha u(1) + u'(1) = \beta. \quad (2.38)$$

Как очевидно следует из (2.33), (2.37), симметричная вариационная задача в данном случае (коэффициент $a(x)$ будем считать теперь гладким) сводится к минимизации функционала

$$J(v) = \int_0^1 [a(v')^2 + cv^2 - 2fv] dx + a(1)[\beta v(1) - \alpha v^2(1)] \quad (2.39)$$

на функциях $v(x)$ из пространства V в (2.34). При этом, в соответствии с (2.34), (2.35), первая вариация функционала (2.39) равна

$$0 = \int_0^1 [-(au')' + cu - f]v \, dx + a(1)[\beta - \alpha u(1) - u'(1)]v(1) \quad (2.40)$$

$$\forall v \in V.$$

Таким образом, неоднородное краевое условие третьего рода (2.38) является также естественным в том смысле, что оно оказывается следствием вариационной постановки, без наложения каких-либо условий на пробные функции $v(x)$ в точке $x = 1$.

Рассмотрим теперь переход от однородного условия Дирихле к неоднородному:

$$u(0) = g, \quad \alpha u(1) + u'(1) = \beta.$$

Очевидно, что для такой краевой задачи вариационная постановка должна включать минимизацию функционала энергии на функциях, удовлетворяющих неоднородному граничному условию первого рода. Но поскольку множество таких функций не является линейным пространством, то с формулировкой поступим следующим образом. Определим в $\Omega = (0, 1)$ некоторую функцию $v_g(x) \in H^1(\Omega)$, удовлетворяющую требуемому условию $v_g(0) = g$. Тогда для искомого решения мы имеем $\tilde{u}(x) = u(x) - v_g(x) \in V$, и условия минимума для $J(v)$ вида (2.39) опять можно записать в форме

$$J(\tilde{u}) \leq J(\tilde{u} + \varepsilon v) \quad \forall v \in V, \quad \forall \varepsilon \in R.$$

Отсюда первая вариация функционала энергии снова приводит к соотношению (2.36), т. е. неоднородное условие третьего рода в точке $x = 1$ и здесь оказывается естественным следствием задачи минимизации.

В заключение данного пункта обратимся к операторной формулировке симметричной вариационной задачи.

Пусть H — вещественное гильбертово пространство со скалярным произведением $(\cdot, \cdot)$ и нормой $\|\cdot\|$, A — самосопряженный и положительно определенный в $D(A)$ оператор, а его область определения является всюду плотной в H , т. е. $D(A) \subset H$, $\overline{D(A)} = H$ ($\forall v \in H$ и $\forall \varepsilon > 0$ существует $u \in D(A)$ такое, что

$\|v - u\| < \varepsilon$). Пусть также $f(v)$ — линейный ограниченный функционал с областью определения $D(f) = H$. Рассмотрим задачу о нахождении элемента $u \in H$ (если он существует), реализующего минимум определенного на $D(A)$ функционала

$$\begin{aligned} J(u) &= \inf_{v \in H} \{J(v) = (Av, v) - 2f(v)\}, \\ (Av, v) &\geq \alpha(v, v), \quad \alpha > 0. \end{aligned} \quad (2.41)$$

Поскольку оператор A определен не на всех $v \in H$, приведем построение нового *энергетического* (гильбертового) *пространства* согласно реализованному ранее приему, см. п. 1.2.

На элементах $u, v \in D(A)$ определим скалярное произведение и норму

$$(u, v)_A = (Au, v), \quad \|u\|_A = (Au, u)^{1/2}, \quad (2.42)$$

которые определяют некоторое новое пространство. Если оно не полно (вследствие $D(A) \neq H$), пополняя $D(A)$ по норме $\|\cdot\|_A$, получаем гильбертовое пространство H_A со скалярным произведением $(\cdot, \cdot)_A$, которое называется энергетическим пространством.

Теорема 2.5. *Пространство H_A вложено в H , т. е. все элементы H_A принадлежат также пространству H , причем*

$$\|v\| \leq \alpha^{1/2} \|v\|_A \quad \forall v \in H_A, \quad \alpha > 0. \quad (2.43)$$

При $D(A) = H$ очевидным образом имеем $H_A = H$.

Таким образом, функционал энергии (2.41) можно рассматривать как функционал в H_A и записать в виде

$$J(v) = \|v\|_A^2 - 2f(v). \quad (2.44)$$

Линейный функционал $f(v)$, будучи ограниченным на $D(A)$, является ограниченным на этом множестве и в пространстве H_A , ибо

$$|f(v)| \leq \|f\|_{H'} \|v\| \leq \alpha^{-1/2} \|f\|_{H'} \|v\|_A.$$

Рассмотрим его продолжение по непрерывности со всюду плотного в H_A множества $D(A)$ на все пространство H_A . Тогда по теореме Рисса существует единственный элемент $u \in H_A$ такой, что $f(v) = (u, v)_A$, и квадратичный функционал (2.44) можно записать как

$$J(v) = (v, v)_A^2 - 2(u, v)_A = \|u - v\|_A^2 - \|u\|_A^2. \quad (2.45)$$

Последняя формула показывает, что $J(v)$ достигает своего наименьшего значения $-\|u\|_A^2$ при единственном условии $v = u$, т. е. фактически получено независимое доказательство о существовании и единственности решения симметричной абстрактной вариационной задачи.

Квадратичный функционал

$$\|u - v\|_A^2 = (A(u - v), (u - v)) \quad (2.46)$$

называется *функционалом ошибок*, его минимум равен нулю и достигается при $v = u$.

Пусть теперь $u \in D(A)$ — (единственное!) решение операторного уравнения $Au = f$, а $v \in D(A)$ — произвольный элемент. Поскольку

$$(A(u - v), (u - v)) = (Av, v) - 2(f, v) + (f, u) = J(v) + (f, u) \geq 0,$$

то из получаемого отсюда неравенства

$$J(v) \geq -(f, u) = (Au, u) = \|u\|_A^2 = J(u) \quad (2.47)$$

имеем следующий результат.

Теорема 2.6. Если $u \in D(A)$, $\overline{D(A)} = H$, есть решение уравнения $Au = f$ с самосопряженным и положительно определенным в $D(A)$ оператором, то u минимизирует в $D(A)$ функционал энергии $J(v)$.

Справедливо и обратное утверждение.

Теорема 2.7. Если $u \in D(A)$ минимизирует функционал $J(v)$ в $D(A)$, то u есть решение операторного уравнения $Au = f$.

Очевидно, что функционал $J(v)$ с самосопряженным и положительно определенным в $D(A)$ оператором A всегда имеет минимум в H_A . Но если данный минимум достигается на элементе $u \notin D(A)$, то этот элемент не является классическим решением уравнения $Au = f$, а только обобщенным.

Определение 2.5. Последовательность $\{u_n\} \in H_A$ называется минимизирующей энергетический функционал $J(v)$, если при $n \rightarrow \infty$

$$J(u_n) \rightarrow \inf_{v \in H_A} J(v) = J(u), \quad u \in H_A.$$

Теорема 2.8. Если последовательность $\{u_n\}$ — минимизирующая, то $u_n \rightarrow u$ в норме пространства H_A .

2.4. Приближенные методы решения вариационных задач

Данный параграф посвящен изложению общих принципов построения приближенных решений абстрактных вариационных задач: метода Ритца для симметричного случая и Галеркина — для несимметричного, а также вопросам сходимости и оценкам погрешностей получаемых приближений.

2.4.1. Метод Ритца

Определение 2.6. Пусть V есть гильбертово пространство и $a(\cdot, \cdot)$ — симметричная непрерывная билинейная V -эллиптическая форма. Пусть также $V_h \subset V$ есть конечномерное подпространство и $f(v) \in V' \forall v \in V$. Тогда $u_h \in V_h$ называется приближенным решением Ритца симметричной абстрактной вариационной задачи, если

$$a(u_h, v) = f(v) \quad \forall v \in V_h. \quad (2.48)$$

Теорема 2.9. Приближенное решение Ритца симметричной вариационной задачи из определения (2.48) существует и единственно.

Очевидно также, что аналогично (2.22), (2.23) приближенная симметричная вариационная задача формулируется как задача минимизации функционала энергии в конечномерном подпространстве:

$$J(u_h) = \min_{v \in V_h} \{J(v)\}. \quad (2.49)$$

Само приближенное решение u_h при этом можно формально определить как

$$u_h = \arg \left\{ \min_{v \in V_h} \{J(v)\} \right\}.$$

Другая возможная формулировка данных постановок — соотношения (2.48), (2.49) определяют метод Ритца приближенного решения симметричной вариационной задачи минимизации.

Теорема 2.10. Пусть u есть точное решение симметричной вариационной задачи

$$a(u, v) = f(v) \quad \forall v \in V, \quad (2.50)$$

а u_h — определяемое из (2.48) соответствующее приближенное решение. Тогда ошибка приближенного решения $z_h = u - u_h$ ортогональна конечномерному подпространству V_h , т. е.

$$a(z_h, v) = 0 \quad \forall v \in V_h.$$

Поскольку рассматриваемая билинейная форма порождает энергетическое пространство со скалярным произведением и нормой

$$\|v\|_E = (v, v)_E^{1/2} = [a(v, v)]^{1/2}, \quad (2.51)$$

то в силу теоремы 1.7 о минимальном расстоянии, а также непосредственно из (2.45) имеем, что в данной метрике приближенное решение Ритца имеет наименьшую ошибку, т. е.

$$\|z_h\|_E = \min_{v \in V_h} \{\|u - v\|_E\}. \quad (2.52)$$

Теорема 2.11 (о погрешности метода Ритца). Пусть $a(\cdot, \cdot)$ есть симметричная, ограниченная и V -эллиптическая билинейная форма. Тогда для погрешности приближенного решения Ритца вариационной задачи (3.1) справедлива оценка

$$\|z_h\|_V \leq \sqrt{\frac{c_1}{c_2}} \min_{v \in V_h} \{\|u - v\|_V\}, \quad (2.53)$$

где c_1, c_2 — определяемые в (1.5), (1.6) постоянные.

Принципиальное достоинство оценок (2.52), (2.53) в том, что они дают оценки погрешности метода Ритца в разных нормах через соответствующие оценки каких-либо других аппроксимаций точного решения u исходной симметричной вариационной задачи. Здесь одним из естественных подходов является использование интерполанта $u_I(x)$ искомого решения, т. е. следствием оценок (2.52), (2.53) являются неравенства

$$\|z_h\|_E \leq \|u - u_I\|_E, \quad u_I \in V_h, \quad (2.54)$$

$$\|z_h\|_V \leq \sqrt{\frac{c_1}{c_2}} \|u - u_I\|_V, \quad u_I \in V_h. \quad (2.55)$$

Очевидно, что их использование для вывода конкретных оценок требует каких-то предположений о гладкости искомого решения, а также знания свойств интерполяционных приближений в соответствующих пространствах.

Главным принципом нахождения приближенных решений в методе Рунца является построение минимизирующих последовательностей для функционала энергии, см. определение 2.5, на основе конструирования последовательностей конечномерных подпространств.

Определение 2.7. Последовательность конечномерных подпространств $\{V_k^h\}_{k=1}^\infty$ называется полной в H_A , если для любых $v \in H_A$, $\varepsilon > 0$ существует такое $k_\varepsilon > 0$, что при любом $k > k_\varepsilon$ в V_k^h найдется элемент v_k^h , для которого $\|v - v_k^h\|_A < \varepsilon$.

Наиболее распространенные принципы формирования конечномерных подпространств V_k^h основаны на построении последовательности сеточных функций, определенных на сгущающихся сетках с характерными шагами $h_k \rightarrow 0$, при $k \rightarrow \infty$. В частности, с их помощью можно строить сходящиеся интерполяционные приближения, а понятие полноты в данном случае означает, что любой элемент $v \in H_A$ может быть аппроксимирован с любой требуемой точностью сеточной функцией из V_k^h с достаточно малым шагом h_k .

Теорема 2.12. Если последовательность конечномерных подпространств полна в H_A , а оператор A является самосопряженным и положительно определенным в $D(A)$, $\overline{D(A)} = H$, то последовательность приближений u_k^h по методу Рунца

$$J(u_k^h) = \min_{v \in V_k^h} \{J(v)\} \quad (2.56)$$

сходится при $k \rightarrow \infty$ в H_A к обобщенному решению уравнения $Au = f$.

Пусть N_k есть размерность подпространства V_k^h , а совокупность элементов $\varphi_i^{(k)}$, $i = 1, \dots, N_k$, образует его базис, так что для любого $v_k^h \in V_k^h$ имеет место представление

$$v_k^h = \sum_{i=1}^{N_k} c_i^{(k)} \varphi_i^{(k)}. \quad (2.57)$$

Тогда метод Ритца эквивалентен нахождению коэффициентов $c_i^{(k)}$, реализующих минимум энергетического функционала в V_k^h , что приводит к системе уравнений

$$\frac{\partial J(v_k^h)}{\partial c_i^{(k)}} = 0, \quad i = 1, \dots, N_k. \quad (2.58)$$

Далее, опуская для краткости индекс « k », получаем

$$\begin{aligned} \frac{\partial}{\partial c_i} \left[\left(\sum_{j=1}^N c_j \varphi_j, \sum_{j=1}^N c_j \varphi_j \right)_A - 2 \left(\sum_{j=1}^N c_j \varphi_j, f \right) \right] \\ = 2 \left[\sum_{j=1}^N c_j (\varphi_i, \varphi_j)_A - (f, \varphi_i) \right] = 0, \quad i = 1, \dots, N. \end{aligned}$$

Поскольку отсюда видно, что вторая производная функционала положительна, то его минимум обеспечивается системой уравнений

$$\sum_{j=1}^N c_j (\varphi_i, \varphi_j)_A = (f, \varphi_i), \quad i = 1, \dots, N, \quad (2.59)$$

которую можно записать в матричной форме

$$\begin{aligned} A^h c^h &= f^h, \quad A^h = \{a_{ij} = (\varphi_i, \varphi_j)_A\}, \\ f^h &= \{f_i = (f, \varphi_i)\}. \end{aligned} \quad (2.60)$$

Очевидно, что матрица A^h системы уравнений метода Ритца симметрична и положительно определена, что непосредственно следует из самосопряженности и положительной определенности оператора A . Действительно, для любого N -мерного вектора $c^h = \{c_j\}$ имеем

$$(A^h c^h, c^h) \equiv \sum_{i,j=1}^N a_{ij} c_i c_j = \left(\sum_{i=1}^N c_i \varphi_i, \sum_{i=1}^N c_i \varphi_i \right)_A = \|v\|_A^2,$$

где элемент $v = \sum_{i=1}^N c_i \varphi_i \in H$, если только все c_i не равны нулю, и отличен от нуля в силу линейной независимости φ_i .

Таким образом, мы можем сформулировать следующее утверждение.

Теорема 2.13. Пусть оператор A самосопряжен и положительно определен в $D(A)$, $\overline{D(A)} = H$, а элементы $\varphi_i \in V^h \subset H$ составляют базис конечномерного подпространства. Тогда алгебраическая система уравнений (2.60) для коэффициентов s_i приближенного решения метода Рунца имеет единственное решение.

Отметим также, что квадратная матрица, составленная из скалярных произведений (φ_i, φ_j) , $i, j = 1, \dots, N$, элементов гильбертового пространства $\varphi_i \in V$, $i = 1, \dots, N$, называется *матрицей Грама* и обладает следующими общими свойствами: она всегда симметрична и положительно полуопределена (следовательно, имеет неотрицательные собственные числа), а если φ_i — линейно независимые элементы, то матрица Грама положительно определена (собственные числа больше нуля).

2.4.2. Метод Галеркина

Определение 2.8. Пусть V есть гильбертовое пространство и $a(\cdot, \cdot)$ — билинейная непрерывная коэрцитивная (V -эллиптическая) форма, не обязательно симметричная. Пусть также $V_h \subset V$ есть конечномерное пространство и $f(v) \in V'$ $\forall v \in V$ — ограниченная (непрерывная) линейная форма. Тогда $u_h \in V_h$ называется приближенным решением Галеркина абстрактной вариационной задачи (2.10), если

$$a(u_h, v) = f(v) \quad \forall v \in V_h. \quad (2.61)$$

Формула (2.61) может быть также названа *методом Галеркина* решения абстрактной вариационной задачи. Очевидно, в силу формального совпадения (2.49) и (2.61), что метод Галеркина является прямым обобщением метода Рунца в случае несимметричной билинейной формы $a(\cdot, \cdot)$. И именно отсутствие симметричности лишает вариационную задачу минимизационных свойств, что требует привлечения других алгоритмов и способов их исследования.

Предваряя вопросы теоретического обоснования метода Галеркина, рассмотрим формальную алгебраическую запись получаемой системы линейных уравнений. Предположим, что решение u_h задачи (2.61) существует и в разложении по базисным функциям J_i пространства V_h размерности N имеет аналогичное (2.57) представление:

$$u_h = \sum_{i=1}^N c_i \varphi_i.$$

Поскольку любой элемент $v \in V_h$ также может быть разложен по этому базису, то соотношения (2.61) эквивалентны следующей системе линейных уравнений:

$$a\left(\sum_{j=1}^N c_j \varphi_j, \varphi_i\right) = (f, \varphi_i), \quad f \in V', \quad i = 1, \dots, N,$$

которая в векторно-матричной форме имеет вид

$$A^h c^h = f^h, \quad A^h = \{a_{i,j} = a(\varphi_i, \varphi_j)\}, \quad f^h = \{f_i = (f, \varphi_i)\},$$

совпадающий с (2.60) в случае симметричности билинейной формы $a(\cdot, \cdot)$.

Теорема 2.14. *Приближенное решение Галеркина абстрактной вариационной задачи, сформулированное в определении 3.3, существует и единственно.*

Пусть теперь u и u_h — точное решение вариационной задачи (2.10) и приближенное решение Галеркина соответственно. Вопрос об оценке ошибки приближения $z_h = u - u_h$ в определенной степени закрывается следующим образом.

Теорема 2.15 (Сеа). *При выполнении теоремы Вишика–Лакса–Мильграма справедливо неравенство*

$$\|u - u_h\|_V \leq \frac{c_1}{c_2} \min_{v \in V_h} \{\|u - v\|_V\}. \quad (2.62)$$

Отметим, что теорема Сеа устанавливает квазиоптимальность приближенного метода Галеркина в том смысле, что ошибка $\|u - u_h\|_V$ пропорциональна той наименьшей погрешности, которая может быть достигнута в подпространстве V_h .

Напомним, что в симметричной вариационной задаче для приближенного метода Ритца было доказано, что

$$\|u - u_h\|_A = \min_{v \in V_h} \{\|u - v\|_A\},$$

на основе чего в норме $\|\cdot\|_V$ была получена оценка ошибки z_h (2.53), несколько лучшая, чем (2.62).

Как видно из приведенных неравенств, ошибка приближенного решения вариационной задачи может быть оценена через погрешности аппроксимации (например, интерполяции) искомых решений в конечномерных пространствах. При этом необходимо использовать априорную или апостериорную информацию о свойствах неизвестного решения. Более привлекательно выражать итоговую оценку через нормы исходных данных. В этом смысле полезным оказывается следующий результат, устанавливаемый в предположении справедливости неравенства:

$$[a(u - u_h, u - u_h)]^{1/2} \leq \varepsilon \|f\|_V, \quad (2.63)$$

где используется обозначение

$$\|f\|_V = \|f(v)\|_{V'} = \sup_{v \neq 0} \left\{ \frac{|f(v)|}{\|v\|} \right\}, \quad v \in V,$$

и которое во многих приложениях можно установить с одновременной оценкой малости величины ε . При этом точную и приближенную вариационные задачи также будем записывать в виде

$$a(u, v) = (f, v) \quad \forall v \in V, \quad a(u_h, v) = (f, v) \quad \forall v \in V_h \subset V, \quad (2.64)$$

где $f \in V$ — определяемый из теоремы Рисса элемент, который можно назвать «правой частью» соответствующего операторного уравнения.

Наряду с «основными» вариационными задачами (2.64) введем сопряженные им вариационные задачи:

- точная сопряженная — найти элемент $u^* \in V$ такой, что справедливо соотношение

$$a(v, u^*) = (f, v) \quad \forall v \in V; \quad (2.65)$$

- приближенная сопряженная — найти элемент $u_h^* \in V_h \subset V$ такой, что выполняется равенство

$$a(v, u_h^*) = (f, v) \quad \forall v \in V_h. \quad (2.66)$$

Пример. Пусть вариационная задача заключается в нахождении $u \in H_0^1(a, b) \forall v \in H_0^1(a, b)$, причем билинейная и линейная формы определяются как

$$\begin{aligned} a(u, v) &= \int_a^b (u'v' + bu'v + cuv)dx, \\ (f, v) &= \int_a^b fv dx, \quad u, v \in H_0^1(a, b). \end{aligned}$$

Тогда с помощью равенства

$$\int_a^b bu'v dx = - \int_a^b u(bv)' dx \quad \forall u, v \in H_0^1(a, b)$$

определяется сопряженная билинейная форма

$$a^*(u, v) = \int_a^b (v'u' - (bv) + cvu)dx, \quad u, v \in H_0^1(a, b),$$

обладающая свойством $a^*(u, v) = a(v, u)$. При этом вариационная задача (2.65) порождает сопряженную дифференциальную краевую задачу

$$\begin{aligned} L^*u^* &\equiv -(u^*)'' - (bu^*)' + cu^* = 0, \\ x &\in (a, b), \quad u^*(a) = u^*(b) = 0. \end{aligned} \quad (2.67)$$

Теорема 2.16 (Обэна–Ницше). Пусть выполняются условия теоремы Сеа и неравенство (2.63). Тогда справедлива оценка

$$\|u - u_h\|_V \leq \varepsilon^2 \|f\|_V. \quad (2.68)$$

Отметим, что в случае симметричности и коэрцитивности билинейной формы левая часть неравенства (2.63) есть энергетическая норма ошибки $\|u - u^h\|_E$. Таким образом, теорема 2.16 устанавливает, при наличии оценки в энергетической норме, более высокую скорость сходимости, в смысле порядка малой величины ε , в норме гильбертового пространства V . Подчеркнем также тот немаловажный факт, что доказанная оценка (2.68) справедлива и для несимметричной вариационной задачи.

2.4.3. «Несогласованный» метод Галеркина

Общим для рассмотренных в пп.2.4.1 и 2.4.2 методов Ритца и Галеркина являлось то, что пространство приближенных решений V_h было вложено в пространство точных решений V вариационной задачи. Однако условие $V_h \subset V$ в ряде приближенных алгоритмов, представляющих практический интерес, может и не выполняться. Оказывается при этом, что если пространство V_h «почти вложено» в V в некотором смысле, то получаемый приближенный метод является оправданным и сходящимся. Рассмотрим такое «несогласование» пространств V_h и V на примере абстрактной формулировки более общего метода Галеркина. Справедливость получаемых ниже результатов для метода Ритца в случае симметричной вариационной постановки очевидна.

Пусть H обозначает гильбертово пространство и $V \subset H$ — подпространство. Пусть также $V_h \not\subset V$ есть конечномерное пространство, в котором будем искать приближенное решение вариационной задачи, порождаемой билинейной формой $a(\cdot, \cdot)$, которая определена в H и не является симметричной.

Теорема 2.17. Пусть $V \subset H$ и $V_h \subset H$, а $a(\cdot, \cdot)$ есть билинейная форма, непрерывная на $V + V_h$, коэрцитивная как в V , так и в V_h . Обозначим через c_1 и c_2 константы непрерывности и коэрцитивности, а через $u \in V$ — решение вариационной задачи

$$a(u, v) = f(v) \quad \forall v \in V, \quad f \in V'.$$

Тогда для решения $u_h \in V_h$ задачи

$$a(u_h, v) = f(v) \quad \forall v \in V_h$$

справедливо неравенство

$$\|u - u_h\|_H \leq \left(1 + \frac{c_1}{c_2}\right) \inf_{v \in V_h} \{\|u - v\|_H\} + \frac{1}{c_2} \sup_{w \in V_h \setminus \{0\}} \left\{ \frac{a(u - u_h, w)}{\|w\|_H} \right\}. \quad (2.69)$$

Отметим, что последнее слагаемое в правой части (2.69) характеризует вклад «несогласованности» пространств V_h , V и обращается в нуль, если $V_h \subset V$.

Типичным примером нарушения вариационного принципа, заключающегося в отсутствии вложенности V_h в V , является применение простейших аппроксимаций при решении краевых задач в расчетных областях с криволинейной границей.

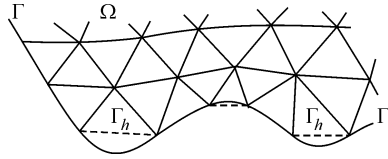


Рис. 2.1. Иллюстрация «спрямления» криволинейной границы двумерной области

Пусть в двумерной области Ω с границей Γ решается однородная задача Дирихле, т.е. $u|_{\Gamma}$, а приближенное решение u_h строится в пространстве V_h кусочно-гладких функций, которые являются линейными на каждом из треугольников, составляющих сеточную расчетную

область Ω_h с границей Γ_h , изображенную на рис. 3.1 штриховой ломаной и аппроксимирующей криволинейную границу Γ .

Простейшей и естественной аппроксимацией краевого условия является $u_h|_{\Gamma_h} = 0$, в результате чего получаем

$$u_h \in V_h \not\subset V.$$

Другой возможный способ аппроксимации краевой задачи с криволинейной границей — оставлять приграничные треугольники с одной криволинейной стороной, используя при этом какие-то приближенные квадратурные формулы для вычисления билинейной формы $a(\cdot, \cdot)$, фактически модифицируя ее к некоторому другому виду $a_h(\cdot, \cdot)$. Это будет другой тип нарушения вариационного принципа, рассмотрение которого требует отдельной теоремы, которую мы приведем для случая симметричной задачи.

Теорема 2.18. Пусть $a_h(\cdot, \cdot)$ будет непрерывная невырожденная (положительно определенная) симметричная билинейная форма на $V + V_h$, которая приводится к $a(\cdot, \cdot)$ на V . Пусть $u \in V$ есть решение вариационной задачи

$$a(u, v) = f(v) \quad \forall v \in V,$$

где $f \in V' \cap V'_h$, и $u_h \in V_h$ — решение приближенной задачи

$$a_h(u_h, v) = f(v) \quad \forall v \in V_h.$$

Тогда справедливо неравенство

$$\|u - u_h\|_h \leq \inf_{v \in V_h} \{\|u - v\|\} + \sup_{w \in V_h \setminus \{0\}} \left\{ \frac{|a_h(u - u_h, w)|}{\|w\|_h} \right\}, \quad (2.70)$$

в котором используется обозначение $\|\cdot\|_h = \sqrt{a_h(\cdot, \cdot)}$.

Как и в предыдущей теореме, последнее слагаемое отвечает за факт $V_h \not\subset V$ и обращается в нуль при $V_h \subset V$.

Отметим, что аналогичную теорему можно доказать также для несимметричного случая. И наоборот, предыдущую теорему можно было бы несколько усилить, переформулировав ее для симметричной билинейной формы.

Рассмотрим еще один пример рассогласования пространств V_h и V , обусловленного применением так называемых неконформных кусочно-линейных конечных элементов. В этом случае расчетная полигональная область разбивается на треугольники $\Omega = \cup_k T_k$, а пространство приближенных решений V_h определяется как совокупность функций, линейных в каждом треугольнике T_k и непрерывных в точках — серединах общих сторон смежных треугольников. Если решается краевая задача для уравнения Пуассона с однородным условием Дирихле, то искомое решение ищется в пространстве $V = H^1(\Omega)$, а функции $v \in V_h$ полагаются равными нулю в серединах сторон треугольников, лежащих на границе области. Так как при этом функции в V_h не являются непрерывными, то модифицированная билинейная форма и соответствующая норма вводятся следующим естественным образом:

$$a_h(v, w) = \sum_k \int_{T_k} a(v, w) d\vec{x} = \sum_k \int_{T_k} \nabla v \cdot \nabla w d\vec{x},$$

$$\|v\|_h = \sqrt{a_h(v, v)}.$$

Форма $a_h(\cdot, \cdot)$ коэрцитивна, так как она равна $a_h(\cdot, \cdot)$ в V , и является невырожденной, поскольку равенство $a_h(v, v) = 0$ означает кусочно-постоянный характер функции v , а вследствие нулевых граничных условий и непрерывности в серединах сторон треугольников получаем $v \equiv 0$.

Обоснование таких разрывных аппроксимаций можно провести с помощью теоремы 2.18, см. подробнее [26] и цитируемую там литературу.

2.4.4. Смешанные вариационные методы

Под смешанными формулировками краевых задач традиционно понимаются такие, которые включают сведение дифференциальных уравнений к системам уравнений более низкого порядка.

В этом случае при построении численных методов удается находить приближения не только для исходных искомых решений, но и для их производных, или потоков субстанции, определяемой конкретным прикладным контекстом.

Однако в некоторых приложениях, связанных чаще всего с векторными задачами, как раз смешанные постановки являются исходными и естественными с физической точки зрения.

Примеры смешанных формулировок

В качестве иллюстрации можно рассмотреть эллиптическое n -мерное уравнение

$$-\sum_{i,j=1}^n \frac{\partial}{\partial x_i} \left(a_{i,j}(\vec{x}) \frac{\partial w}{\partial x_j} \right) = f(\vec{x}), \quad \vec{x} \in \Omega. \quad (2.71)$$

Будем предполагать, что матрица коэффициентов $A = \{a_{i,j}\}$ невырождена, и обозначим через $\bar{a}_{i,j}$ элементы обратной к ней матрицы $A^{-1} = \{\bar{a}_{i,j}\}$.

Вводя обозначения

$$\sum_{j=1}^n a_{i,j} \frac{\partial w}{\partial x_j} = u_i(\vec{x}), \quad \vec{u} = (u_1, \dots, u_n)^t,$$

можно определить вектор градиента решения

$$\text{grad } w = A^{-1} \vec{u}$$

и записать вариационную формулировку исходной задачи в следующем виде:

$$\begin{aligned} a(\bar{u}, \bar{v}) + b(\bar{v}, w) &= 0 \quad \forall \bar{v} \in V, \\ b(\bar{u}, q) &= f(q) \quad \forall q \in Q. \end{aligned} \quad (2.72)$$

Здесь используются обозначения для билинейных форм

$$\begin{aligned}
a(\bar{u}, \bar{v}) &= \int_{\Omega} \sum_{i=1}^n (\text{grad } u)_i (\text{grad } v)_i d\vec{x} = \int_{\Omega} \sum_{i,j=1}^n \bar{a}_{i,j} u_i v_j d\vec{x}, \\
b(u, q) &= - \int_{\Omega} (\text{div } \bar{v}) q d\vec{x},
\end{aligned} \tag{2.73}$$

$$Q = \left\{ q \in L^2(\Omega) : \int_{\Omega} q d\vec{x} = 0 \right\}$$

есть пространство пробных скалярных функций, а пространство векторных функций V определяется как

$$V = \{ \bar{v} = (v_1, \dots, v_n) \in L_2(\Omega)^n : \text{div } \bar{v} \in L_2(\Omega) \}.$$

Это пространство обозначается также через $H(\div)$ и снабжается нормой

$$\|\bar{v}\|_{H(\text{div})}^2 = \|\bar{v}\|_{L_2(\Omega)^n}^2 + \|\text{div } \bar{v}\|_{L_2(\Omega)}^2. \tag{2.74}$$

Отметим, что билинейная форма $a(\cdot, \cdot)$ не коэрцитивна во всем пространстве V , но она оказывается коэрцитивной в подпространстве функций с нулевой дивергенцией.

Второй характерной задачей, порождающей естественным образом смешанную вариационную постановку, является система уравнений Стокса

$$\begin{aligned}
-\Delta \bar{u} + \text{grad } p &= \bar{f}(\vec{x}), \\
\text{div } \bar{u} &= 0, \quad \vec{x} \in \Omega \subset R^n,
\end{aligned} \tag{2.75}$$

которая формально описывает стационарное векторное поле скоростей $\bar{u}$ и распределение давления p бесконечно вязкой жидкости под действием объемной силы $\bar{f}$. На границе Γ расчетной области Ω для простоты выбираем однородное условие Дирихле $\bar{u} = 0$.

После умножения уравнений системы (3.29) на $\bar{v} \in V$ и $q \in Q$ соответственно с помощью интегрирования по частям получаем систему вариационных равенств

$$\begin{aligned}
a(\bar{u}, \bar{v}) + b(\bar{v}, p) &= \int_{\Omega} (\bar{f}, \bar{v}) d\vec{x} \quad \forall \bar{v} \in V, \\
b(\bar{u}, q) &= 0 \quad \forall q \in Q,
\end{aligned}$$

где $a(\cdot, \cdot)$ и $b(\cdot, \cdot)$ суть введенные в (2.73) билинейные формы.

Таким образом, в рассмотренных примерах мы определили два гильбертовых пространства (V — пространство векторных функций и Q — пространство скалярных функций) и две билинейные формы

$$a(\cdot, \cdot) : V \times V \rightarrow R, \quad b(\cdot, \cdot) : V \times Q \rightarrow R,$$

которые будем считать непрерывными:

$$\begin{aligned} a(\bar{u}, \bar{v}) &\leq C \|\bar{u}\|_V \cdot \|\bar{v}\|_V \quad \forall \bar{u}, \bar{v} \in V, \\ b(\bar{v}, p) &\leq C \|\bar{v}\|_V \cdot \|p\|_Q \quad \forall \bar{v}, p \in Q, \end{aligned} \quad (2.76)$$

Сформулируем абстрактную смешанную вариационную задачу как нахождение функций $\bar{u} \in V$ и $p \in Q$, удовлетворяющих равенствам

$$\begin{aligned} a(\bar{u}, \bar{v}) + b(\bar{u}, p) &= f(v) \quad \forall \bar{v} \in V, \\ b(\bar{u}, q) &= g(q) \quad \forall q \in Q, \end{aligned} \quad (2.77)$$

где $f \in V'$ и $g \in Q'$.

Определим замкнутое подпространство $W \subset V$:

$$W = \{\bar{v} \in V : b(\bar{v}, q) = 0 \quad \forall q \in Q\}.$$

Если в (2.77) $g = 0$, то эта вариационная формулировка эквивалентна задаче о нахождении функции $\bar{u} \in W$ такой, что

$$a(\bar{u}, \bar{v}) = f(v) \quad \forall v \in W. \quad (2.78)$$

Эта задача корректна, в силу теоремы Вишика–Лакса–Мильграма, если билинейная форма коэрцитивна, т. е.

$$a(\bar{v}, \bar{v}) \geq \alpha \|\bar{v}\|_V^2, \quad \alpha > 0, \quad \forall \bar{v} \in W. \quad (2.79)$$

В приведенном выше примере (2.71)–(2.73) для скалярной эллиптической задачи билинейная форма $a(\cdot, \cdot)$ является коэрцитивной в подпространстве функций с нулевой дивергенцией:

$$W = \{\bar{v} \in H(\text{div}) : \text{div } \bar{v} = 0\}. \quad (2.80)$$

Если же в (2.77) $g \neq 0$, можно эту задачу свести к случаю $g = 0$ следующим образом. Пусть $\bar{u}_0 \in V$ есть любое решение уравнения $b(\bar{u}_0, q) = g(q) \quad \forall q \in Q$. Если также функция $\bar{u}_0$ каким-либо образом найдена, то решение задачи (2.77) имеет вид $\bar{u} = \bar{u}_1 + \bar{u}_0$, где $\bar{u}_1$ удовлетворяет вариационному уравнению

$$a(\bar{u}, v) = f(\bar{v}) - a(\bar{u}, \bar{v}) \quad \forall v \in W,$$

и мы снова приходим к задаче типа (2.78):

$$\begin{aligned} a(\bar{u}, \bar{v}) + b(\bar{v}, p) &= f(\bar{v}) \quad \forall \bar{v} \in V, \\ b(\bar{u}, q) &= 0 \quad \forall q \in Q. \end{aligned} \quad (2.81)$$

Если функция $\bar{u}$ из (2.81) уже найдена, мы можем определить $p \in Q$ как решение уравнения

$$b(\bar{v}, p) = f(\bar{v}) - a(\bar{u}, \bar{v}) \quad \forall v \in V. \quad (2.82)$$

Чтобы установить корректность этой задачи, необходимо ввести новый тип коэрцитивности. Для этого полезно заметить, что из неравенства (2.79) следует соотношение

$$\sup_{\bar{w} \in V} \left\{ \frac{a(\bar{w}, \bar{v})}{\|\bar{w}\|_V} \right\} \geq \alpha \|v\|_V.$$

Применяя это соображение к $b(\cdot, \cdot)$, мы определяем условие

$$\sup_{w \in V} \left\{ \frac{b(\bar{w}, p)}{\|\bar{w}\|_V} \right\} \geq \beta \|p\|_Q, \quad (2.83)$$

которое позволяет сформулировать следующий результат (см. доказательство в [26]).

Теорема 2.19. *Если билинейная форма $b(\cdot, \cdot)$ удовлетворяет условиям непрерывности и коэрцитивности (2.76), (2.83), то вариационное уравнение*

$$b(\bar{v}, p) = f(\bar{v}) \quad \forall \bar{v} \in V \quad (2.84)$$

имеет единственное решение.

Дискретная смешанная вариационная задача

Введем конечномерные пространства $V_h \subset V$ и $Q_h \subset Q$, для которых рассмотрим приближенную вариационную постановку о нахождении функций $\bar{u}_h \in V_h$ и $p_h \in Q_h$ таких, что

$$\begin{aligned} a(\bar{u}_h, \bar{v}) + b(\bar{v}, p_h) &= f(\bar{v}) \quad \forall \bar{v} \in V_h, \\ b(\bar{u}_h, q) &= 0 \quad \forall q \in Q_h. \end{aligned} \quad (2.85)$$

Аналогично предыдущему определим пространство

$$W_h = \{\bar{v} \in V_h : b(\bar{v}, q) = 0 \quad \forall q \in Q_h\}. \quad (2.86)$$

Тогда задача (2.85) может быть переформулирована следующим образом: найти решение $\bar{u}_h \in W_h$ из вариационного равенства

$$a(\bar{u}_h, \bar{v}) = f(\bar{v}) \quad \forall \bar{v} \in W_h, \quad (2.87)$$

а затем определить $p_h \in Q_h$ такое, что выполняется соотношение

$$b(\bar{v}, p_h) = f(\bar{v}) - a(u_h, \bar{v}) \quad \forall \bar{v} \in W_h. \quad (2.88)$$

Если $W_h \subset W$, с помощью теоремы Сеа получается неравенство

$$\|\bar{u} - \bar{u}_h\|_V \leq \frac{C}{\alpha} \inf_{\bar{v} \in W_h} \{\|\bar{u} - \bar{v}\|_V\}. \quad (2.89)$$

В случае $W_h \not\subset W$ происходит нарушение вариационного принципа, и после применения рассмотренных в п. 2.4.3 положений приходим к оценке

$$\begin{aligned} \|\bar{u} - \bar{u}_h\|_V &\leq \left(1 + \frac{C}{\alpha}\right) \inf_{\bar{v} \in W_h} \{\|\bar{u} - \bar{v}\|_V\} + \\ &\quad \frac{1}{\alpha} \sup_{\bar{w} \in W_h \setminus \{0\}} \left\{ \frac{|a(u - u_h, \bar{w})|}{\|\bar{w}\|_V} \right\}. \end{aligned} \quad (2.90)$$

Последний член в (2.90) может быть преобразован с помощью следующих справедливых для $w \in W_h$ соотношений:

$$\begin{aligned} a(\bar{u} - \bar{u}_h, \bar{w}) &= a(\bar{u}, \bar{w}) - f(\bar{w}) \\ &= -b(\bar{w}, p) = -b(\bar{w}, p - q) \quad \forall q \in Q_h. \end{aligned}$$

Отсюда в силу условий непрерывности (2.76) имеем неравенство

$$|b(\bar{w}, p - q)| \leq C \|\bar{w}\|_V \cdot \|p - q\|_Q,$$

а поскольку оно справедливо для любого $q \in Q_h$, то мы получаем соотношение

$$|a(\bar{u} - \bar{u}_h, \bar{w})| \leq C \|\bar{w}\|_V \inf_{q \in Q_h} \{\|p - q\|_Q\},$$

которое позволяет сформулировать следующее утверждение.

Теорема 2.20. Пусть $V_h \subset V$ и $Q_h \subset Q$, а пространства W и W_h определены в соответствии с (2.80) и (2.86). Пусть также выполнены условия (2.76) и (2.79) для всех $\bar{v} \in W \cup W_h$. Тогда для определенных в (2.81) функций $\bar{u}$, p и функций $\bar{u}_h$ из (2.91) справедливо неравенство

$$\|\bar{u} - \bar{u}_h\|_V \leq \left(1 + \frac{C}{\alpha}\right) \inf_{\bar{v} \in W_h} \{\|\bar{u} - \bar{v}\|_V\} + \frac{C}{\alpha} \inf_{q \in Q_h} \{\|p - q\|_Q\}. \quad (2.91)$$

Главный результат данной теоремы заключается в том, что ошибка приближенного решения $\bar{u} - \bar{u}_h$ определяется, при выполнении указанных свойств непрерывности и коэрцитивности, аппроксимационными свойствами пространств W_h и Q_h .

Перейдем теперь к анализу ошибки скалярной функции $p - p_h$, где p и p_h определяются из вариационных равенств (2.82) и (2.88) соответственно. Для этого предварительно сформулируем так называемое inf-sup, или ЛББ (Ладыженская–Бабушка–Брецци), условие для билинейной формы $b(\cdot, \cdot)$:

$$\inf_{q \in Q_h} \sup_{\bar{v} \in V_h} \left\{ \frac{b(\bar{v}, q)}{\|\bar{v}\|_V \cdot \|q\|_Q} \right\} = \beta > 0. \quad (2.92)$$

Это условие играет роль коэрцитивности для формы $b(\cdot, \cdot)$ и является естественным для ряда характерных задач математической физики. Через константу β устанавливается следующий результат, доказательство которого можно найти, например, в [26].

Теорема 2.21. Пусть $V_h \subset V$ и $Q_h \subset Q$, а функции u , p и u_h , p_h определим из соотношений (2.78), (2.87), (2.82) и (2.88). Тогда при выполнении условий (2.76) и (2.92) вариационное уравнение (2.88) имеет единственное решение p_h , а также справедливо неравенство

$$\|p - p_h\|_Q \leq \frac{C}{\beta} \left[\|\bar{u} - \bar{u}_h\|_V + \left(1 + \frac{C}{\beta}\right) \inf_{q \in Q_h} \{\|p - q\|_Q\} \right]. \quad (2.93)$$

Если теперь в соотношении (2.93) использовать оценку для $\bar{u} - \bar{u}_h$ из теоремы 3.12, то получаем итоговый результат:

$$\|p - p_h\|_Q \leq \left(1 + \frac{C}{\beta}\right) \left(1 + \frac{C}{\alpha}\right) \left[\inf_{\bar{v} \in W_h} \{\|u - \bar{v}\|_V\} + \inf_{q \in Q_h} \{\|p - q\|_Q\} \right].$$

Глава 3

Конечно-разностные аппроксимации

Как известно, основатели дифференциального и интегрального исчисления И. Ньютон и Г. Лейбниц исходили из понятия бесконечно малых. И поэтому вполне естественно, что отправной точкой приближенных решений дифференциальных уравнений, прародителем которых явился явный метод Л. Эйлера (1768 г.), стали конечные достаточно малые величины и конечно-разностные аппроксимации производных. Функциональный анализ за века своего развития стал мощнейшим орудием изучения реальных процессов и явлений, однако переход от непрерывных к дискретным моделям окружающей действительности явился неизбежным не только из методических соображений, но также из понимания идеализированности таких математических объектов, как производные, при описании атомарного микромира. Появились даже философско-методологические концепции об изначальной роли именно конечно-разностных, а не дифференциальных уравнений. Не случайно дискретная математика стала одной из важнейших научных дисциплин, а также появилось такое самостоятельное направление, как исчисление конечных разностей, которому посвящена, например, фундаментальная монография А. О. Гельфонда [43] 1967 г.

Однако мы в данной главе придерживаемся прагматического представления о конечно-разностных методах как об одном из главных подходов к решению дифференциальных уравнений. Поэтому основой рассматриваемых вычислительных алгоритмов является построение сеточной модели исходной непрерывной задачи. Генерация сеток для решения сложных практических задач требует своих вычислительных методов, и этим вопросам посвящена отдельная глава нашей книги. Здесь же мы будем рассматривать только относительно простые дискретизации, поскольку эффективное использование адаптивных и неструктурированных сеток — это одно из главных предназначений методов конечных объемов и конечных элементов, которым посвящены следующие главы.

Активное развитие конечно-разностных методов пришлось на вторую половину XX-го века и нашло отражение в книгах

русских математиков Г. И. Марчука [121, 123], А. А. Самарского с коллегами [156]–[162], С. К. Годунова и В. Е. Рябенского [49], Е. Г. Дьяконова [68, 69], Н. Н. Яненко [206] а также зарубежных авторов В. Вазова и Дж. Форсайта [35], Э. Дулана и Дж. Миллера [66], Р. Рихтмайра и К. Мортонна [150], Э. Хайрера [192]–[194] и многих других. Среди более поздних изданий можно указать книги Н. С. Бахвалова с коллегами [17], П. Н. Вабищевича [30]–[34], А. И. Задорина [75], автора [81], В. И. Лебедева [105], К. М. Магометова и А. С. Холодова [115], Е. А. Новикова [134], В. В. Смелова [167], А. И. Толстых [178], В. В. Шайдурова с учениками [11], [12], Г. И. Шишкина [201], Ю. И. Шокина [205], в которых имеется достаточно содержательный анализ публикаций своего времени. По этим вопросам существует огромное количество журнальной литературы, не нашедшей еще своего монографического изложения, но на их обзоре мы не останавливаемся, а только иногда делаем ссылки в самих описаниях оригинальных результатов.

Пусть Ω – область n -мерного пространства точек x с декартовыми ортогональными координатами $x_1, \dots, x_n$, $n \geq 2$, а $u(x)$ – определенная в этой области действительная функция, непрерывная вместе со своими производными. Мы будем рассматривать уравнения в частных производных второго порядка

$$\sum_{p,q=1}^n a_{pq} \frac{\partial^2 u}{\partial x_p \partial x_q} + \sum_{p=1}^n b_p \frac{\partial u}{\partial x_p} + cu = f(x),$$

где коэффициенты a_{pq} , b_p , c и правая часть f – заданные в Ω вещественные функции от x . Уравнение называется *эллиптическим в области Ω* , если характеристическая квадратичная форма

$$Q = \sum_{p,q=1}^n a_{pq}(x) \xi_p \xi_q$$

для любых вещественных параметров $\xi_1, \dots, \xi_n$ и $x \in \Omega$ строго положительна.

Главными объектами наших исследований являются двумерные задачи с $n = 2$. Основные результаты теории разностных методов переносятся и на более громоздкий для изложения трехмерный случай. Одномерные задачи формально относятся к обыкновенным дифференциальным уравнениям, но мы их использу-

ем не только из удобства в методологическом плане, но и в силу их самостоятельного практического интереса.

Область Ω предполагаем ограниченной, а ее границу Γ – состоящей из частей $\Gamma_1, \Gamma_2, \Gamma_3$, на каждой из которых задано условие или первого рода (Дирихле) $u = g_1(x)$, или второго (Неймана) $\frac{\partial u}{\partial \vec{n}} = g_2(x)$, или третьего (Ньютона) $\frac{\partial u}{\partial \vec{n}} + \alpha u = g_3(x)$; здесь $\frac{\partial}{\partial \vec{n}}$ означает производную по направлению внешней нормали к границе области. Отметим еще один характерный класс задач – нахождение решений, периодических по какой-либо одной переменной. В таком случае краевое условие *периодичности* имеет вид, например, $u(x_1 + H) = u(x_1)$, а встречается оно при наличии расчетных областей с периодически меняющимися геометрическими и материальными свойствами (H – длина периода).

При исследовании конечно-разностных аппроксимаций от свойств краевой задачи нам требуется формально только достаточная гладкость решения, т. е. ограниченность необходимого числа производных. Мы не останавливаемся специально на нелинейных дифференциальных уравнений, так как для них справедливы все рассматриваемые аппроксимационные качества МКР, поскольку они определяются только свойствами частных производных по независимым переменным.

Фактически единственный инструмент функционального анализа, необходимый нам в гл. 3, – это ряд Тейлора, который в случае нескольких независимых переменных можно записать в виде

$$\begin{aligned} u(x_1 + h_1, \dots, x_n + h_n) &= u(x_1, \dots, x_n) + \sum_{p=1}^n h_p \frac{\partial}{\partial x_p} u(x_1, \dots, x_n) \\ &+ \frac{1}{2} \left(\sum_{p=1}^n h_p \frac{\partial}{\partial x_p} \right)^2 u(x_1, \dots, x_n) + \dots + \frac{1}{k!} \left(\sum_{p=1}^n h_p \frac{\partial}{\partial x_p} \right)^k u(x_1, \dots, x_n) \\ &+ \frac{1}{(k+1)!} \left(\sum_{p=1}^n h_p \frac{\partial}{\partial x_p} \right)^{k+1} u(\xi_1, \dots, \xi_n). \end{aligned} \quad (3.1)$$

Здесь h_p – произвольные приращения соответствующих аргументов, $\xi_p \in [x_p, x_p + h_p]$ и, естественно, предполагается, что функция $u(x)$ обладает ограниченными производными до $(k+1)$ -го порядка включительно. Подчеркнем, что при любом конечном значении k равенство (3.1) является не приближенным, а

точным, в котором, однако, неизвестны величины ξ_p . Последний член (будем обозначать его через ψ) называется погрешностью отрезка ряда Тейлора. Обозначая далее через h максимальный модуль приращения (шага) аргумента, а через M_{k+1} – максимальный модуль смешанной производной $(k+1)$ -го порядка в окрестности $[x, x+h]$, при $h \ll 1$ можно записать

$$|\psi| \leq \frac{n^{k+1}}{(k+1)!} h^{k+1} M_{k+1} = O(h^{k+1}), \quad (3.2)$$

где величина $k+1$ называется *порядком погрешности*. В дальнейшем вместо x_i для независимых переменных будем использовать обозначения x, y, z .

3.1. Приближения производных и дифференциальных выражений

Пусть $u(x)$ – функция одной переменной, определенная на некотором отрезке $\bar{\Omega} = \Omega \cup \Gamma = \{x \in [a, b]\}$. Определим в Ω сетку как совокупность точек узлов x_i с шагами $h_i = x_{i+1} - x_i > 0$ (в дальнейшем через h будем обозначать, по контексту, или шаг равномерной сетки, или максимальный из шагов h_i , а через Ω_h – множество узлов, составляющих сетку).

3.1.1. Разностные аппроксимации производной

Введем определения *конечных разностей первого порядка*:

$$\begin{aligned} \nabla_h^+ u_i &\equiv \frac{u(x_{i+1}) - u(x_i)}{h_i} = u'(x_i) + \psi_i^+, \\ \nabla_h^- u_i &\equiv \frac{u(x_i) - u(x_{i-1}))}{h_{i-1}} = u'(x_i) + \psi_i^-. \end{aligned} \quad (3.3)$$

Первая из них называется *правой разностью*, а вторая – *левой* (употребляются еще наименования *разность вперед* и *разность назад*). *Погрешности аппроксимации* ψ_i^+ , ψ_i^- этих разностей, называемых также *односторонними*, легко выводятся из разложения величин $u(x_{i\pm 1})$ в ряд Тейлора относительно точки x_i ($k = p = 1$ в формуле (3.1)):

$$\begin{aligned} \psi_i^+ &= \frac{1}{2} h_i u''(\xi_i^+), \quad \xi_i^+ \in [x_i, x_{i+1}], \\ \psi_i^- &= \frac{-1}{2} h_{i-1} u''(\xi_i^-), \quad \xi_i^- \in [x_{i-1}, x_i]. \end{aligned} \quad (3.4)$$

Отсюда уже можно сформулировать первое утверждение: односторонние разности первого порядка (слово “конечные” для краткости опускаем) аппроксимируют первую производную на функциях из класса $C^2[a, b]$ (имеющих ограниченные первые и вторые производные) с погрешностью первого порядка: $|\psi_i^+| = O(h)$, $|\psi_i^-| = O(h)$.

Нетрудно видеть, что для аппроксимации $u'(x_i)$ можно применять и так называемую *двустороннюю разность первого порядка*

$$\bar{\nabla}_h u_i \equiv \frac{u(x_{i+1}) - u(x_{i-1}))}{h_i + h_{i-1}} = u(x_i) + \bar{\psi}_i. \quad (3.5)$$

Для величины $\bar{\psi}_i$, при использовании еще одного дополнительного члена в ряде Тейлора, легко получить выражение

$$\bar{\psi}_i = \frac{1}{2}(h_i - h_{i-1})u''(x_i) + \frac{1}{6}(h_i^2 - h_i h_{i-1} + h_{i-1}^2)u'''(\xi_i),$$

$$\xi_i \in [x_{i-1}, x_{i+1}].$$

Отсюда следует, что двусторонняя разность аппроксимирует первую производную на неравномерной сетке также с погрешностью $O(h)$. Однако в частном случае $h_i = h_{i-1} = h$ имеем $\bar{\psi}_i = \frac{h^2}{6}u'''(\xi_i)$, а получаемое при этом из (3.5) выражение

$$\bar{\nabla}_h u_i \equiv \frac{u(x_{i+1}) - u(x_{i-1}))}{2h} \quad (3.6)$$

называется *центральной разностью*, и последняя обладает погрешностью $O(h^2)$.

Поскольку для погрешностей $\psi_i^\pm$ из (3.3) при использовании трех членов в отрезке ряда Тейлора можно получить вместо (3.4) новые соотношения

$$\psi_i^+ = \frac{1}{2}h_i u_i'' + \frac{1}{6}h_i^2 u'''(\xi_i^+),$$

$$\psi_i^- = \frac{-1}{2}h_{i-1} u_i'' + \frac{1}{6}h_{i-1}^2 u'''(\xi_i^-)$$

(здесь величины $\xi_i^\pm$ отличаются, строго говоря, от предыдущих, а u_i'' – как и далее u_i – введены для обозначений $u''(x_i)$ и $u(x_i)$), то понятно, что линейная комбинация односторонних разностей дает новую – *трехточечную* в общем случае – *аппроксимацию первой производной*:

$$\begin{aligned}
\nabla_h u_i &= \frac{1}{h_i + h_{i-1}} (h_{i-1} \nabla^+ u_i + h_i \nabla^- u_i) \\
&= \frac{1}{h_i + h_{i-1}} \left[\frac{h_{i-1}}{h_i} u_{i+1} - \left(\frac{h_{i-1}}{h_i} - \frac{h_i}{h_{i-1}} \right) u_i - \frac{h_i}{h_{i-1}} u_{i-1} \right] \\
&= u'_i + \frac{h_i h_{i-1}}{6} u'''(\xi_i), \quad \xi_i \in [x_{i-1}, x_{i+1}], \quad (3.7)
\end{aligned}$$

которая имеет уже погрешность второго порядка, если только функция $u(x)$ обладает ограниченной третьей производной. Если сетка равномерная, то выражение $\nabla_h u_i$ становится двухточечным:

$$\nabla_h u_i \equiv \frac{u_{i+1} - u_{i-1}}{2h} = u'_i + \psi_i, \quad |\psi_i| \leq \frac{h^2}{6} M_3,$$

и совпадает с центральной разностью (3.6) относительно точки x_i (величина M_3 – это максимум модуля третьей производной на интервале $[x_{i-1}, x_{i+1}]$).

Трехточечная аппроксимация первой производной (3.7) может быть выведена также с помощью применения интерполяции и центральных разностей. Для этого нужно воспользоваться выражением для линейной интерполяции

$$u(x) = \frac{(x-a)u(b) + (b-x)u(a)}{b-a} + \psi^1, \quad \psi^1 = \frac{(x-a)(x-b)}{2} u''(\xi), \quad \xi \in [a, b],$$

а также аппроксимациями первых производных:

$$\begin{aligned}
\bar{\nabla}_h u_{i+1/2} &\equiv \frac{u_{i+1} - u_i}{h_i} = u'_{i+1/2} + \bar{\psi}_{i+1/2}, \\
\bar{\psi}_{i+1/2} &= \frac{h_i^2}{24} u'''(\xi_{i+1/2}), \quad \xi_{i+1/2} \in [x_i, x_{i+1}].
\end{aligned}$$

Интерполируя теперь значение u'_i по величинам $u'_{i\pm 1/2}$, т. е. полагая $a = x_{i-1/2}$ и $b = x_{i+1/2}$, с помощью этих формул получаем соотношение

$$\begin{aligned}
\bar{\nabla}_h u_i &= \frac{1}{x_{i+1/2} - x_{i-1/2}} \left[(x_i - x_{i-1/2}) (\bar{\nabla}_h u - \bar{\psi})_{i-1/2} \right. \\
&\quad \left. + (x_i - x_{i+1/2}) (\bar{\nabla}_h u - \bar{\psi})_{i+1/2} \right] + \psi_i^1,
\end{aligned}$$

которое после проведения простых преобразований приводит к (3.7).

3.1.2. Конечно-разностные выражения второго порядка

Через конечные разности первого порядка рекуррентно могут быть определены разности второго и более высокого порядков, аппроксимирующие различные производные. Аналогично предыдущему на неравномерной сетке *разность второго порядка* записывается в виде

$$\begin{aligned}\Delta_h u_i &\equiv \frac{\bar{\nabla}_h u_{i+1/2} - \bar{\nabla}_h u_{i-1/2}}{(h_i + h_{i-1})/2} \\ &= \frac{2u_{i-1}}{h_{i-1}(h_i + h_{i-1})} - \frac{2u_i}{h_i h_{i-1}} + \frac{2u_{i+1}}{h_i(h_i + h_{i-1})} \\ &= u_i'' + \frac{h_i - h_{i-1}}{3} u_i''' + \frac{h_i^2 - h_i h_{i-1} + h_{i-1}^2}{12} u^{(4)}(\xi_i), \\ &\quad \xi_i \in [x_{i-1}, x_{i+1}].\end{aligned}$$

Здесь $\nabla_h u_{i\pm 1/2}$ суть центральные разности относительно средних точек сетки $x_{i\pm 1/2} = \frac{1}{2}(x_i + x_{i\pm 1})$. Как видно, вторая производная аппроксимируется соответствующей разностью на функциях из класса $C^3[x_{i-1}, x_{i+1}]$ с погрешностью порядка $O(h)$. Если же сетка равномерная, то

$$\Delta_h u_i = \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} = u_i'' + \frac{h^2}{12} u^{(4)}(\xi_i) \quad (3.8)$$

и погрешность аппроксимации имеет уже второй порядок, но только если функция $u(x)$ обладает ограниченной производной четвертого порядка.

Таким же образом, т. е. последовательной формальной заменой производных разностями, строятся конечно-разностные приближения более сложных дифференциальных выражений. Так, если $p(x)$ – некоторая дифференцируемая функция, то можно вывести аппроксимацию для “диффузионного члена”

$$\begin{aligned}(p(x)u')'_i &= \frac{(pu')_{i+1/2} - (pu')_{i-1/2}}{(h_i + h_{i-1})/2} + \psi_i^{(1)} = \frac{2p_{i-1/2}}{h_{i-1}(h_i + h_{i-1})} u_{i-1} \\ &\quad - \frac{2}{h_i + h_{i-1}} \left(\frac{p_{i-1/2}}{h_{i-1}} + \frac{p_{i+1/2}}{h_i} \right) u_i + \frac{2p_{i+1/2}}{h_i(h_i + h_{i-1})} u_{i+1} + \psi_i. \quad (3.9)\end{aligned}$$

Здесь $\psi_i^{(1)}$ – погрешность аппроксимации “внешней” производной, равная

$$\psi_i^{(1)} = \frac{h_i - h_{i-1}}{4} (pu'_i)' + O(h^2),$$

а ψ_i – полная погрешность, в которой дополнительно присутствуют ошибки приближения величин $(pu)'_{i\pm 1/2}$ с помощью центральных разностей:

$$\psi_i = \frac{h_i - h_{i-1}}{12} [pu''' + 3(pu')'']_i + O(h^2).$$

Отсюда, в частности, следует, что на равномерной сетке погрешность аппроксимации в случае переменного коэффициента диффузии $p(x)$ есть величина второго порядка, как и для второй производной. Более того, если сетка “почти равномерная” т. е. $h_i - h_{i-1} = O(h^2)$, то погрешность аппроксимации также имеет второй порядок.

В качестве примера приведем еще *разность четвертого порядка* на равномерной сетке:

$$\begin{aligned} \Delta_h^2 u_i &\equiv \Delta_h \frac{u_{i-1} - 2u_i + u_{i+1}}{h^2} = \frac{u_{i-2} - 4u_{i-1} + 6u_i - 4u_{i+1} + u_{i+2}}{h^4} = u_i^{(4)} + \psi_i, \\ |\psi_i| &\leq \frac{h^2}{6} M_6, \quad M_6 = \max_{x \in [x_{i-2}, x_{i+2}]} \{|u^{(6)}(x)|\}. \end{aligned} \quad (3.10)$$

Если в выражении (3.8) для второй разности использовать следующие члены разложения в ряд Тейлора и подставить аппроксимацию четвертой производной из (3.10), то для второй производной получается на равномерной сетке пятиточечное выражение с погрешностью четвертого порядка (на функциях $u(x)$ из класса C^6):

$$u''_i = \Delta_h^{(5)} u_i - \psi_h,$$

где $\psi_h = O(h^4)$ и использовано обозначение

$$\begin{aligned} \Delta_h^{(5)} u_i &= \Delta_h u_i - \frac{h^2}{12} \Delta_h^2 u_i \\ &= \frac{1}{12h^2} (-u_{i-2} + 16u_{i-1} - 30u_i + 16u_{i+1} - u_{i+2}). \end{aligned} \quad (3.11)$$

Естественно, что если дифференциальное выражение состоит из нескольких слагаемых, то каждое из них можно аппроксимировать отдельно, а общая погрешность будет равна сумме погрешностей слагаемых. Очевидно, что ее порядок будет не хуже минимального из порядков слагаемых ошибок. В то же время теоретически можно ожидать (и мы это действительно увидим), что в результате группирования отдельных компонент ошибок итоговая погрешность аппроксимации может уменьшаться.

3.1.3. Двумерный разностный оператор Лапласа

Аналогичные рассуждения проводятся и для дифференциальных выражений, определенных на функциях от нескольких независимых переменных. Если Ω – некоторая ограниченная область изменения аргумента $x = (x_1, \dots, x_n)$, то сетка Ω_h определяется как совокупность точек $x_k = (x_1^{(k)}, \dots, x_n^{(k)})$, расположенных достаточно плотно в Ω . Не вдаваясь в общий случай, когда строгость изложения требует достаточно сложного формализма, мы остановимся на функциях двух переменных $u(x, y)$ и *прямоугольных сетках* вида $x_{i+1} = x_i + h_i^x$, $y_{j+1} = y_j + h_j^y$. Для *равномерных* сеток будем использовать обозначения $h_x = h_i^x$ и $h_y = h_j^y$, а для *квадратных* – $h = h^x = h^y$. По контексту h может обозначать и максимальный шаг сетки, т. е.

$$h = \sup \left\{ \max_i \{h_i^x\}, \max_j \{h_j^y\} \right\}.$$

Для двумерного оператора Лапласа

$$\Delta u \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$$

его дискретный аналог на неравномерной прямоугольной сетке можно определить пятиточечным разностным выражением

$$\begin{aligned} \Delta_h u_{i,j} \equiv & \frac{2u_{i-1,j}}{h_{i-1}^x(h_i^x + h_{i-1}^x)} + \frac{2u_{i,j-1}}{h_{j-1}^y(h_j^y + h_{j-1}^y)} + \frac{2u_{i+1,j}}{h_i^x(h_i^x + h_{i-1}^x)} \\ & + \frac{2u_{i,j+1}}{h_j^y(h_j^y + h_{j-1}^y)} - \left(\frac{2}{h_{i-1}^x h_i^x} + \frac{2}{h_{j-1}^y h_j^y} \right) u_{i,j} = (\Delta u)_{ij} + \psi_{ij}, \end{aligned} \quad (3.12)$$

где для погрешности аппроксимации аналогично одномерному анализу в (3.8) получаем

$$\begin{aligned} |\psi_{i,j}| \leq & \frac{h_i^x - h_{i-1}^x}{3} \frac{\partial^3 u(\xi_i, y_j)}{\partial x^3} + \frac{h_j^y - h_{j-1}^y}{3} \frac{\partial^3 u(x_i, \eta_j)}{\partial y^3} + O(h^2), \\ & \xi_i \in [x_{i-1}, x_{i+1}], \quad \eta_j \in [y_{j-1}, y_{j+1}]. \end{aligned}$$

На равномерной сетке *пятиточечный разностный оператор Лапласа* упрощается, а порядок его погрешности повышается:

$$\begin{aligned} \Delta_h u_{i,j} \equiv & \frac{u_{i-1} - 2u_{i,j} + u_{i+1,j}}{h_x^2} + \frac{u_{i,j-1} - 2u_{i,j} + u_{i,j+1}}{h_y^2} = (\Delta u)_{i,j} + \psi_{i,j}, \\ |\psi_{i,j}| < & \frac{h^2}{6} M_4, \end{aligned} \quad (3.13)$$

что очевидно следует из (3.9).

Здесь мы без формального определения начали использовать понятие разностного оператора как некоторого преобразования в пространстве сеточных функций – в данном случае совокупности значений $u_h = \{u_{i,j}\}$ функции непрерывного аргумента в узлах сетки Ω_h . Соответственно совокупность значений сеточной функции вида $L_h u_{i,j}$ будем обозначать через $L_h u_h$. Поскольку на алгебраическом языке u_h есть вектор, размерность которого равна количеству узлов сетки Ω_h , а линейное преобразование в конечномерном пространстве представляется матрицей, то понятия разностного оператора и матрицы мы будем считать эквивалентными (естественно, с точностью до конкретной упорядоченности компонент сеточной функции u_h).

3.1.4. Определение погрешности аппроксимации разностного выражения

Теперь мы можем сформулировать достаточно общее понятие.

Определение 3.1. Пусть $Lu(x)$ – некоторое дифференциальное выражение, определенное в области изменения аргумента Ω , а $L_h u_k$ – определенное в точке x_k из Ω_h конечно-разностное выражение. Если при достаточно малых h имеют место соотношения

$$L_h u_k = Lu(x_k) + \psi_k^u, \quad |\psi_k^u| \leq C_k h^\gamma, \quad \gamma > 0, \quad (3.14)$$

где C_k – не зависящая от h постоянная, то ψ_k^u называется погрешностью аппроксимации разностного выражения в точке x_k на функции $u(x)$, а γ – порядком погрешности.

Поскольку выражение $L_h u_k$ на множестве сеточных функций $u_h = \{u_k\}$ определяет разностный оператор L_h , то предыдущему определению эквивалентно следующее: при выполнении условий (3.14) разностный оператор L_h аппроксимирует дифференциальный оператор L на функции $u(x)$ с погрешностью аппроксимации $\psi_h = \{\psi_k\}$ порядка γ . Соответственно можно записать

$$L_h u_h = (Lu)_h + \psi_h^u. \quad (3.15)$$

Заметим, что в (3.14) и (3.15) при обозначении погрешности используется, в отличие от предыдущего, верхний индекс u для акцентирования того, что аппроксимация осуществляется на функции $u(x)$ (и выражается именно через ее производные).

3.2. Аппроксимации дифференциальных уравнений

Перейдем теперь от дифференциальных выражений к уравнениям, которые абстрактно будем записывать в форме

$$Lu = f, \quad (3.16)$$

где L есть некоторый дифференциальный оператор, а f – известная функция.

3.2.1. Разностные уравнения Пуассона

Простейшим примером для нас будет одномерное уравнение Пуассона

$$-\frac{\partial^2 u}{\partial x^2} = f(x). \quad (3.17)$$

Аппроксимируя на неравномерной сетке Ω_h левую часть (3.17) с помощью соотношения (3.8), а на равномерной – трехточечным или пятиточечным разностным выражением вида (3.8) или (3.10), мы получим определенные на сеточных функциях уравнения

$$L_h u_i = f_i + \psi_i, \quad (3.18)$$

в которых величины ψ_i зависят от производных искомого решения. Если в предположении малости погрешностей аппроксимации ими пренебречь, то получим новые соотношения

$$(L_h v)_i = f_i, \quad (3.19)$$

которые и называются *разностными уравнениями*. Сеточная функция $v_h = \{v_i\}$, которая этим уравнениям удовлетворяет, называется *разностным решением*.

Очевидно, что если в (3.18) оператор L_h есть Δ_h из (3.8), то погрешность ψ_i имеет второй порядок, а при $L_h = \Delta_h^{(5)}$ – четвертый.

Структура оператора $\Delta_h^{(5)}$, с учетом вида уравнения (3.17), подсказывает, что четвертый порядок аппроксимации можно получить и для трехточечного разностного уравнения путем изменения его правой части. Такое уравнение имеет вид

$$-\Delta_h v_i \equiv \frac{-v_{i-1} + 2v_i - v_{i+1}}{h^2} = f_i + \frac{h^2}{12} \Delta_h f_i. \quad (3.20)$$

Понятно, что того же четвертого порядка будет погрешность и у разностного уравнения

$$-\Delta_h v_i = \left(f + \frac{h^2}{12} f'' \right)_i, \quad (3.21)$$

которое может быть получено и непосредственно из (3.8).

Приведем еще представляющие практический интерес аппроксимации уравнений Пуассона

$$-\frac{1}{x^\alpha} \frac{d}{dx} \left(x^\alpha \frac{du}{dx} \right) = f(x)$$

в цилиндрической ($\alpha = 1$) или сферической ($\alpha = 2$) системе координат. С помощью соотношения (3.9) легко проверить, что разностные уравнения

$$\frac{2}{x_i(h_i + h_{i-1})} \left[-\frac{x_{i-1/2}^\alpha}{h_{i-1}} u_{i-1} + \left(\frac{x_{i-1/2}^\alpha}{h_{i-1}} + \frac{x_{i+1/2}^\alpha}{h_i} \right) u_i - \frac{x_{i+1/2}^\alpha}{h_i} u_{i+1} \right] = f_i$$

при $x_{i-1/2} > 0$ аппроксимируют исходные дифференциальные уравнения с погрешностью $O(h)$ на неравномерной сетке и $O(h^2)$ на равномерной.

3.2.2. Погрешность аппроксимации разностного уравнения

В определенном смысле принципиальное отличие уравнения (3.19) от (3.20), (3.21) заключается в том, что первое в правой части содержит просто значения в узлах (проекцию на сетку) функции из правой части исходного дифференциального уравнения (3.17), а последние – некоторые разностные выражения, которые имеют уже ненулевую погрешность аппроксимации функции $f(x)$.

Поэтому определение погрешности аппроксимации разностного уравнения требует некоторого обобщения. Прежде всего, само разностное уравнение, аппроксимирующее дифференциальное уравнение (3.16), записываем в виде

$$L_h v_h = P_h f_h, \quad (3.22)$$

где P_h – оператор аппроксимации правой части:

$$P_h f_h = f_h + \psi_h^f. \quad (3.23)$$

Тогда сеточная функция u_h – проекция на Ω_h решения дифференциальной задачи – удовлетворяет разностному уравнению с погрешностью, включающей ψ_h^f , а именно

$$L_h u_h = P_h f_h + \psi_h, \quad \psi_h = \psi_h^u - \psi_h^f, \quad (3.24)$$

где ψ_h^u – определенная в (3.15) погрешность аппроксимации разностного оператора L_h на решении u . Сеточная функция ψ_h и составляет полную погрешность аппроксимации разностного уравнения. Несколько повторяясь, мы сформулируем введенное понятие формально.

Определение 3.2. Пусть $u(x)$ – определенное в области Ω решение дифференциального уравнения (3.16) и в любой точке x_k сетки Ω_h определено разностное уравнение

$$L_h v_k = P_h f_k. \quad (3.25)$$

Если при достаточно малом h после подстановки $v_k = u(x_k)$ в (3.25) выполняется равенство

$$L_h u_k = P_h f_k + \psi_k, \quad |\psi_k| \leq C h^\gamma, \quad \gamma > 0, \quad (3.26)$$

где постоянная $C > 0$ не зависит от x_k и h , то сеточная функция $\psi_h = \{\psi_k\}$ называется погрешностью аппроксимации разностного уравнения, а γ – порядком погрешности.

Немаловажно отметить, что если погрешность аппроксимации разностного выражения определяется на любой, вообще говоря, достаточно гладкой функции, то для разностного уравнения – только на решении дифференциального уравнения.

3.2.3. Аппроксимации диффузионно-конвективного уравнения. Конечно-разностные уравнения положительного типа и монотонные (абсолютно монотонные и условно монотонные) сеточные уравнения

Рассмотрим теперь так называемое *диффузионно-конвективное уравнение*, содержащее производные от нулевого до второго порядков включительно:

$$-(p(x)u')' + q(x)u' + r(x)u = f(x), \quad (3.27)$$

где первые два члена имеют физический смысл диффузионного и конвективного переносов, а остальные – источники некоторой субстанции (один из них линейно зависит от решения, а второй – “внешний источник” не зависит). Коэффициент диффузии $p(x)$ предполагается строго положительным. Аппроксимацию будем проводить для простоты на равномерной сетке, а акцент сделаем на особенностях разностных уравнений, связанных с использованием конечных разностей первого порядка различных типов.

Если в (3.27) конвективный член аппроксимируется правой разностью, а диффузионный – в соответствии с (3.9), то для i -го узла сетки получаем *третточечное разностное уравнение* вида

$$-a_i v_{i-1} + b_i v_i - c_i v_{i+1} = f_i, \quad (3.28)$$

коэффициенты которого описываются формулами

$$a_i = \frac{p_{i-1/2}}{h^2}, \quad c_i = \frac{p_{i+1/2}}{h^2} - \frac{q_i}{h}, \quad b_i = \frac{p_{i-1/2} + p_{i+1/2}}{h^2} - \frac{q_i}{h} + r_i, \quad (3.29)$$

а погрешность аппроксимации имеет порядок $O(h)$. Коэффициент b_i при неизвестной величине будем называть диагональным, а остальные коэффициенты ($-a_i$, $-c_i$ в данном случае) – внедиагональными.

Определение 3.3. Разностное уравнение называется *уравнением положительного типа*, если диагональный коэффициент является положительным, внедиагональные – неположительными, и имеется диагональное преобладание (в общем случае – модуль диагонального коэффициента уравнения не меньше суммы модулей внедиагональных).

Свойство положительности типа, как мы увидим, является очень полезным как в теоретическом плане, так и в практическом. Для уравнения (3.28) оно сводится к неравенствам

$$b_i > 0, \quad a_i, c_i \geq 0, \quad b_i \geq a_i + c_i. \quad (3.30)$$

Если в *неразложимой системе* все уравнения положительного типа, то она называется *системой положительного типа*.

Неразложимой называется система уравнений, в которой не может быть выделена подсистема меньшего порядка такая, что соответствующие искомым компоненты могут быть найдены вне зависимости от остальных неизвестных.

Система уравнений положительного типа является частным случаем более широкого класса монотонных систем уравнений.

Определение 3.4. Система линейных алгебраических уравнений с матрицей L_h называется монотонной, если из векторного неравенства $L_h v_h \geq 0$ следует $v_h \geq 0$ (здесь и далее векторные неравенства понимаются покомпонентно, как и равенства, а нуль по контексту понимается как вектор соответствующей размерности с нулевыми компонентами).

Замечание 3.1. Определения 3.3 и 3.4 содержат методологическое отличие: понятие положительного типа существует и для отдельного уравнения, а понятие монотонности – только для системы. Первое из них, естественно, распространяется на системы или подсистемы уравнений, если, например, условия (3.30) выполняются для некоторой совокупности индексов. Вообще же оба свойства не зависят от правых частей уравнений и характеризуются только коэффициентами при неизвестных. Доказательство того, что система положительного типа является монотонной, может быть проведено от противного. Если система разностных уравнений положительного типа (монотонная), то определяемый ею разностный оператор также положительного типа (монотонный). Другими словами, в этом случае матрица системы также называется монотонной.

Определение 3.5. Система разностных уравнений называется абсолютно монотонной, если она монотонна при любых шагах сетки, и условно монотонной, если она монотонна при выполнении некоторых условий на шаги сетки.

Замечание 3.2. Понятие положительности типа можно ввести не только для уравнения, но и для разностного выражения.

Аналогично интуитивно оправдано и употребление термина *монотонная аппроксимация* (или разностная схема, под этим фактически будет подразумеваться, что соответствующие разностные уравнения являются уравнениями положительного типа). Более строго категория “монотонность” характеризуется в теории матриц, и ниже мы активно будем ее эксплуатировать при алгебраическом анализе разностных методов.

Из формул (3.29) видно, что если $r_i \geq 0$ (это мы будем предполагать и в дальнейшем), то уравнение является абсолютно монотонным при $q_i \leq 0$, а при положительных q_i монотонность достигается только в случае выполнения условия

$$h \leq \frac{p_{i+1/2}}{q_i}, \quad (3.31)$$

которое означает, что при относительно больших конвективных членах шаг сетки должен быть достаточно мал. Важно подчеркнуть, что это требование на “малость” h может быть гораздо жестче, чем это следует из оценки погрешности аппроксимации. А использование чрезмерно большого числа узлов – это значительное удорожание вычислительного алгоритма.

Если в дифференциальном уравнении (3.27) $q(x_i) \geq 0$, то получить абсолютно монотонное разностное уравнение очень легко – достаточно для аппроксимации u' использовать не правую, а левую разность. В результате этого формулы для разностных коэффициентов заменятся с (3.29) на аналогичные:

$$a_i = \frac{p_{i-1/2}}{h^2} + \frac{q_i}{h}, \quad c_i = \frac{p_{i+1/2}}{h^2}, \quad b_i = \frac{p_{i-1/2} + p_{i+1/2}}{h^2} + \frac{q_i}{h} + r_i. \quad (3.32)$$

Погрешность аппроксимации здесь будет тоже первого порядка, а условие монотонности при $q_i < 0$ вместо (3.31) принимает форму

$$h \leq \frac{p_{i-1/2}}{|q_i|}. \quad (3.33)$$

Когда конвективный коэффициент $q(x)$ является знакопеременной функцией, то легко догадаться, как сконструировать и здесь абсолютно монотонную аппроксимацию. Это делается с помощью так называемых *направленных односторонних разностей* – в точках с $q(x_i) \leq 0$ используется правая разность, а при

$q_i > 0$ – левая. Получаемые разностные уравнения записываются формально следующим единым образом:

$$\begin{aligned} a_i &= \frac{p_{i-1/2}}{h^2} + \frac{|q_i| + q_i}{2h}, & c_i &= \frac{p_{i+1/2}}{h^2} + \frac{|q_i| - q_i}{2h}, \\ b_i &= \frac{p_{i-1/2} + p_{i+1/2}}{h^2} + \frac{|q_i|}{h} + r_i. \end{aligned} \quad (3.34)$$

Так как во всех трех случаях – (??) и (3.34) – погрешность аппроксимации имеет первый порядок, то совершенно естественно попытаться ее повысить в результате применения не односторонних разностей, а центральной. Тогда уравнение имеет по-прежнему трехточечный вид (3.28), а его коэффициенты имеют вид

$$a_i = \frac{p_{i-1/2}}{h^2} + \frac{q_i}{2h}, \quad c_i = \frac{p_{i+1/2}}{h^2} - \frac{q_i}{2h}, \quad b_i = \frac{p_{i-1/2} + p_{i+1/2}}{h^2} + r_i. \quad (3.35)$$

Достижимый здесь эффект – второй порядок погрешности – омрачается тем обстоятельством, что разностное уравнение при любом знаке q_i становится условно монотонным. Правда, условия (3.31), (3.33) при этом в два раза “смягчаются”:

$$h \leq \frac{2p_{i\pm 1/2}}{|q_i|}. \quad (3.36)$$

Замечание 3.3. Строго говоря, нарушение условий (3.31) или (3.36) свидетельствует о нарушении положительного типа уравнений, т. е. только об отсутствии достаточного условия монотонности. Однако мы увидим ниже (см. § 6.8), что для трехдиагональных матриц с положительными диагональными элементами и свойством диагонального преобладания неположительность внедиагональных элементов является необходимым условием монотонности.

3.2.4. Свойства симметричности и симметризуемости разностных уравнений и выражений

Остановимся теперь на такой важной характеристике разностных уравнений, как симметричность. Понятие *симметричности системы уравнений* обычно связывается с ее матрицей. Однако ее можно ввести также и для пары любой совокупности уравнений как “равенство коэффициента при k -й переменной

в i -м уравнении коэффициенту при i -й переменной в k -м уравнении" (аналогично можно определить и симметричность разностных выражений). Для совокупности трехточечных уравнений вида (3.28) это означает выполнение равенств $a_i = c_{i-1}$.

Очевидно, что для разностных уравнений Пуассона и даже уравнения диффузии с переменным коэффициентом условия симметрии выполняются, если только шаг сетки постоянный.

Если же величины h_i различны, то оказывается, что эти системы уравнений несимметричны, но относятся к другому важному классу – *симметризуемым системам уравнений*. Последнее означает, что они становятся симметричными после того, как каждое из них умножается на некоторое число (масштабирующий множитель). Как видно из (3.9), для разностного уравнения диффузии таким множителем является сумма $h_i + h_{i-1}$, поскольку

$$\tilde{a}_i = a_i(h_i + h_{i-1}) = \frac{2p_{i-1/2}}{h_{i-1}} = \tilde{c}_{i-1}(h_{i-1} + h_{i-2}).$$

Однако разностные диффузионно-конвективные уравнения (при любых рассмотренных выше или других способах аппроксимации) в отличие от предыдущих не являются симметричными даже в случае равномерной сетки и постоянства коэффициентов $p(x)$ и $q(x)$. Не симметризуются они и с помощью введения масштабирующего множителя $(h_i + h_{i-1})$.

Тем не менее системы трехчленных уравнений – и в этом их отличительное качество – могут быть симметризованы с помощью умножения каждого i -го уравнения на *масштабирующий множитель* $\varkappa_i$, если только разрешимы рекуррентные соотношения $\varkappa_i = \varkappa_{i-1}c_{i-1}/a_i$, которые следуют из условий симметрии отмасштабированных уравнений $\tilde{a}_i = \varkappa_i a_i = \tilde{c}_i = \varkappa_{i-1}c_{i-1}$.

Такая *симметризация* возможна (и только!) для *якобиевых систем трехчленных уравнений*, определяемых свойством $a_i c_{i-1} > 0$.

3.2.5. Разностные аппроксимации двумерного уравнения Пуассона

Перейдем теперь от аппроксимации одномерных дифференциальных уравнений к двумерным и начнем в определенном смысле с простейшего – уравнения Пуассона в декартовой системе координат

$$-\Delta u \equiv -\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = f(x, y). \quad (3.37)$$

При использовании тех же прямоугольных сеток, которые применялись в предыдущем параграфе, мы легко получаем, что пятиточечные разностные уравнения

$$-\Delta_h v_{ij} = f_{ij}, \quad (3.38)$$

где разностный оператор Δ_h определяется из выражений (3.12) на неравномерной сетке и (3.13) – на равномерной, будут иметь соответственно первый и второй порядки аппроксимации.

Для дальнейшего полезно представить пятиточечный разностный оператор Δ_h в виде суммы трехточечных:

$$\Delta_h = \Delta_h^x + \Delta_h^y, \quad (3.39)$$

где первый представляет вторую разность по x , а второй – по y .

На равномерной прямоугольной сетке, используя пятиточечные аппроксимации вторых производных вида (3.11), получаем *девятиточечное разностное уравнение Пуассона типа “большой крест”* :

$$\begin{aligned} -\Delta_h^{(k)} v_{i,j} \equiv & \frac{v_{i-2,j} - 16v_{i-1,j} + 30v_{ij} - 16v_{i+1,j} + v_{i+2,j}}{12h_x^2} \\ & + \frac{v_{i,j-2} - 16v_{i,j-1} + 30v_{i,j} - 16v_{i,j+1} + v_{i,j+2}}{12h_y^2} = f_{ij}, \end{aligned} \quad (3.40)$$

которое, как легко проверить, при ограниченности шестых производных решения дифференциального уравнения имеет погрешность $O(h^4)$.

На рис. 3.1 изображены сеточные шаблоны, где точками обозначены узлы, используемые в пятиточечном уравнении, и крестиками – дополнительные узлы для девятиточечного уравнения.

Как видно из (3.40), девятиточечное уравнение “большой крест” ни при каких соотношениях шагов h_x, h_y не является уравнением положительного типа.

Для уравнения Пуассона обнаруживается еще другая, замечательная по своим качествам девятиточечная аппроксимация высокого порядка, впервые предложенная Ш. Е. Микеладзе в 1934 г. Она основана на следующем свойстве разностных операторов: если использовать введенные в (3.39) обозначения Δ_h^x, Δ_h^y , то для достаточно гладкой функции $u(x, y)$ справедливы соотношения

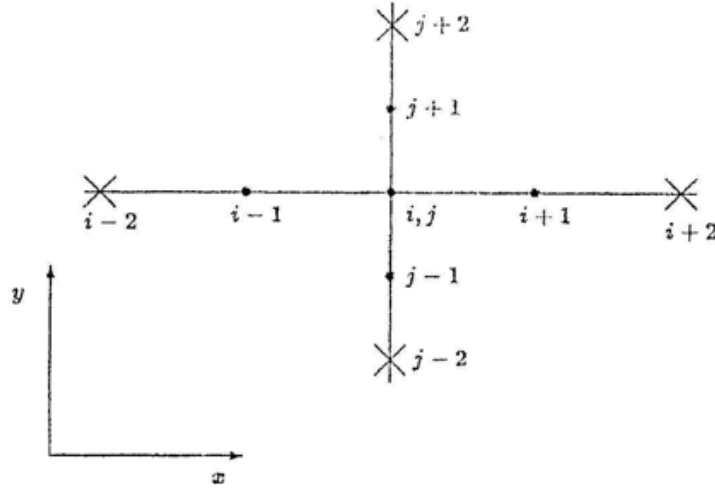


Рис. 3.1. Сеточные шаблоны для пятиточечного уравнения и девятиточечного уравнения типа "большой крест"

$$\begin{aligned} \Delta_h^{(9)} u_{ij} &\equiv \left(\Delta_h^x + \Delta_h^y + \frac{h_x^2 + h_y^2}{12} \Delta_h^x \Delta_h^y \right) u_{ij} = \\ &= \left[\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{h_x^2}{12} \frac{\partial^2}{\partial x^2} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) + \frac{h_y^2}{12} \frac{\partial^2}{\partial y^2} \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \right]_{i,j} \\ &\quad + O(h^4), \quad (3.41) \end{aligned}$$

которые следуют из равенства

$$\Delta_h^x u_{ij} = \left(\frac{\partial^2 u}{\partial x^2} + \frac{h_x^2}{12} \frac{\partial^4 u}{\partial x^4} \right)_{i,j} + \frac{2h_x^4}{6!} \frac{\partial^6 u(\xi_i, y_j)}{\partial x^6}$$

и такого же выражения для $\Delta_h^y u_{ij}$. Заменяя теперь в правой части (3.41) члены $\Delta u_{i,j}$ через $f_{i,j}$ в силу исходного уравнения и отбрасывая погрешность $O(h^4)$, получаем, что разностное уравнение

$$\begin{aligned} -\Delta_h^{(9)} v_{i,j} &\equiv \frac{5(h_x^2 + h_y^2)}{3h_x^2 h_y^2} v_{i,j} - \frac{1}{6h_x^2} \left(5 - \frac{h_x^2}{h_y^2} \right) (v_{i-1,j} + v_{i+1,j}) - \\ &\quad - \frac{1}{6h_y^2} \left(5 - \frac{h_y^2}{h_x^2} \right) (v_{i,j-1} + v_{i,j+1}) \\ &\quad - \frac{h_x^2 + h_y^2}{12h_x^2 h_y^2} (v_{i-1,j-1} + v_{i-1,j+1} + v_{i+1,j-1} + v_{i+1,j+1}) \end{aligned}$$

$$= f_{i,j} + \frac{h_x^2}{12} \frac{\partial^2 f_{i,j}}{\partial x^2} + \frac{h_y^2}{12} \frac{\partial^2 f_{i,j}}{\partial y^2} \quad (3.42)$$

аппроксимирует исходное уравнение (3.37) с четвертым порядком. Очевидно, что порядок погрешности в (3.42) не изменится, если в правой части производные от функции f заменить на вторые разности. Сеточный шаблон полученного *девятиточечного уравнения Пуассона типа "ящик"* изображен на рис. 3.2.

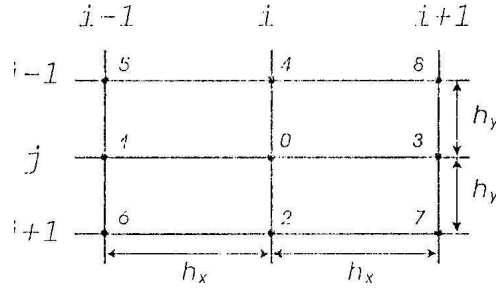


Рис. 3.2. Сеточные шаблоны для девятиточечного уравнения типа "ящик"

Простой анализ знаков коэффициентов показывает, что уравнение (3.42) будет положительного типа только при не очень сильном отличии прямоугольной сетки от квадратной, а именно при выполнении условия

$$\frac{1}{\sqrt{5}} \leq \frac{h_x}{h_y} \leq \sqrt{5}. \quad (3.43)$$

Замечательно, что на квадратной сетке ($h_x = h_y = h$) аппроксимационные свойства разностного уравнения типа "ящик" еще усиливаются. Это следует из соотношения, которое легко проверяется при использовании большего числа членов ряда Тейлора:

$$\Delta_h^{(9)} u_{i,j} = \Delta u_{i,j} + \frac{2}{4!} h \Delta^2 u_{i,j}^2 + \frac{2}{6!} h^4 [\Delta^3 u + 2(\Delta u)_{xxyy}]_{i,j} + O(h^6). \quad (3.44)$$

Таким образом, если решение уравнения Пуассона обладает ограниченными производными до восьмого порядка включительно, то разностное уравнение

$$\begin{aligned}
-\Delta_h^{(9)} v_{i,j} &\equiv \frac{1}{6h^2} \left[20v_{i,j} - 4(v_{i-1,j} + v_{i,j-1} + v_{i+1,j} + v_{i,j+1}) - \right. \\
&\quad \left. - v_{i-2,j} - v_{i,j-2} - v_{i+2,j} - v_{i,j+2} \right] = \\
&= f_{i,j} + \frac{2h^2}{4!} \Delta f_{i,j} + \frac{2h^4}{6!} \left(\Delta^2 f_{i,j} + 2 \frac{\partial^4 f_{i,j}}{\partial x^2 \partial y^2} \right) \quad (3.45)
\end{aligned}$$

имеет погрешность аппроксимации шестого (!) порядка.

Переходя от уравнения Пуассона к более общему – *диффузионному уравнению*

$$-\frac{\partial}{\partial x} p \frac{\partial u}{\partial x} - \frac{\partial}{\partial y} q \frac{\partial u}{\partial y} = f(x, y), \quad (3.46)$$

с помощью соотношения (3.9) получаем, что разностное уравнение

$$\begin{aligned}
&\left[\frac{2}{h_i^x + h_{i-1}^x} \left(\frac{p_{i-1/2,j}}{h_{i-1}^x} + \frac{p_{i+1/2,j}}{h_i^x} \right) + \frac{2}{h_j^y + h_{j-1}^y} \left(\frac{q_{i,j-1/2}}{h_{j-1}^y} + \frac{q_{i,j+1/2}}{h_j^y} \right) \right] v_{i,j} \\
&\quad - \frac{2p_{i-1/2,j}}{h_{i-1}^x (h_i^x + h_{i-1}^x)} v_{i-1,j} - \frac{2q_{i,j-1/2}}{h_{j-1}^y (h_j^y + h_{j-1}^y)} v_{i,j-1} \\
&\quad - \frac{2p_{i+1/2,j}}{h_i^x (h_i^x + h_{i-1}^x)} v_{i+1,j} - \frac{2q_{i,j+1/2}}{h_j^y (h_j^y + h_{j-1}^y)} v_{i,j+1} = f_{i,j} \quad (3.47)
\end{aligned}$$

имеет погрешность аппроксимации первого порядка на неравномерной сетке и второго – на равномерной (при достаточно гладких коэффициентах диффузии $p(x, y)$ и $q(x, y)$).

Чтобы не загромождать изложение, мы не будем выписывать пятиточечные разностные диффузионно-конвективные уравнения. Они конструируются по образу и подобию рассмотренных выше одномерных аппроксимаций, с применением односторонних или центральных разностей для производных по x , y , и условия монотонности для них выглядят аналогично (3.31) или (3.36).

Остановимся теперь на свойствах симметричности двумерных разностных уравнений. Система на пятиточечном шаблоне в общей форме может быть записана как

$$-a_{i,j} u_{i-1,j} - b_{i,j} u_{i,j-1} - c_{i,j} u_{i+1,j} - d_{i,j} u_{i,j+1} + e_{i,j} u_{i,j} = f_{i,j}, \quad (3.48)$$

а условия симметрии сводятся к равенствам

$$a_{i,j} = c_{i-1,j}, \quad b_{i,j} = d_{i,j-1}. \quad (3.49)$$

Отсюда видно, что свойства двумерных задач аналогичны свойствам одномерных. Разностные уравнения Пуассона и диффузии являются симметричными на равномерных сетках и несимметричными – на неравномерных. В последнем случае они оказываются симметризуемыми, а именно путем умножения на масштабирующий коэффициент $(h_i^x + h_{i-1}^x)(h_j^y + h_{j-1}^y)$. Рассмотренные девятиточечные уравнения повышенной точности на равномерных сетках также являются симметричными. Двумерные диффузионно-конвективные разностные уравнения несимметричны на любых сетках, а вопрос с их симметризуемостью более тонкий, чем в одномерном случае, и мы на нем остановимся ниже.

3.2.6. Примеры трехмерных разностных уравнений

Остановимся – очень коротко – на аппроксимации трехмерных дифференциальных уравнений. Принципиально нового здесь ничего не появляется, но технически весь анализ, естественно, усложняется. Сетка рассматривается теперь *параллелепипедоидальная*:

$$x_{i+1} = x_i + h_i^x, \quad y_{j+1} = y_j + h_j^y, \quad z_{k+1} = z_k + h_k^z,$$

а компоненты сеточных функций обозначаются тремя индексами: $u_h = \{u_{i,j,k} = u(x_i, y_j, z_k)\}$.

Ограничимся уравнением Пуассона

$$-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = f(x, y, z), \quad (3.50)$$

конечно-разностные аналоги которого приведем только на равномерной сетке. Аналогом пятиточечного уравнения здесь является *семиточечное сеточное уравнение Пуассона*

$$\begin{aligned} -\Delta_h^{(7)} v_{i,j,k} \equiv & -\frac{v_{i-1,j,k} + v_{i+1,j,k}}{h_x^2} - \frac{v_{i,j-1,k} + v_{i,j+1,k}}{h_y^2} \\ & - \frac{v_{i,j,k-1} + v_{i,j,k+1}}{h_z^2} + 2\left(\frac{1}{h_x^2} + \frac{1}{h_y^2} + \frac{1}{h_z^2}\right)v_{i,j,k} = f_{i,j,k}. \end{aligned} \quad (3.51)$$

Также подобно двумерному случаю (см. (3.47)) выписывается разностное уравнение диффузии, которое имеет погрешность

аппроксимации $O(h)$ на неравномерной сетке и $O(h^2)$ на равномерной.

Что более любопытно – удастся построить на равномерной сетке прямой аналог девятиточечного разностного уравнения Пуассона (3.42) повышенного порядка точности. А именно из непосредственного проверяемого соотношения

$$\begin{aligned} \Delta_h^{(19)} u_{i,j,k} &\equiv \left(\Delta_h^x + \Delta_h^y + \Delta_h^z + \frac{h_x^2 + h_y^2}{12} \Delta_h^x \Delta_h^y \right. \\ &\quad \left. + \frac{h_y^2 + h_z^2}{12} \Delta_h^y \Delta_h^z + \frac{h_x^2 + h_z^2}{12} \Delta_h^x \Delta_h^z \right) u_{i,j,k} = \\ &= \Delta u_{i,j,k} + \left(\frac{h_x^2}{12} \frac{\partial^2}{\partial x^2} + \frac{h_y^2}{12} \frac{\partial^2}{\partial y^2} + \frac{h_z^2}{12} \frac{\partial^2}{\partial z^2} \right) \Delta u_{i,j,k} + O(h^4), \end{aligned} \quad (3.52)$$

где Δ_h^z – вторая разность ”по z ” а $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ – трехмерный дифференциальный оператор Лапласа, следует, что 19-точечное разностное уравнение Пуассона

$$-\Delta_h^{(19)} v_{i,j,k} = f_{i,j,k} + \frac{1}{12} \left(h_x^2 \frac{\partial^2}{\partial x^2} + h_y^2 \frac{\partial^2}{\partial y^2} + h_z^2 \frac{\partial^2}{\partial z^2} \right) f_{i,j,k} \quad (3.53)$$

имеет погрешность аппроксимации $O(h^4)$. Сеточный шаблон для такого уравнения изображен на рис. 3.3, где жирными точками обозначены узлы семиточечного шаблона, а простыми – узлы, дополняющие его до 19-точечного.

Уравнения (3.53) являются уравнениями положительного типа при выполнении неравенств $\frac{1}{\sqrt{3}} \leq \frac{h_z}{h_x}, \frac{h_y}{h_z}, \frac{h_x}{h_y} \leq \sqrt{3}$.

Отметим, что если сетка кубическая ($h_x = h_y = h_z = h$), то с помощью 27-точечного уравнения

$$\begin{aligned} -\Delta_h^{27} u &\equiv - \left[\Delta_h^x + \Delta_h^y + \Delta_h^z + \frac{h^2 (\Delta_h^x \Delta_h^y + \Delta_h^x \Delta_h^z + \Delta_h^y \Delta_h^z)}{6} + \frac{h^4}{30} \Delta_h^x \Delta_h^y \Delta_h^z \right] u \\ &= f + \frac{h^2}{12} \Delta f + \frac{h^4}{360} \Delta^2 f + \frac{h^4}{180} \left(\frac{\partial^4 f}{\partial x^2 \partial y^2} + \frac{\partial^4 f}{\partial x^2 \partial z^2} + \frac{\partial^4 f}{\partial y^2 \partial z^2} \right) + O(h^6) \end{aligned}$$

строится схема порядка $O(h^6)$.

Естественно, что в этих уравнениях шестой порядок погрешности сохраняется, если производные от функции f заменить на соответствующие разностные аппроксимации второго порядка точности.

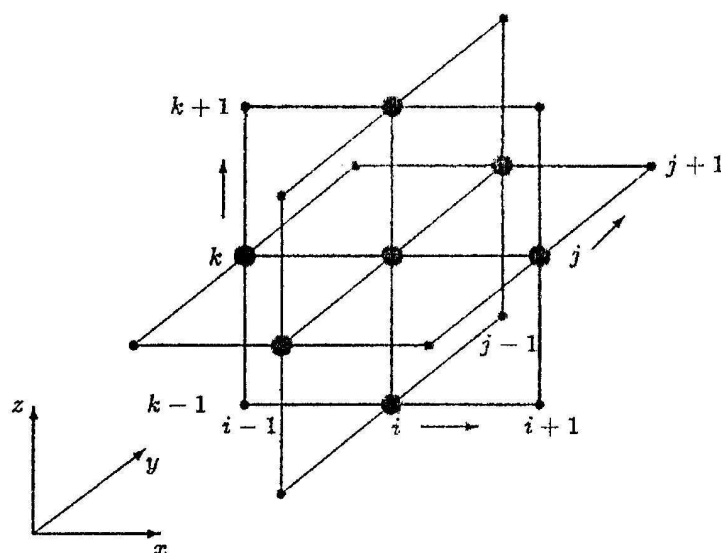


Рис. 3.3. Сеточные шаблоны для трехмерных семи- и девятнадцатиточечного уравнения

3.3. Методы аппроксимации граничных условий

До сих пор мы говорили только об отдельных уравнениях или их совокупности, не акцентируя даже внимания на термине “система уравнений” хотя на алгебраическом языке любую совокупность уравнений формально можно назвать системой. Исторически или по другим причинам, но в методах конечных разностей применяются и исследуются только системы, имеющие одинаковое число уравнений и неизвестных.

В теории дифференциальных уравнений краевая задача определена полностью, когда сформулированы граничные условия, причем их количество и вид, необходимый и достаточный для существования единственного решения, зависят от характера исходных уравнений.

Мы пока рассматривали конечно-разностные уравнения в точках сетки Ω_h , внутренних по отношению к расчетной области Ω , где определено дифференциальное уравнение. Однако в этих разностных уравнениях неизбежно участвуют неизвестные, соответствующие узлам, лежащим или на границе Γ

области Ω , или даже вне ее, так что в таком виде полученная алгебраическая система является недоопределенной и имеет бесконечное число решений.

Для замыкания системы к ней надо добавить уравнения, аппроксимирующие граничные условия, с тем чтобы уравнять число неизвестных и разностных уравнений (если при этом система окажется невырожденной, то она будет иметь единственное решение).

3.3.1. Аппроксимация условий 1-го рода в одномерных задачах

Двигаясь от простого к сложному, начнем с одномерных задач, а здесь, в свою очередь, – с простейшего в алгоритмическом плане условия Дирихле. Для понимания нам достаточно остановиться на уравнении Пуассона (3.17) в области $\Omega = \{a < x < b\}$, которое мы теперь дополним условиями первого рода

$$u|_{x=a} = u_a, \quad u|_{x=b} = u_b, \quad (3.54)$$

где u_a и u_b – заданные числа. Даже в такой тривиальной, казалось бы, задаче имеются методологические нюансы.

- а) Самый простой случай такой, когда граница области совпадает с узлами сетки, т. е.

$$x_0 = a, \quad x_{N+1} = b, \quad x_i = x_{i-1} + h_{i-1}, \quad i = 1, \dots, N+1. \quad (3.55)$$

При этом разностные уравнения дополняются соотношениями, точно передающими краевые условия (3.54):

$$v_0 = u_0 = u_a, \quad v_{N+1} = u_{N+1} = u_b. \quad (3.56)$$

- б) *Интерполяция нулевого порядка.* Если граница не совпадает с узлом сетки, то самым тривиальным способом является *простой снос* граничных значений функции в ближайшие узлы сетки. Например, для левой границы в силу ряда Тейлора

$$u_0 = u_a + (x_0 - a)u'(\xi_0), \quad \xi_0 \in [a, x_0],$$

граничное “разностное” уравнение $v_0 = u_a$ имеет погрешность аппроксимации первого порядка.

- в) *Аппроксимация Коллатца*. Для предыдущего случая естественно использовать более точную линейную интерполяцию:

$$v_0 = \frac{u_a(x_1 - x_0) - v_1(a - x_0)}{x_1 - a}, \quad (3.57)$$

которая, как легко проверить, имеет погрешность $O(h^2)$.

Если в пунктах б) и в) $x_0 < a$, то приведенные аппроксимации представляют собой экстраполяцию решения за пределы расчетной области, а сама точка x_0 в этом случае называется фиктивным узлом сетки.

- г) Логическим развитием рассмотренного подхода является применение интерполяции (экстраполяции) второго и более высокого порядков с помощью нескольких внутренних узлов. Примером может быть квадратичная аппроксимация с погрешностью $O(h^3)$:

$$v_0 = u_a + (u_a - u_1) \frac{a - x_0}{a - x_2} + \frac{(a - x_0)^2}{2} \times \left[\frac{u_a}{(x_1 - a)(x_2 - a)} + \frac{u_2}{(x_2 - x_1)(x_2 - a)} - \frac{2u_1}{(x_1 - a)(x_2 - x_1)} \right]. \quad (3.58)$$

Совершенно очевидно, что учет граничного условия на правой границе $x = b$ проводится аналогично.

Дополнительные вопросы возникают при аппроксимации граничных условий, когда используются разностные уравнения повышенной точности типа (3.11). Поскольку в первом и предпоследнем узлах сетки около границы пятиточечные выражения не определены, один из способов преодоления возникающих трудностей – использование в таких точках трехточечных уравнений, после чего аппроксимация краевых условий может делаться “обычным” образом.

3.3.2. Условия 2-го и 3-го рода (одномерный случай)

Перейдем теперь к рассмотрению краевых условий, содержащих производные. Поскольку условия 2-го рода формально являются частным случаем граничных условий Ньютона, рассмотрим только последние:

$$\alpha u + \beta \frac{\partial u}{\partial n} = \gamma,$$

где $\frac{\partial}{\partial n}$ означает $-\frac{\partial}{\partial x}$ при $x = a$ и $\frac{\partial}{\partial x}$ при $x = b$. Применяя линейную интерполяцию для решения и заменяя производную конечной разностью первого порядка, для левой границы получаем

$$\alpha \frac{v_0(x_1 - a) + v_1(a - x_0)}{x_1 - x_0} - \beta \frac{v_1 - v_0}{x_1 - x_0} = \gamma. \quad (3.59)$$

Первый член здесь при любом значении координаты границы a имеет погрешность $O(h^2)$ (кроме $a = x_0$ или $a = x_1$, когда она нулевая). Вторым же член при $a = x_{1/2}$ представляет собой центральную разность, аппроксимирующую нормальную производную со вторым порядком, а во всех остальных случаях его ошибка есть $O(h)$. Таким образом, с рассмотренной точки зрения при наличии краевых условий второго или третьего рода границу целесообразно располагать в середине шага сетки.

Переписывая разностное соотношение (3.59) для нулевого узла сетки с явным выделением диагонального коэффициента:

$$[\alpha(x_1 - a) + \beta]v_0 - [\beta - \alpha(a - x_0)]v_1 = \gamma(x_1 - x_0), \quad (3.60)$$

мы можем отметить, что при $\alpha\beta < 0$ это уравнение не является уравнением положительного типа (мы предполагаем $x_0 \leq a \leq x_1$), а при условии $\alpha\beta \geq 0$ оно этим свойством обладает, если только шаг сетки достаточно мал (конкретнее, при $[\alpha(x_1 - a) + \beta][\beta - \alpha(a - x_0)] \geq 0$). Такой же вывод получается и при аппроксимации условия Ньютона на правой границе области. Отметим, что совпадение знаков коэффициентов α, β является "хорошим" качеством и с точки зрения корректности дифференциальной постановки задачи.

Что касается влияния уравнения (3.60) на свойство симметрии, то здесь достаточно вспомнить разностное уравнение Пуассона или диффузии для первого узла сетки. Сравнение показывает, что в общем случае симметрия нарушается, но для одномерных трехточечных уравнений возможна симметризация.

Аналогично условию Дирихле, для граничных условий второго и третьего родов технически несложно использовать интерполяцию более высокого порядка с привлечением дополнительных узлов и повышением порядка аппроксимации. Однако свойства монотонности и симметричности при этом практически всегда теряются.

3.3.3. Аппроксимации граничных условий в двумерных задачах

Рассмотрим теперь особенности аппроксимации краевых условий для двумерных задач. Ситуация здесь определяется главным образом конфигурацией границ. Наиболее простые задачи – когда расчетная область представляет собой прямоугольник со сторонами, параллельными координатным осям. В этом случае направление нормали границы совпадает с одной из координатных линий и методы аппроксимации фактически сводятся к одномерным.

Для иллюстрации существенно двумерных эффектов рассмотрим изображенные на рис. 3.4 узлы сетки, расположенные около криволинейной границы.

Условия Дирихле для пятиточечных разностных уравнений могут быть аппроксимированы достаточно просто. При использовании интерполяции нулевого, первого или более высоких порядков во внутренних околограничных узлах само дифференциальное уравнение дискретизируется на обычном сеточном шаблоне (узлы с номерами 0, 1, 2, 3, 4 на рис. 3.4), но с использованием внешних (фиктивных) узлов. Затем для последних выписывается “одномерный” многочлен Лагранжа вдоль соответству-

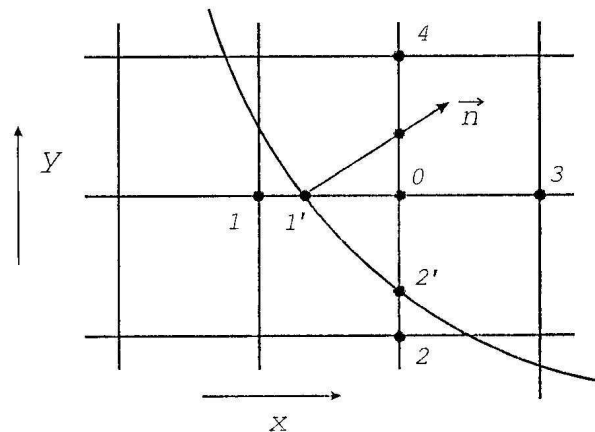


Рис. 3.4. Сеточный шаблон вблизи криволинейной границы области

ющих координатных линий.

В качестве примера рассмотрим аппроксимацию Коллатца (линейная экстраполяция) для изображенного на рис. 3.4 расположения узлов:

$$v_1 = u_{1'} \frac{l_{01}}{l_{01'}} - v_0 \frac{l_{11'}}{l_{01'}}, \quad v_2 = u_{2'} \frac{l_{02}}{l_{02'}} - v_0 \frac{l_{22'}}{l_{02'}}. \quad (3.61)$$

Здесь $u_{1'}$, $u_{2'}$ – известные значения решения в точках $1'$ и $2'$ пересечения координатных линий с границей, а l_{mn} – расстояния между узлами с номерами m и n . Полезно отметить, что если величины v_1 и v_2 из (3.61) подставить в разностное уравнение (ради простоты – Пуассона) для точки θ , то в итоге имеем

$$-\frac{2}{l_{01'}l_{13}}v_{1'} - \frac{2}{l_{02'}l_{24}}v_{2'} - \frac{2}{l_{03}l_{13}}v_3 - \frac{2}{l_{04}l_{24}}v_4 + \left(\frac{2}{l_{01'}l_{03}} + \frac{2}{l_{02'}l_{04}} \right) v_0 = f_0. \quad (3.62)$$

Парадоксальным, на первый взгляд, является следующий факт: если соотношения (3.61) имеют погрешность $O(h^2)$, то разностное уравнение (3.62) аппроксимирует исходное уравнение Пуассона с погрешностью только нулевого порядка (ошибка равна величине $O(1)$, т. е. при $h \rightarrow 0$ она не стремится к нулю, но, что тоже важно, как мы увидим позже, – не стремится и к бесконечности, оставаясь ограниченной).

Другое важное качество аппроксимации Коллатца – сохранение свойств симметричности (при $h_i^x = h_x$, $h_j^y = h_y$) или симметризуемости системы разностных уравнений; чтобы убедиться в этом, достаточно сравнить (3.62) с аналогичными уравнениями для узлов 3 и 4 .

Здесь уместно сделать одно замечание методологического характера. Данная аппроксимация использует экстраполяцию решения во внешние узлы, хотя оно определено только внутри Ω . Это предполагает возможность гладкого продолжения решения за область, что в одномерных задачах не вызывает сомнения, а в двумерных требует, вообще говоря, отдельного изучения.

Альтернативный способ учета граничных условий первого рода – это так называемая *аппроксимация Шортли–Уэллера*. Суть ее заключается в том, что для околограничного узла строится пятиточечное разностное уравнение на шаблоне, в котором вместо фиктивных (внешних) соседних узлов используются точки пересечения соответствующих линий сетки с границей области (точки $1'$ и $2'$ вместо 1 и 2 на рис. 3.4). В таком случае

строящееся по обычным правилам разностное уравнение Пуассона, например, принимает вид

$$-\frac{2v_{1'}}{l_{01'}(l_{01'} + l_{03})} - \frac{2v_{2'}}{l_{02'}(l_{02'} + l_{04})} - \frac{2v_3}{l_{03}(l_{01'} + l_{03})} - \frac{2v_4}{l_{04}(l_{02'} + l_{04})} + \left(\frac{2}{l_{01'}l_{03}} + \frac{2}{l_{02'}l_{04}} \right) v_0 = f_0. \quad (3.63)$$

При этом учет краевых условий уравнениями вида $v_{1'} = u_{1'}$, $v_{2'} = u_{2'}$ производится точно, а само уравнение (3.63) имеет погрешность аппроксимации $O(h)$, т. е. на порядок выше, чем (3.62). Общее выигрышное качество этих уравнений – это то, что оба они положительного типа при любых соотношениях шагов сетки и расстояния узлов от границы.

Большой минус аппроксимации Шортли–Уэллера заключается в том, что уравнение (3.63) нарушает свойство симметризуемости разностной системы, даже для равномерной сетки (разумеется, если только $l_{01'} \neq l_{01}$ или $l_{02'} \neq l_{02}$, точнее, если граница пересекает координатные линии не в узлах сетки).

Аппроксимация условий, содержащих нормальные производные к границам, не параллельным координатным линиям сетки, технически усложняется, хотя принципиально это выглядит просто. Достаточно нормаль к граничной точке сетки продолжить внутрь области до пересечения с каким-либо отрезком, соединяющим два внутренних узла сетки. Затем в полученной вспомогательной точке решение линейно проинтерполировать через узловое значения, а саму нормальную производную аппроксимировать по вспомогательной и граничной точкам. Необходимое разностное уравнение, аппроксимирующее условие 3-го рода в точке $1'$ на рис. 3.4, выглядит следующим образом:

$$\alpha v_{1'} + \frac{\beta}{l_{51'}} \left(\frac{l_{54}v_0 + l_{05}v_4}{l_{04}} - v_{1'} \right) = \gamma. \quad (3.64)$$

Погрешность аппроксимации его, очевидно, есть $O(h)$. Очевидно также, что использование такого уравнения нарушает симметрию системы и ее симметризуемость в общем случае не представляется возможной.

Однако если граница области параллельна координатным линиям, то аппроксимация условий 2-го и 3-го рода делается просто по “одномерным” формулам (3.59), (3.60).

На трехмерных задачах мы останавливаться не будем, так как методика остается той же, только необходимо выполнять более сложные геометрические построения.

3.3.4. Погрешности аппроксимации разностных краевых задач

Приведенные выше примеры резюмируем теперь более общими рассуждениями. Пусть краевая задача описывается уравнениями

$$Lu(x) = f(x), \quad x \in \Omega, \quad lu(x) = g(x), \quad x \in \Gamma, \quad (3.65)$$

где l – дифференциальный оператор краевых условий, а $g(x)$ – известная функция, как и $f(x)$.

Обозначая через l_h оператор аппроксимации левой части граничного уравнения, а через Q_h – его правой части (чаще всего Q_h – единичный оператор), второе уравнение из (3.65) мы можем привести к виду

$$l_h u_h = Q_h g_h + \psi_h^g, \quad (3.66)$$

где ψ_h^g – сеточная функция погрешности аппроксимации краевого условия, которая в общем случае состоит из ошибок приближения левой и правой частей.

Разностные “граничные” уравнения получаются из (3.66) отбрасыванием остаточных членов, и в итоге дискретизированная краевая задача описывается алгебраической системой

$$L_h v_h = P_h f_h, \quad l_h v_h = Q_h g_h. \quad (3.67)$$

Мы всегда будем считать, что разностная краевая задача вида (3.67) имеет одинаковое число уравнений и неизвестных. Относительно последних следует провести еще некоторую внутреннюю классификацию.

Через $\bar{\Omega}_h$ мы обозначали множество всех узлов сетки, участвующих при построении разностных уравнений, в том числе внутренних по отношению к расчетной области Ω , лежащих на ее границе Γ или даже за ней. Соответственно проведем разбиение $\bar{\Omega}_h = \Omega_h \cup \Gamma_h$, где через Ω_h обозначено множество внутренних узлов, а через Γ_h – множество внешних. Соответственно

полное количество узлов N складывается из соответствующих слагаемых: $\bar{N} = N + N_\Gamma$.

Иногда оказывается удобным часть неизвестных из системы (3.67) исключить в силу тривиальности отдельных уравнений (например, для граничных узлов с условиями Дирихле) или из других соображений. Тогда получается новая система разностных уравнений, которую можно записать в виде

$$\tilde{L}_h \tilde{v}_h = \tilde{f}_h, \quad (3.68)$$

где оператор $\tilde{L}_h$ объединяет уже разностные аналоги и исходного дифференциального уравнения, и краевых условий, а сеточная функция $\tilde{v}_h$ содержит число компонент $\bar{N}$, удовлетворяющих неравенству $\bar{N} > \tilde{N} \geq N$. Необходимо отметить, что формальное рассмотрение погрешности аппроксимации при переходе к (3.68) меняется, как мы это видели на примере аппроксимации Коллатца в представлении (3.62).

В заключение данного параграфа укажем еще на один характерный класс задач, возникающих при наличии расчетных областей с периодическими повторяющимися – по одной из координат – геометрическими и материальными свойствами. В этом случае само решение является также периодическим, что, например, для отрезка $a \leq x \leq b$ записывается в виде $u(x) = u(x + b - a)$. Естественным путем аппроксимации таких “краевых условий” (вообще говоря, постановку задачи при этом можно рассматривать и для бесконечной области) является совмещение узлов с границами области $x_0 = a$, $x_{N+1} = b$ и использование точных соотношений $u_0 = u_{N+1}$, $u_{-1} = u_N$, $u_{N+2} = u_1$ (в предположении равенства шагов сетки около левой и правой границ).

3.4. Компактные схемы повышенной точности для уравнений с переменными коэффициентами

Напомним, что в п. 3.2.5 была рассмотрена девятиточечная разностная схема Микеладзе (3.42), обеспечивающая на сеточном шаблоне типа “ящик” равномерной прямоугольной сетки аппроксимацию порядка $O(h^4)$. Оказывается, что используемая при ее построении идея переносится и на более слож-

ные дифференциальные уравнения. Получаемые при этом схемы стали называться *компактными* в отличие от схем на шаблоне вида “большой крест”.

3.4.1. Аппроксимации диффузионно-конвективных уравнений

Рассмотрим аппроксимацию диффузионно-конвективного дифференциального уравнения

$$Lu = (L_x + L_y)u \equiv -\frac{\partial}{\partial x}p\frac{\partial u}{\partial x} + a\frac{\partial u}{\partial x} - \frac{\partial}{\partial y}q\frac{\partial u}{\partial y} + b\frac{\partial u}{\partial y} = f, \quad (3.69)$$

где коэффициенты p, q являются положительными функциями. На прямоугольной равномерной сетке

$$x_i = x_0 + ih_x, \quad i = 1, \dots, I+1; \quad y_j = y_0 + jh_y, \quad j = 1, \dots, J+1,$$

определим сначала “стандартный” одномерный разностный оператор

$$(L_x^h u)_{i,j} \equiv \frac{1}{h_x^2} [p_{i+1/2,j}(u_{i,j} - u_{i+1,j}) + p_{i-1/2,j}(u_{i,j} - u_{i-1,j})] + \frac{a_{i,j}}{2h_x}(u_{i+1,j} - u_{i-1,j}), \quad (3.70)$$

где $p_{i\pm 1/2,j} = p(\frac{x_i + x_{i\pm 1}}{2}, y_j)$, а погрешность аппроксимации имеет второй порядок и находится с помощью ряда Тейлора из соотношений

$$L_x^h u = L_x u - \frac{h_x^2}{12} R_x u + O(h^4), \\ R_x u = p \frac{\partial^4 u}{\partial x^4} + 2 \frac{\partial p}{\partial x} \frac{\partial^3 u}{\partial x^3} + \frac{3}{2} \frac{\partial^2 p}{\partial x^2} \frac{\partial^2 u}{\partial x^2} + \frac{1}{2} \frac{\partial^3 p}{\partial x^3} \frac{\partial u}{\partial x} - 2a \frac{\partial^3 u}{\partial x^3}. \quad (3.71)$$

Следуя идее А. А. Самарского, будем искать разностный оператор $L_{1,x}^h$ со следующими аппроксимационными свойствами:

$$L_{1,x}^h u = L_x u - \frac{h_x^2}{12} L_x \left(\frac{1}{p} L_x u \right) + O(h^4). \quad (3.72)$$

Для этого выпишем легко проверяемое соотношение

$$L_x\left(\frac{1}{p}L_x u\right) = \left(-\frac{\partial}{\partial x}p\frac{\partial}{\partial x} + a\frac{\partial}{\partial x}\right)\frac{1}{p}\left(-\frac{\partial}{\partial x}p\frac{\partial u}{\partial x} + a\frac{\partial u}{\partial x}\right) = R_x u + M_x u,$$

$$M_x u = \frac{\partial}{\partial x}\left[\left(\frac{1}{2}\frac{\partial^2 p}{\partial x^2} - \frac{1}{p}\left(\frac{\partial p}{\partial x}\right)^2 - \frac{\partial a}{\partial x}\right)\frac{\partial u}{\partial x}\right] + a\frac{\partial}{\partial x}\left(\frac{a}{p}\frac{\partial u}{\partial x}\right) - \frac{\partial a}{\partial x}\frac{\partial^2 u}{\partial x^2}. \quad (3.73)$$

Из уравнений (3.71) и (3.73) видно, что для обеспечения условия (3.72) необходимо положить

$$L_{1,x}^h u = L_x^h u - \frac{h_x^2}{12} M_x^h u, \quad (3.74)$$

где M_x^h – разностный оператор, определяющий аппроксимацию дифференциального оператора M_x с погрешностью $O(h^2)$:

$$M_x^h u = \frac{1}{h_x^4} \left\{ d_{i+1/2,j}(u_{i+1,j} - u_{i,j}) - d_{i-1/2,j}(u_{i,j} - u_{i-1,j}) \right. \\ \left. + h_x^2 a_{i,j} \left[\left(\frac{a}{p}\right)_{i+1/2,j} (u_{i+1,j} - u_{i,j}) - \left(\frac{a}{p}\right)_{i-1/2,j} (u_{i,j} - u_{i-1,j}) \right] \right. \\ \left. - \frac{h_x}{2} (a_{i+1,j} - a_{i-1,j})(u_{i+1,j} - 2u_{i,j} + u_{i-1,j}) \right\},$$

$$d_{i+1/2,j} = 2(p_{i,j} - 2p_{i+1/2,j} + p_{i+1,j}) - \\ - \frac{(p_{i+1,j} - p_{i,j})^2}{p_{i+1/2,j}} - h_x(a_{i+1,j} - a_{i,j}). \quad (3.75)$$

Получаемый новый одномерный оператор (3.74) является трехточечным и может быть представлен формулой

$$L_{1,x}^h u = \frac{1}{h_x^2} [\check{p}_{i-1/2,j}(u_{i,j} - u_{i-1,j}) + \hat{p}_{i+1/2,j}(u_{i,j} - u_{i+1,j})], \quad (3.76)$$

в которой коэффициенты $\hat{p}_{i+1/2,j}$, $\check{p}_{i-1/2,j}$ имеют вид

$$\hat{p}_{i+1/2,j} = p_{i+1/2,j} - \frac{a_{i,j}}{2} h_x \\ + \frac{1}{12} \left[d_{i+1/2,j} + h_x^2 a_{i,j} \left(\frac{a}{p}\right)_{i+1/2,j} - \frac{h_x}{2} (a_{i+1,j} - a_{i-1,j}) \right],$$

$$\check{p}_{i-1/2,j} = p_{i-1/2,j} + \frac{a_{i,j}}{2} h_x \\ + \frac{1}{12} \left[d_{i-1/2,j} + h_x^2 a_{i,j} \left(\frac{a}{p}\right)_{i-1/2,j} - \frac{h_x}{2} (a_{i+1,j} - a_{i-1,j}) \right] \quad (3.77)$$

Определяя аналогичным образом разностный оператор $L_{1,y}^h$ с коэффициентами $\hat{q}_{i,j+1/2}$, $\check{q}_{i,j-1/2}$, мы можем построить следующую аппроксимацию порядка $O(h^4)$ для уравнения (3.69):

$$\begin{aligned}
L_1^h u &\equiv L_{1,x}^h u + L_{1,y}^h u - \frac{h_x^2}{12} L_{1,x}^h \frac{1}{p} L_{1,y}^h u - \frac{h_y^2}{12} L_{1,y}^h \frac{1}{q} L_{1,x}^h u \\
&= f - \frac{h_x^2}{12} L_x \frac{f}{p} - \frac{h_y^2}{12} L_y \frac{f}{q} + O(h^4).
\end{aligned} \tag{3.78}$$

В правой части этого уравнения дифференциальные операторы можно заменить разностными аналогами, что не изменит общего порядка аппроксимации.

Отметим, что данная система является симметричной, если только в (3.69) отсутствуют конвективные члены ($a = b \equiv 0$), а диффузионные коэффициенты p, q не зависят от x, y соответственно.

В соответствии с локальной нумерацией узлов на рис. 3.2 левая часть уравнения (3.78) может быть записана в виде

$$(L_1^h u)_{i,j} = a_{0,i,j} u_0 - \sum_{k=1}^8 a_{k,i,j} u_k, \tag{3.79}$$

причем коэффициенты разностного оператора L_1^h имеют следующий вид:

$$\begin{aligned}
a_{1,i,j} &= \frac{\check{p}_{i-1/2,j}}{h_x^2} \left(1 - \frac{\check{q}_{i,j}}{6q_{i,j}} - \frac{\check{q}_{i-1,j}}{6p_{i-1,j}} \frac{h_x^2}{h_y^2} \right), \\
a_{3,i,j} &= \frac{\hat{p}_{i+1/2,j}}{h_x^2} \left(1 - \frac{\check{q}_{i,j}}{6q_{i,j}} - \frac{\check{q}_{i-1,j}}{6p_{i-1,j}} \frac{h_x^2}{h_y^2} \right), \\
a_{2,i,j} &= \frac{\check{q}_{i,j-1/2}}{h_y^2} \left(1 - \frac{\check{p}_{i,j}}{6p_{i,j}} - \frac{\check{p}_{i,j-1}}{6q_{i,j-1}} \frac{h_y^2}{h_x^2} \right), \\
a_{4,i,j} &= \frac{\hat{q}_{i,j+1/2}}{h_y^2} \left(1 - \frac{\check{p}_{i,j}}{6p_{i,j}} - \frac{\check{p}_{i,j+1}}{6q_{i,j+1}} \frac{h_y^2}{h_x^2} \right), \\
a_{5,i,j} &= \frac{1}{12} \left(\frac{\check{p}_{i-1/2,j}}{p_{i-1,j}} \frac{\hat{q}_{i-1,j+1/2}}{h_y^2} + \frac{\hat{q}_{i,j+1/2}}{q_{i,j+1}} \frac{\check{p}_{i-1/2,j+1}}{h_x^2} \right), \\
a_{6,i,j} &= \frac{1}{12} \left(\frac{\check{p}_{i-1/2,j}}{p_{i-1,j}} \frac{\check{q}_{i-1,j-1/2}}{h_y^2} + \frac{\check{q}_{i-1/2}}{q_{i,j-1}} \frac{\check{p}_{i-1/2,j-1}}{h_x^2} \right), \\
a_{7,i,j} &= \frac{1}{12} \left(\frac{\hat{p}_{i+1/2,j}}{p_{i+1,j}} \frac{\check{q}_{i+1,j-1/2}}{h_y^2} + \frac{\check{q}_{i,j-1/2}}{q_{i,j-1}} \frac{\hat{p}_{i+1/2,j-1}}{h_x^2} \right), \\
a_{8,i,j} &= \frac{1}{12} \left(\frac{\hat{p}_{i+1/2,j}}{p_{i+1,j}} \frac{\hat{q}_{i+1,j+1/2}}{h_y^2} + \frac{\hat{q}_{i,j+1/2}}{q_{i,j+1}} \frac{\hat{p}_{i+1/2,j+1}}{h_x^2} \right), \\
a_{0,i,j} &= \sum_{k=1}^8 a_{k,i,j} = 2 \left[\frac{\check{p}_{i,j}}{h_x^2} \left(1 - \frac{\check{q}_{i,j}}{6q_{i,j}} \right) + \frac{\check{q}_{i,j}}{h_y^2} \left(1 - \frac{\check{p}_{i,j}}{6p_{i,j}} \right) \right],
\end{aligned} \tag{3.80}$$

$$\tilde{p}_{i,j} = (\hat{p}_{i+1/2,j} + \check{p}_{i-1/2,j})/2; \tilde{q}_{i,j} = (\hat{q}_{i,j+1/2} + \check{q}_{i,j-1/2})/2.$$

При достаточной гладкости коэффициентов исходного дифференциального уравнения мы можем записать

$$\begin{aligned} \hat{p}_{i+1/2,j} &= p_{i+1/2,j} + \frac{h_x}{2} \left(\frac{h_x}{6} \frac{a_{i,j}^2}{p_{i+1/2,j}} - a_{i,j} \right) \\ &\quad + \frac{h_x^2}{12} \left(\frac{1}{2} \frac{\partial^2 p}{\partial x^2} - \frac{1}{p} \left(\frac{\partial p}{\partial x} \right)^2 - \frac{\partial a}{\partial x} \right)_{i,j} + O(h_x^3), \\ \check{p}_{i-1/2,j} &= p_{i-1/2,j} + \frac{h_x}{2} \left(\frac{h_x}{6} \frac{a_{i,j}^2}{p_{i-1/2,j}} + a_{i,j} \right) \\ &\quad + \frac{h_x^2}{12} \left(\frac{1}{2} \frac{\partial^2 p}{\partial x^2} - \frac{1}{p} \left(\frac{\partial p}{\partial x} \right)^2 - \frac{\partial a}{\partial x} \right)_{i,j} + O(h_x^3). \end{aligned} \quad (3.81)$$

Отсюда следует, что неравенства

$$\rho_{i,j} = \min\{\hat{p}_{i+1/2,j}, \check{p}_{i-1/2,j}, \hat{q}_{i,j+1/2}, \check{q}_{i,j-1/2}\} \geq \rho > 0, \quad (3.82)$$

необходимые для того, чтобы уравнения (3.78) были положительного типа, могут в принципе нарушаться при определенных соотношениях коэффициентов p, q, a, b и шагов сетки h_x, h_y . Однако из (3.81) очевидно, что при преобладающих конвективных членах условия удовлетворения (3.82) более “мягкие чем те, которые требуются для обеспечения положительного типа “стандартного” разностного оператора L_x^h из (3.70).

Действительно, выражение для $\hat{p}_{i+1/2,j}$, например, можно переписать в виде

$$\hat{p}_{i+1/2,j} = \frac{1}{12p_{i+1/2,j}} \left[12p_{i+1/2,j}^2 - 6(a_{i,j} - \varphi_{i,j})h_x p_{i+1/2,j} + h_x^2 \left(a^2 - \left(\frac{\partial p}{\partial x} \right)^2 \right)_{i,j} \right],$$

$$\text{где } \varphi_{i,j} = \frac{h_x}{6} \left(\frac{1}{2} \frac{\partial^2 p}{\partial x^2} - \frac{\partial a}{\partial x} \right)_{i,j} + O(h_x^2).$$

Выписывая аналогичные выражения для $\check{p}_{i-1/2,j}$, $\hat{q}_{i,j+1/2}$, $\check{q}_{i,j-1/2}$ и вводя обозначения $p_0 = \min_{x,y}\{p(x,y)\}$, $q_0 = \min_{x,y}\{q(x,y)\}$, мы можем сформулировать следующий результат.

Лемма 3.1. *Величины $\hat{p}_{i+1/2,j}$, $\check{p}_{i-1/2,j}$, $\hat{q}_{i,j+1/2}$, $\check{q}_{i,j-1/2}$ при выполнении условий*

$$\begin{aligned}
& 12p_0^2 - \left| \pm 6a_{i,j} - h_x \left(\frac{1}{2} \frac{\partial^2 p}{\partial x^2} - \frac{\partial a}{\partial x} \right) \right|_{i,j} h_x p_0 + h_x^2 \left[a^2 - \left(\frac{\partial p}{\partial x} \right)^2 \right]_{i,j} \\
& \quad + O(h_x^3) > 0, \\
& 12q_0^2 - \left| \pm 6b_{i,j} - h_y \left(\frac{1}{2} \frac{\partial^2 q}{\partial y^2} - \frac{\partial b}{\partial y} \right) \right|_{i,j} h_y q_0 + h_y^2 \left[b^2 - \left(\frac{\partial q}{\partial y} \right)^2 \right]_{i,j} \\
& \quad + O(h_y^3) > 0
\end{aligned}$$

являются положительными.

Из леммы 3.1 следует, в частности, такой важный результат, что если все диффузионные и конвективные коэффициенты постоянны, то значения $\hat{p}, \check{p}, \hat{q}, \check{q}$ будут положительны при любых соотношениях величин p, q, a, b, h_x, h_y .

Далее, требование положительности коэффициентов a_1, a_2, a_3, a_4 из (3.80), как нетрудно видеть, при достаточно малых значениях $h = \max\{h_x, h_y\}$ имеет форму, близкую к условиям положительности коэффициентов схемы Микеладзе для уравнения Пуассона:

$$\frac{1}{5} + O(h) < \frac{p_{i,j} h_y^2}{q_{i,j} h_x^2} < 5 + O(h). \quad (3.83)$$

Представляет значительный интерес вопрос, может ли система уравнений (3.79) при условии (3.81) быть монотонной, если не выполняются условия (3.83). Ответ на этот вопрос оказывается положительным, но доказательство этого факта мы здесь выпускаем.

Если вместо $\hat{p}_{i+1/2,j}, \check{p}_{i-1/2,j}$ в (3.76) использовать какую-нибудь другую дискретизацию (шаблонный функционал, в терминологии А. А. Самарского), обладающую свойством

$$\hat{p}_{i+1/2,j}^{(k)} = \hat{p}_{i+1/2,j} + h_x^2 \varphi_k(x_{i+1/2,j}) + O(h^4), \quad (3.84)$$

где $\varphi_k(x)$ – некоторая гладкая функция, то аналогичным образом можно получать вместо $L_{1,x}^h$ и другие схемы повышенного порядка, применяя очевидное соотношение

$$\begin{aligned}
(\bar{L}_{k,x}^h u)_{i,j} & \equiv \frac{-1}{h_x^2} [\hat{p}_{i+1/2,j}^{(k)} (u_{i+1,j} - u_{i,j}) - \check{p}_{i-1/2,j}^{(k)} (u_{i,j} - u_{i-1,j})] \\
(\bar{L}_{k,x}^h u)_{i,j} & = (L_{1,x}^h u)_{i,j} - \frac{h_x^2}{12} (M_{k,x} u)_{i,j} + O(h^4),
\end{aligned}$$

$$M_{k,x}u = \frac{\partial}{\partial x} \left(\varphi_k \frac{\partial u}{\partial x} \right). \quad (3.85)$$

Действительно, определяя через $M_{k,x}^h u$ естественную разностную аппроксимацию $M_{k,x}u$, с помощью соотношений (3.71), (3.74), (3.85) легко убедиться, что следующий разностный оператор удовлетворяет условию (3.72):

$$L_{k,x}^h u \equiv \bar{L}_{k,x}^h u + M_{k,x}^h u - M_x^h = L_{1,x}u - \frac{h_x^2}{12}(R_x + M_x)u + O(h^4). \quad (3.86)$$

А. А. Самарский для чисто диффузионной задачи ($a = b = 0$ в (3.69)) предложил для построения схемы четвертого порядка использовать шаблонный функционал вида

$$\begin{aligned} p_{i\pm 1/2,j}^{(2)} &= 6(p_{i,j}^{-1} + 4p_{i\pm 1/2,j}^{-1} + p_{i\pm 1,j}^{-1}) \\ &= p_{i\pm 1/2,j} + \frac{h_x^2}{12} \left[\frac{1}{2} \frac{\partial^2 p}{\partial x^2} + \frac{1}{p} \left(\frac{\partial p}{\partial x} \right)^2 \right]_{i\pm 1/2,j} + O(h^4). \end{aligned} \quad (3.87)$$

Из последнего разложения и (3.71) легко следует, что соответствующий разностный оператор $L_{2,x}^h = \bar{L}_{2,x}^h$ действительно удовлетворяет условию (3.72).

Рассмотрим еще один параметризованный способ дискретизации функции $p(x, y)$:

$$\begin{aligned} \tilde{p}_{i\pm 1/2,j}^{(3)} &= \frac{2}{\frac{\theta}{p_{i,j}} + \frac{4(1-\theta)}{p_{i,j} + p_{i\pm 1,j}} + \frac{\theta}{p_{i\pm 1,j}}} \\ &= p_{i\pm 1/2,j} + \frac{h_x^2}{4} \left[\frac{1}{2} \frac{\partial^2 p}{\partial x^2} - \frac{\theta}{p} \left(\frac{\partial p}{\partial x} \right)^2 \right]_{i\pm 1/2,j} + O(h^4). \end{aligned}$$

Соответствующий разностный оператор, удовлетворяющий условию (3.72), например, при $\theta = \frac{1}{3}$ имеет вид

$$\begin{aligned} (L_{3,x}^h u)_{i,j} &= \frac{-1}{h_x^2} \left[p_{i+1/2,j}^{(3)}(u_{i+1,j} - u_{i,j}) - p_{i-1/2,j}^{(3)}(u_{i,j} - u_{i-1,j}) \right], \\ p_{i\pm 1/2,j}^{(3)} &= 6 \left[p_{i,j}^{-1} + 8(p_{i,j} + p_{i\pm 1,j})^{-1} + p_{i\pm 1,j}^{-1} \right]^{-1} \\ &\quad - \frac{1}{3} (p_{i\pm 1,j} - 2p_{i\pm 1/2,j} + p_{i,j}). \end{aligned} \quad (3.88)$$

Очевидно, что любая линейная комбинация операторов, удовлетворяющих условию (3.72), также обладает этим свойством.

Однако нетрудно убедиться, что на основе применения $L_{1,x}^h, L_{2,x}^h, L_{3,x}^h$ невозможно получить разностную схему без использования значений $p_{i\pm 1/2,j}$.

Данный вопрос имеет практическое значение в задачах, где функции p, q не заданы явно в аналитическом виде, а вычисляются, например, через значения самого решения или других сеточных функций, которые известны только в узлах.

Рассмотренный принцип аппроксимации легко переносится и на трехмерные уравнения более общего вида

$$Lu + \varkappa u \equiv -\frac{\partial}{\partial x} p \frac{\partial u}{\partial x} - \frac{\partial}{\partial y} q \frac{\partial u}{\partial y} - \frac{\partial}{\partial z} r \frac{\partial u}{\partial z} + a \frac{\partial u}{\partial x} + b \frac{\partial u}{\partial y} + c \frac{\partial u}{\partial z} + \varkappa u = f. \quad (3.89)$$

В этом случае четвертый порядок обеспечивается схемами

$$\begin{aligned} L^h u_h + (\varkappa u)_h &\equiv L_{k,x}^h u_h + L_{k,y}^h u_h + L_{k,z}^h u_h - \frac{h_x^2}{12} L_{k,x}^h \frac{1}{p} (L_{k,y}^h + L_{k,z}^h) u_h \\ &- \frac{h_y^2}{12} L_{k,y}^h \frac{1}{q} (L_{k,x}^h + L_{k,z}^h) u_h - \frac{h_z^2}{12} L_{k,z}^h \frac{1}{r} (L_{k,x}^h + L_{k,y}^h) u_h + \varkappa u_h \\ &= f - \left(\frac{h_x^2}{12} L_{k,x}^h \frac{1}{p} + \frac{h_y^2}{12} L_{k,y}^h \frac{1}{q} + \frac{h_z^2}{12} L_{k,z}^h \frac{1}{r} \right) (f - \varkappa u_h). \end{aligned} \quad (3.90)$$

Здесь $L_{k,x}^h$ — один из одномерных разностных операторов, описываемых формулами (3.74), (3.84)–(3.87), а одномерные операторы $L_{k,y}^h, L_{k,z}^h$ определяются аналогично.

Как видно из последнего члена в (3.90), данная система уравнений будет симметрична только при постоянном "гельмгольцовском" коэффициенте $\varkappa$ (даже если диффузионные коэффициенты p, q, r — постоянные, а конвективные a, b, c — нулевые).

Вопрос о монотонности разностной схемы (3.90) требует дополнительных исследований. Для простейшего случая чисто диффузионной задачи с постоянными коэффициентами p, q, r и $\varkappa = a = b = c = 0$ легко показать, что данная система уравнений будет положительного типа при выполнении следующих неравенств:

$$\frac{1}{3} \leq \frac{qh_x^2}{ph_y^2}, \quad \frac{rh_x^2}{ph_z^2}, \quad \frac{rh_y^2}{qh_z^2} \leq 3. \quad (3.91)$$

3.4.2. Уравнение Пуассона в цилиндрической системе координат

Пусть u удовлетворяет дифференциальному уравнению

$$L_r u + L_z u \equiv -\frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial u}{\partial r} - \frac{\partial^2 u}{\partial z^2} = f(r, z) \quad (3.92)$$

в области $(r_0, r_{I+1}) \times (z_0, z_{J+1})$, $r_0 \geq 0$, и граничному условию $\frac{\partial u}{\partial r} = 0$ на оси симметрии при $r_0 = 0$. Формально уравнение (3.92) (после умножения на r) сводится к виду (3.69), однако мы остановимся на этом случае подробнее ввиду его самостоятельной прикладной значимости.

Рассмотрим сначала погрешность аппроксимации стандартного "радиального" разностного оператора при $r > 0$ (см. п. 3.1.2):

$$\begin{aligned} (L_r^h u)_i &= \frac{1}{r_i h_r^2} \left[r_{i+1/2} (u_i - u_{i+1}) + r_{i-1/2} (u_i - u_{i-1}) \right] \\ &= \frac{-1}{r_i} \left[u' + r u'' + \frac{h_r^2}{6} u''' + r \frac{h_r^2}{12} u^{(4)} \right]_i + O(h^4), \quad r_i = i h_r, \end{aligned}$$

которая имеет, как видно, порядок $O(h^2)$.

Выпишем также выражение для квадрата соответствующего дифференциального оператора:

$$L_r^2 u = \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} \left(\frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial u}{\partial r} \right) = \frac{1}{r} \left(2 \frac{\partial^3 u}{\partial r^3} - \frac{\partial}{\partial r} \frac{1}{r} \frac{\partial u}{\partial r} + r \frac{\partial^4 u}{\partial r^4} \right).$$

Из сопоставления двух последних соотношений легко обнаружить, что разностный оператор

$$\begin{aligned} (L_{1,r}^h u)_i &\equiv \frac{1}{h_r^2 r_i} \left[r_{i+1/2}^{(1)} (u_i - u_{i+1}) + r_{i-1/2}^{(1)} (u_i - u_{i-1}) \right], \\ r_{i\pm 1/2}^{(1)} &= r_{i\pm 1/2} - \frac{h_r^2}{12 r_{i\pm 1/2}}, \end{aligned} \quad (3.93)$$

следующим образом аппроксимирует при $r > 0$ дифференциальный "радиальный" оператор:

$$L_{1,r}^h u = L_r u - \frac{h_r^2}{12} L_r^2 u + O(h^4).$$

Таким образом, разностное уравнение

$$\begin{aligned} L_1^h u &\equiv \left(L_{1,r}^h + L_z^h - \frac{h_r^2 + h_z^2}{12} L_{1,r}^h L_z^h \right) u_h = f - \\ &\quad - \frac{h_r^2}{12} L_r^h f - \frac{h_z^2}{12} L_z^h f \end{aligned} \quad (3.94)$$

имеет погрешность аппроксимации четвертого порядка. Эта алгебраическая система несимметрична, но легко симметризуется после умножения каждого уравнения на величину r_i (другими словами, после умножения обеих частей уравнения (3.94) на соответствующую диагональную матрицу). Коэффициенты получаемой системы имеют следующий вид:

$$\begin{aligned} a_1 &= \left(\frac{r_{i-1/2}}{h_r^2} - \frac{1}{12r_{i-1/2}} \right) \left(1 - \frac{h_r^2 + h_z^2}{6h_z^2} \right), \\ a_3 &= \left(\frac{r_{i+1/2}}{h_r^2} - \frac{1}{12r_{i+1/2}} \right) \left(1 - \frac{h_r^2 + h_z^2}{6h_z^2} \right), \\ a_2 = a_4 &= \left[1 - \frac{h_r^2 + h_z^2}{12h_r^2} \left(2 - \frac{h_r^2}{6r_{i+1/2}r_{i-1/2}} \right) \right] \frac{r_i}{h_z^2}, \\ a_5 = a_6 &= \frac{h_r^2 + h_z^2}{12h_z^2} \left(\frac{r_{i-1/2}}{h_r^2} - \frac{1}{12r_{i-1/2}} \right), \\ a_7 = a_8 &= \frac{h_r^2 + h_z^2}{12h_z^2} \left(\frac{r_{i+1/2}}{h_r^2} - \frac{1}{12r_{i+1/2}} \right). \end{aligned} \quad (3.95)$$

Отсюда видно, что при $r_{i-1/2} > h_r$ условия положительности схемы, т. е. $a_k \geq 0$ в (3.95) те же, что и для схемы Микеладзе для уравнения Пуассона в декартовой системе координат.

Рассмотренный подход неприменим для построения разностных уравнений на оси симметрии, и при $r_i = 0$ необходимо использовать специальные представления. В частности, из (3.92) непосредственно следует, что при $r \rightarrow 0$ уравнение Пуассона принимает форму

$$L_{r,0}u - \frac{\partial^2 u}{\partial z^2} = f(0, z), \quad L_{r,0} = -2 \frac{\partial^2 u}{\partial r^2}. \quad (3.96)$$

Принимая во внимание свойство симметрии решения рассматриваемой задачи, мы определяем на оси разностный оператор

$$(L_{r,0}^h u)_{0,j} \equiv 4 \frac{u_{0,j} - u_{1,j}}{h_r^2} = (L_{r,0}u)_{0,j} - \frac{h_r^2}{16} (L_{r,0}^2 u)_{0,j} + O(h^4), \quad (3.97)$$

в котором используется свойство

$$L_{r,0}^2 u = \lim_{r \rightarrow 0} (L_r^2 u) = \frac{8}{3} \frac{\partial^4 u}{\partial r^4} \Big|_{r=0}.$$

Отсюда разностная аппроксимация четвертого порядка на оси определяется в виде

$$\begin{aligned}
(L^h u)_{0,j} &\equiv \left(L_{r,0}^h u + L_z^h u - \frac{3h_r^2 + 4h_z^2}{48} L_{r,0}^h L_z^h u \right)_{0,j} \\
&= \left(f + \frac{h_r^2}{16} L_{r,0}^h f + \frac{h_z^2}{12} L_z^h f \right)_{0,j} + O(h^4). \quad (3.98)
\end{aligned}$$

В покомпонентном представлении этого уравнения величины a_1 , a_5 , a_6 равны нулю, а остальные коэффициенты выражаются формулами

$$\begin{aligned}
a_2 = a_4 &= \frac{1}{h_z^2} \left(1 - \frac{3h_r^2 + 4h_z^2}{24h_r^2} \right), & a_3 &= \frac{4}{h_r^2} \left(1 - \frac{3h_r^2 + 4h_z^2}{12h_z^2} \right), \\
a_7 = a_8 &= \frac{3h_r^2 + 4h_z^2}{24h_r^2 h_z^2}, & a_0 &= \sum_{k=1}^8 a_k. \quad (3.99)
\end{aligned}$$

Сравнение величин a_1 , a_5 , a_6 из (3.95) для $i = 1$ со значениями a_3 , a_7 , a_8 из (3.99) соответственно показывает, что алгебраическая система, получаемая из (3.97), (3.98), является не только несимметричной, но и несимметризуемой.

Отметим, что на основе приведенных выше подходов можно строить и другие аппроксимации порядка $O(h^4)$ для уравнения (3.92). Если в (3.93) вместо $r_{i\pm 1/2}^{(1)}$ использовать величины

$$r_{i\pm 1/2}^{(k)} = r_{i\pm 1/2}^{(1)} + h_r^2 \varphi_{i\pm 1/2}^{(k)} + O(h^4),$$

где $\varphi^{(k)}$ и остаточный член – достаточно гладкие функции r , то при этом в (3.94) достаточно заменить “радиальный” оператор $L_{1,r}^h$ на оператор $L_{k,r}^h$ следующего вида:

$$\begin{aligned}
L_{k,r}^h u &= \bar{L}_{k,r}^h u + h_r^2 M_{k,r}^h u, \\
(\bar{L}_{k,r}^h u)_{i,j} &= \frac{1}{r_i h_r^2} \left[r_{i+1/2}^{(k)} (u_{i,j} - u_{i+1,j}) + r_{i-1/2}^{(k)} (u_{i,j} - u_{i-1,j}) \right],
\end{aligned}$$

где $M_{k,r}^h u$ – некоторое определяемое аналогично $M_{k,x}^h u$ из (3.85), (3.86) разностное выражение, аппроксимирующее с погрешностью $O(h_r^2)$ величину $M_{k,r} u = \frac{1}{r} \frac{\partial}{\partial r} \varphi^{(k)} \frac{\partial u}{\partial r}$. Получаемый при этом разностный оператор L_k^h (вместо L_1^h в (3.94)) имеет ошибку аппроксимации четвертого порядка в силу легко проверяемого соотношения

$$L_{k,r}^h u = L_{1,r}^h u + O(h_r^4) = L_r u - \frac{h_r^2}{12} L_r^2 u + O(h_r^4).$$

3.4.3. Аппроксимации систем дифференциальных уравнений

Рассмотренный выше подход формально легко переносится и на системы дифференциальных уравнений вида

$$Lu \equiv \sum_{i=1}^n b_i L_i u + su = f, \quad (3.100)$$

где $u = \{u_1, \dots, u_m\}$ – вектор-функция от независимых переменных $x = \{x_1, \dots, x_n\}$, L_i – дифференциальный оператор второго порядка

$$L_i u = - \sum_{j=1}^n \frac{\partial}{\partial x_i} a_{i,j} \frac{\partial u}{\partial x_j} + c_i \frac{\partial u}{\partial x_i}, \quad i = 1, \dots, n, \quad (3.101)$$

а b_i , $a_{i,j}$, c_i , s – матрицы порядка m с переменными коэффициентами, удовлетворяющими условиям существования единственного и достаточного гладкого решения рассматриваемой задачи (для простоты расчетную область будем считать n -мерным параллелепипедом с граничными условиями 1-го рода).

Предположим, что на равномерной сетке с шагом h_i по каждой из переменных существуют разностные операторы со следующими аппроксимационными свойствами:

$$L_i^h u = L_i u - h_i^2 d_i L_i g_i^{-1} L_i u + O(h^4), \quad i = 1, \dots, n, \quad (3.102)$$

где d_i и g_i – некоторые матрицы, причем g_i – невырожденные. Данная запись является формальным обобщением скалярного случая, и с формулами (3.90) для $n = 2$ имеется соответствие: $d_1 = d_2 = \dots = d_n = \frac{1}{12}$, $g_1 = p$, $g_2 = q$.

Из (3.100) и (3.102) при невырожденности матриц b_i можно выписать следующие выражения для слагаемых дифференциальных операторов:

$$L_i u = L_i^h u + h_i^2 d_i L_i g_i^{-1} L_i u + O(h^4) = b_i \left(f - su - \sum_{k \neq i} b_k L_k u \right). \quad (3.103)$$

Отсюда можно получить соотношения

$$d_i L_i g_i^{-1} L_i u = d_i L_i g_i^{-1} b_i^{-1} \left(f - su - \sum_{k \neq i} b_k L_k u \right). \quad (3.104)$$

$$\sum_{i=1}^n b_i (L_i^h u + h_i^2 d_i L_i g_i^{-1} L_i u) + su = f. \quad (3.105)$$

Подставляя теперь (3.104) в (3.105) и заменяя далее дифференциальные операторы на разностные, получаем разностную схему

$$\begin{aligned} L^h u &\equiv \sum_{i=1}^n b_i \left[L_i^h u - h_i^2 d_i L_i g_i^{-1} b_i^{-1} \left(su + \sum_{k \neq i} b_k L_k^h u \right) \right] \\ &= f - \sum_{i=1}^n h_i^2 b_i d_i L_i g_i^{-1} b_i^{-1} f + O(h^4). \end{aligned}$$

Таким образом, мы вывели аппроксимацию системы уравнений (3.100) с погрешностью четвертого порядка, если только справедливы все проведенные выше преобразования. В частности, исходные коэффициенты и искомое решение должны удовлетворять условиям, при которых остаточный член в (3.102) действительно имеет четвертый порядок. Кроме того, проблема нахождения матричных коэффициентов d_i и g_i может оказаться самостоятельной сложной задачей. Так что рассмотренный формализм, можно сказать, указывает общее направление исследований для конкретных случаев и только. Например, при самосопряженности исходной дифференциальной системы могут интересовать вопросы симметричности, положительной определенности и/или монотонности получаемых разностных схем.

3.4.4. Аппроксимация граничных условий 2-го и 3-го родов

Рассмотренные компактные разностные схемы легко реализуются при решении краевых задач Дирихле в областях, составленных из прямоугольников, стороны которых параллельны координатным осям. В этом случае для внутренних узлов выписываются рассмотренные выше девятиточечные уравнения, и при этом в лежащих на границе узлах сетки используются заданные значения искомого решения, без внесения дополнительных погрешностей.

Если же заданы граничные условия Неймана или Ньютона, то их необходимо аппроксимировать (на решении краевой задачи) с повышенным порядком точности, используя при этом тот факт, что решаемые дифференциальные уравнения справедливы вплоть до границы расчетной области.

Мы продемонстрируем такой подход на решении диффузионного уравнения без конвективных членов

$$L_x u + L_y u \equiv -\frac{\partial}{\partial x} p \frac{\partial u}{\partial x} - \frac{\partial}{\partial y} q \frac{\partial u}{\partial y} = f(x, y), \quad (3.106)$$

решение которого на левой границе $x = x_0$ удовлетворяет условию

$$-p \frac{\partial u}{\partial x} + \varkappa u = g, \quad (3.107)$$

где $p > 0$, $\varkappa$ и g – достаточно гладкие функции.

Для этого, вводя аналогичные (3.77) обозначения

$$\begin{aligned} \hat{p}_{i+1/2,j} &= p_{i+1/2,j} + \frac{h_x^2}{24} \varphi_{i+1/2,j} + O(h^4), & (\nabla_x^h u)_{i,j} &= \frac{u_{i+1,j} - u_{i,j}}{h_x}, \\ \varphi_{i+1/2,j} &= \frac{\partial^2 p}{\partial x^2} - \frac{2}{p} \left(\frac{\partial p}{\partial x} \right)^2, \end{aligned}$$

выпишем с помощью непосредственных тейлоровских разложений следующие вспомогательные соотношения:

$$\begin{aligned} -\hat{p}_{1/2,j} \nabla_x^h u_{0,j} &= - \left[\left(p + \frac{h_x^2}{24} \varphi \right) \left(\frac{\partial u}{\partial x} + \frac{h_x^2}{24} \frac{\partial^3 u}{\partial x^3} \right) \right]_{1/2,j} + O(h^4) \\ &= - \left[p \frac{\partial u}{\partial x} + \frac{h_x}{2} \frac{\partial}{\partial x} p \frac{\partial u}{\partial x} + \frac{h_x^2}{24} \left(3 \frac{\partial^2}{\partial x^2} p \frac{\partial u}{\partial x} + p \frac{\partial^3 u}{\partial x^3} + \varphi \frac{\partial u}{\partial x} \right) \right. \\ &\quad \left. + \frac{h_x^3}{48} \frac{\partial}{\partial x} \left(\frac{\partial^2}{\partial x^2} p \frac{\partial u}{\partial x} + p \frac{\partial^3 u}{\partial x^3} + \varphi \frac{\partial u}{\partial x} \right) \right]_{0,j} + O(h^4). \end{aligned} \quad (3.108)$$

Подставляя теперь сюда граничное условие, выражение для φ и проводя преобразования вида

$$\begin{aligned} \frac{\partial}{\partial x} L_x u &= -\frac{\partial^2}{\partial x^2} p \frac{\partial u}{\partial x} = -p \frac{\partial^3 u}{\partial x^3} - 2 \frac{\partial p}{\partial x} \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 p}{\partial x^2} \frac{\partial u}{\partial x}, \\ L_y u &= L_{1,y}^h u + \frac{h_y^2}{12} L_y \frac{1}{q} L_y u + O(h^4), \\ \frac{\partial}{\partial x} L_y u &= \nabla_x^h L_{1,y} u - \frac{h_x}{2} \frac{\partial^2}{\partial x^2} L_y u + O(h^2), \end{aligned}$$

с помощью исходного дифференциального уравнения приходим к следующей форме:

$$\begin{aligned}
-\hat{p}_{1/2,j} \nabla_x^h u_{0,j} &= \hat{f}_{0,j} - \left\{ \varkappa u + \frac{h_x}{2} L_{1,y}^h u + \frac{h_x^2}{6} \left(\nabla_x^h L_{1,y}^h u - \frac{1}{2p} \frac{\partial p}{\partial x} L_{1,y}^h u \right) \right. \\
&+ \frac{h_x}{24} \left[h_y^2 L_y \frac{1}{q} L_y u + h_x^2 \frac{\partial^2}{\partial x^2} \left(2L_y u + L_x u \right) - h_x^2 \frac{\partial}{\partial x} \left(\frac{1}{p} \frac{\partial p}{\partial x} L_y u \right) \right] \left. \right\}_{0,j} \\
&+ O(h^4), \tag{3.109}
\end{aligned}$$

где $L_{1,y}^h u$ определяется аналогично (3.76) и использовано обозначение

$$\hat{f} \equiv g + \frac{h_x}{2} f + \frac{h_x^2}{6} \frac{\partial f}{\partial x} - \frac{h_x^2}{12} \frac{\partial p}{\partial x} \frac{f}{p} + \frac{h_x^3}{24} \frac{\partial}{\partial x} \left(\frac{f}{p} \frac{\partial p}{\partial x} \right).$$

Прибавляя к обеим частям равенства (3.109) величину $\frac{h_y^2}{12} L_{1,y}^h \frac{p}{q} \nabla_x^h u$ и продолжая аналогичные преобразования, приходим к соотношению

$$\begin{aligned}
&-\hat{p}_{1/2,j} \nabla_x^h u_{0,j} + \left[\varkappa u + \frac{h_x}{2} L_{1,y}^h u + \frac{1}{12} \left(h_x^2 \nabla_x^h L_{1,y}^h u + h_y^2 L_{1,y}^h \frac{p}{q} \nabla_x^h u \right) \right. \\
&- \frac{h_x^2}{12p} \frac{\partial p}{\partial x} L_{1,y}^h u + \frac{h_x^2}{24} \nabla_x^h \left(\frac{1}{p} \frac{\partial p}{\partial x} L_{1,y}^h u \right) \\
&- \frac{h_x^2}{12} \left(L_{1,y}^h \varkappa u - \frac{\partial^2 q}{\partial x \partial y} \bar{\nabla}_y^h u - \frac{\partial q}{\partial x} \Delta_y^h u \right) + \frac{h_y^2}{12} L_{1,y}^h \left(\left(\frac{p}{q} - \frac{h_x}{2q} \frac{\partial p}{\partial x} \right) \varkappa u \right) \left. \right]_{0,j} \\
&= \hat{f}_{0,j} + \left[\frac{1}{12} L_{1,y}^h \left(\left(h_y^2 \frac{p}{q} - h_x^2 \right) g \right) - \frac{h_x h_y^2}{24} L_{1,y}^h \left(\frac{\partial p}{\partial x} \frac{g}{q} \right) \right. \\
&\quad \left. + \frac{h_x}{24} \left(h_x^2 \frac{\partial^2 f}{\partial x^2} - h_y^2 L_y \frac{f}{q} \right) \right]_{0,j} + O(h^4). \tag{3.110}
\end{aligned}$$

Здесь $\bar{\nabla}_y^h$ и Δ_y^h обозначают центральную разность первого порядка и вторую разность по переменной y , введенные выше в (3.6) и (3.39).

Таким образом, если в (3.110) отбросить остаточный член $O(h^4)$, то мы получаем шеститочечное разностное уравнение, аппроксимирующее на линии границы $x = x_0$ краевое условие (3.107) (на решении уравнения (3.106)) с погрешностью четвертого порядка.

В частности, для простейшего случая $p = q \equiv 1$, $\varkappa \equiv 0$ (граничное условие Неймана для уравнения Пуассона) разностное уравнение (3.110) принимает простую форму:

$$\begin{aligned}
-\nabla_x^h u - \frac{h_x}{2} \Delta_y^h u - \frac{h_x^2 + h_y^2}{12} \nabla_x^h \Delta_y^h u &= g + \frac{h_x}{2} f + \frac{h_x^2}{6} \frac{\partial f}{\partial x} \\
- \frac{h_y^2 - h_x^2}{12} \Delta_y^h g + \frac{h_x}{24} \left(h_x^2 \frac{\partial^2 f}{\partial x^2} + h_y^2 \frac{\partial^2 f}{\partial y^2} \right) &+ O(h^4),
\end{aligned}$$

рассмотренную в [81]. Там же рассмотрены аналогичные аппроксимации четвертого порядка для условия Неймана на других сторонах прямоугольной расчетной области, а также в ее угловых точках.

Мы на таких подробностях останавливаться не будем и отметим только, что в приведенных разностных выражениях четвертый порядок сохранится, если участвующие в них производные от функций p, q, f заменить соответствующими конечными разностями.

3.5. Метод неопределенных коэффициентов

Наиболее общий подход к построению разностных схем заключается в *методе неопределенных коэффициентов*, который на примере трехточечной аппроксимации первой производной может быть проиллюстрирован следующим образом.

Будем искать искомую разностную формулу в виде

$$\hat{\nabla}_h u_i = c_{-1} u_{i-1} + c_0 u_i + c_1 u_{i+1},$$

где неопределенные коэффициенты c_{-1}, c_0, c_1 находятся из условия совпадения данного выражения (после подстановки в него тейлоровских разложений для $u_{i\pm 1} = u(x_{i\pm 1})$) с аппроксимируемой первой производной – с максимально возможным порядком малости по h .

Действуя по этому принципу, из получаемого соотношения

$$\begin{aligned}
\tilde{\nabla}_h u_i &\equiv (c_{-1} + c_0 + c_1) u_i + c_{-1} \left[-h_{i-1} u'_i + \frac{h_{i-1}^2}{2} u''_i - \frac{h_{i-1}^3}{6} u'''_i + O(h^4) \right] \\
&+ c_1 \left[h_i u'_i + \frac{h_i^2}{2} u''_i + \frac{h_i^3}{6} u'''_i + O(h^4) \right]
\end{aligned}$$

выписываем требуемые условия для коэффициентов при производных до третьего порядка включительно:

$$\begin{aligned} c_{-1} + c_0 + c_1 &= 0, & h_i c_1 - h_{i-1} c_{-1} &= 1, \\ h_i^2 c_1 + h_{i-1}^2 c_{-1} &= 0, & h_i^3 c_1 - h_{i-1}^3 c_{-1} &= 0. \end{aligned} \quad (3.111)$$

Отсюда подсистема из первых трех уравнений дает однозначное определение коэффициентов

$$c_{-1} = \frac{-h_i}{h_{i-1}(h_{i-1} + h_i)}, \quad c_1 = \frac{1 + c_{-1}h_{i-1}}{h_i}, \quad c_0 = -c_1 - c_{-1}.$$

Четвертое уравнение оказывается несовместным при любых соотношениях h_{i-1} , h_i , и в силу $c_k = O(h^{-1})$, $k = -1, 0, 1$, мы окончательно имеем

$$\tilde{\nabla} u_i = u'_i - \frac{1}{6}(c_1 h_i^3 - c_{-1} h_{i-1}^3) u_i''' + O(h^3) = u'_i + \frac{h_i h_{i-1}}{6} u_i''' + O(h^3).$$

Этот результат согласуется с предыдущими выводами и может быть сформулирован следующим образом.

Теорема 3.1. *Трехточечное разностное выражение (3.7) аппроксимирует первую производную $u'(x_i)$ для $u \in C^3[a, b]$ с погрешностью $O(h^2)$ на произвольной неравномерной сетке. При любом соотношении шагов и $u(x) \in C^4[a, b]$ схемы данного вида с погрешностью $O(h^3)$ не существует. На равномерной сетке выражение (3.7) переходит в центральную разность (3.6). Если же любой из c_k положить равным нулю, то оставшиеся два коэффициента из первых двух уравнений (3.111) определяются однозначно как равные по модулю и противоположные по знаку, а погрешность аппроксимации в точке x_i есть $O(h)$ (если только c_0 и $h_i - h_{i-1}$ одновременно не равны нулю).*

Методом неопределенных коэффициентов можно аналогично показать, что с увеличением числа точек в разностном выражении находятся и аппроксимации первой производной более высоких порядков.

Продemonстрируем идею метода на аппроксимации уравнения для функции двух переменных.

Пусть, например, дифференциальный оператор имеет вид

$$Lu = a_{11} \frac{\partial^2 u}{\partial x^2} + a_{12} \frac{\partial^2 u}{\partial x \partial y} + a_{22} \frac{\partial^2 u}{\partial y^2} + a_{10} \frac{\partial u}{\partial x} + a_{01} \frac{\partial u}{\partial y} + a_{00} u. \quad (3.112)$$

Будем искать аппроксимацию L_h в точке P на совокупности узлов (шаблоне) из $r + 1$ точек Q_{kp} в форме

$$(L_h u)_P = a_0 u(P) - \sum_{k=1}^r a_k u(Q_{kp}), \quad (3.113)$$

где коэффициенты $a_0, a_1, \dots, a_r$ пока не определены. Разлагая $u(Q_{kp})$ в окрестности точки P в ряд Тейлора до третьего порядка и группируя члены разложения, можно получить

$$\begin{aligned} (L_h u)_P = & A_{11} \frac{\partial^2 u(P)}{\partial x^2} + A_{12} \frac{\partial^2 u(P)}{\partial x \partial y} + A_{22} \frac{\partial^2 u(P)}{\partial y^2} + A_1 \frac{\partial u(P)}{\partial x} \\ & + A_2 \frac{\partial u(P)}{\partial y} + A_0 u(P) + O(h), \end{aligned} \quad (3.114)$$

где $A_{11}, A_{12}, \dots$ суть линейные комбинации от $a_0, a_1, \dots, a_r$ с коэффициентами, выражающимися через расстояния от точки P до точек $Q_{kp} = (x_p + h_{xk}, y_p + h_{yk})$. Приравнявая коэффициенты при одинаковых производных в выражениях (3.112) и (3.114), получаем систему шести уравнений относительно $r+1$ неизвестных $a_0, a_1, \dots, a_r$:

$$\begin{aligned} A_{11} = \sum_{k=1}^r h_{xk}^2 a_k = a_{11}, \quad A_1 = \sum_{k=1}^r h_{1k} a_k = a_{10}, \\ A_{12} = \sum_{k=1}^r h_{xk} h_{yk} a_k = a_{12}, \quad A_2 = \sum_{k=1}^r h_{yk} a_k = a_{01}, \\ A_{22} = \sum_{k=1}^r h_{yk}^2 a_k = a_{22}, \quad A_0 = a_0 - \sum_{k=1}^r a_k = a_{00}. \end{aligned} \quad (3.115)$$

В зависимости от выбранного шаблона эта система может быть определена, переопределена или недоопределена. Если система разрешима, то мы находим коэффициенты выражения (3.113), которое в этом случае будет аппроксимировать оператор L с погрешностью первого порядка.

Рассмотрим подробнее аппроксимацию уравнения Пуассона

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = f(x, y) \quad (3.116)$$

на девятиточечном шаблоне, состоящем из совокупности следующих точек $P_k = (x_k, y_k)$:

$$\begin{aligned} P_0 = (x_k, y_k), \quad P_1 = (x_0 + \delta_1 h, y_0), \quad P_2 = (x_0, y_0 + \delta_2 h), \\ P_3 = (x_0 - \delta_3 h, y_0), \quad P_4 = (x_0, y_0 - \delta_4 h), \end{aligned}$$

$$\begin{aligned} P_5 &= (x_0 + \delta_5 h, y_0 + \delta_5 h), & P_6 &= (x_0 - \delta_6 h, y_0 + \delta_6 h), \\ P_7 &= (x_0 - \delta_7 h, y_0 - \delta_7 h), & P_8 &= (x_0 + \delta_8 h, y_0 - \delta_8 h). \end{aligned}$$

Для $k = 1, 2, 5, 6$ положим $s_k = \delta_k + \delta_{k+2}$ и $d_k = \delta_k - \delta_{k+2}$. Конечно-разностную аппроксимацию (3.116) в точке P_0 ищем в виде уравнения

$$\sum_{k=0}^8 a_k u_k = \frac{h^2}{2} \sum_{k=0}^8 b_k f(x_k, y_k). \quad (3.117)$$

Если $u(x, y)$ имеет ограниченные производные до пятого порядка, то с помощью ряда Тейлора получаем

$$\begin{aligned} \sum_{k=0}^8 a_k u_k - \frac{h^2}{2} \sum_{k=0}^8 b_k \Delta u_k &= u(P_0) \sum_{k=0}^8 a_k \\ &+ \frac{h}{1!} \frac{\partial u}{\partial x}(P_0) \left[\delta_1 a_1 - \delta_3 a_3 + \delta_5 a_5 - \delta_6 a_6 - \delta_7 a_7 + \delta_8 a_8 \right] \\ &+ \frac{h}{1!} \frac{\partial u}{\partial y}(P_0) \left[\delta_2 a_2 - \delta_4 a_4 + \delta_5 a_5 + \delta_6 a_6 - \delta_7 a_7 - \delta_8 a_8 \right] \\ &+ \frac{h^2}{2!} \frac{\partial^2 u}{\partial x^2}(P_0) \left[\delta_1^2 a_1 + \delta_3^2 a_3 + \sum_{k=5}^8 \delta_k^2 a_k - \sum_{k=0}^8 b_k \right] \\ &+ \frac{h^2}{2!} \frac{\partial^2 u}{\partial x \partial y}(P_0) \left[2\delta_5^2 a_5 - 2\delta_6^2 a_6 + 2\delta_7^2 a_7 - 2\delta_8^2 a_8 \right] \\ &+ \frac{h^2}{2!} \frac{\partial^2 u}{\partial y^2}(P_0) \left[\delta_2^2 a_2 + \delta_4^2 a_4 + \sum_{k=5}^8 \delta_k^2 a_k - \sum_{k=0}^8 b_k \right] \\ &+ \frac{h^3}{3!} \frac{\partial^3 u}{\partial x^3}(P_0) \left[\delta_1^3 a_1 - \delta_3^3 a_3 + \delta_5^3 a_5 - \delta_6^3 a_6 - \delta_7^3 a_7 + \delta_8^3 a_8 \right. \\ &\quad \left. - 3\delta_1 b_1 + 3\delta_3 b_3 - 3\delta_5 b_5 + 3\delta_6 b_6 + 3\delta_7 b_7 - 3\delta_8 b_8 \right] \\ &+ \frac{h^3}{3!} \frac{\partial^3 u}{\partial x^2 \partial y}(P_0) \left[3\delta_5^3 a_5 + 3\delta_6^3 a_6 - 3\delta_7^3 a_7 - 3\delta_8^3 a_8 - 3\delta_2 b_2 + 3\delta_4 b_4 \right. \\ &\quad \left. - 3\delta_5 b_5 - 3\delta_6 b_6 + 3\delta_7 b_7 + 3\delta_8 b_8 \right] \\ &+ \frac{h^3}{3!} \frac{\partial^3 u}{\partial x \partial y^2}(P_0) \left[3\delta_5^3 a_5 - 3\delta_6^3 a_6 - 3\delta_7^3 a_7 + 3\delta_8^3 a_8 - 3\delta_1 b_1 \right. \\ &\quad \left. + 3\delta_3 b_3 - 3\delta_5 b_5 + 3\delta_6 b_6 + 3\delta_7 b_7 - 3\delta_8 b_8 \right] \\ &+ \frac{h^3}{3!} \frac{\partial^3 u}{\partial y^3}(P_0) \left[\delta_2^3 a_2 - \delta_4^3 a_4 + \delta_5^3 a_5 + \delta_6^3 a_6 - \delta_7^3 a_7 - \delta_8^3 a_8 \right. \\ &\quad \left. - 3\delta_2 b_2 + 3\delta_4 b_4 - 3\delta_5 b_5 - 3\delta_6 b_6 + 3\delta_7 b_7 + 3\delta_8 b_8 \right] \end{aligned} \quad (3.118)$$

$$\begin{aligned}
& + \frac{h^4}{4!} \frac{\partial^4 u}{\partial x^4}(P_0) \left[\delta_1^4 a_1 + \delta_3^4 a_3 + \sum_{k=5}^8 \delta_k^4 a_k - 6\delta_1^2 b_1 - 6\delta_3^2 b_3 - 6 \sum_{k=5}^8 \delta_k^2 b_k \right] \\
& + \frac{h^4}{4!} \left(\frac{\partial^4 u(P_0)}{\partial x^3 \partial y} + \frac{\partial^4 u(P_0)}{\partial x \partial y^3} \right) \left[4 \sum_{k=5}^8 (-1)^{k+1} \delta_k^4 a_k + 12 \sum_{k=5}^8 (-1)^k \delta_k^2 b_k \right] \\
& + \frac{h^4}{4!} \frac{\partial^4 u}{\partial x^2 \partial y^2}(P_0) \left[-6 \sum_{k=1}^4 \delta_k^2 b_k - 12 \sum_{k=5}^8 \delta_k^2 b_k + 6 \sum_{k=5}^8 \delta_k^4 a_k \right] \\
& + \frac{h^4}{4!} \frac{\partial^4 u(P_0)}{\partial y^4} \left[\delta_2^4 a_2 + \delta_4^4 a_4 + \sum_{k=5}^8 \delta_k^4 a_k - 6\delta_2^2 b_2 - 6\delta_4^2 b_4 - 6 \sum_{k=5}^8 \delta_k^2 b_k \right] + O(h^5).
\end{aligned}$$

Приравнявая коэффициенты при одинаковых производных в выражениях (3.116) и (3.118) до членов порядка h^4 , получаем систему уравнений относительно a_k и b_k .

Исследование этой довольно сложной системы показывает, что при выполнении условий

$$|d_k| \leq \frac{1}{2} \quad \text{и} \quad \frac{1}{2} \leq \delta_k \delta_{k+2} \leq 4 \quad \text{для} \quad k = 1, 2, 5, 6 \quad (3.119)$$

решение имеет вид

$$\begin{aligned}
a_k &= \frac{\rho_0 - \delta_{k+2} \mu_1 (\rho_2 + (-1)^k \rho_3)}{\delta_k s_k}, & a_{k+2} &= \frac{\rho_0 + \delta_k \mu_1 (\rho_2 + (-1)^k \rho_3)}{\delta_{k+2} s_k}, \\
a_{k+4} &= \frac{\rho_1 + \delta_{k+6} \rho_{k+1}}{\delta_{k+4} s_{k+4}} \mu_1, & a_{k+6} &= \frac{\rho_1 - \delta_{k+4} \rho_{k+1}}{\delta_{k+6} s_{k+4}} \mu_1, \\
\rho_0 &= \rho_1 \mu_2 + \rho_2 \mu_3 + \rho_3 \mu_4, & \rho_1 &= \frac{1}{12} (\alpha_{12} \alpha_{23} - \alpha_{13} \alpha_{22}) > 0, \\
\rho_2 &= \frac{1}{42} (\alpha_{13} \alpha_{21} - \alpha_{11} \alpha_{23}), & \rho_3 &= \frac{1}{72} (\alpha_{11} \alpha_{22} - \alpha_{12} \alpha_{21}), \\
\mu_1 &= \delta_1 \delta_3 d_1 + d_1^2 + d_2^2 + \delta_2 \delta_4 > 0, & \mu_2 &= 4(d_5^2 + \delta_5 \delta_7 + d_6^2 + \delta_6 \delta_8) > 0, \\
\mu_3 &= \delta_1 \delta_3 d_1 + \delta_2 \delta_4 d_2 + 4\delta_5 \delta_7 d_5, & \mu_4 &= -\delta_1 \delta_3 d_1 + \delta_2 \delta_4 d_2 + 4\delta_6 \delta_8 d_6, \\
\alpha_{k1} &= \mu_2 d_k + 2\mu_1 (-d_5 - (-1)^k d_6), & \alpha_{k2} &= \mu_3 d_k + \mu_1 (-\delta_1 \delta_{k+2} - 2\delta_5 \delta_7), \\
\alpha_{k3} &= \mu_4 d_k - (-1)^k \mu_1 (\delta_k \delta_{k+2} + 2\delta_6 \delta_8), \\
a_0 &= - \sum_{k=1}^8 a_k, & b_0 &= \delta_1^2 a_1 + \delta_3^2 a_3 + \sum_{k=5}^8 \delta_k^2 a_k - \sum_{k=1}^8 b_k, \\
b_k &= \frac{c_k \delta_{k+2} + c_{k+2}}{\delta_k s_k}, & b_{k+2} &= \frac{c_{k+2} - c_k \delta_k}{\delta_{k+2} s_k}, \\
b_{k+4} &= \frac{(-1)^{k+1} c_5}{2\delta_{k+4} s_{k+4}}, & b_{k+6} &= \frac{(-1)^{k+1} c_5}{2\delta_{k+6} s_{k+4}}, \quad k = 1, 2,
\end{aligned} \quad (3.120)$$

$$\begin{aligned}
c_1 &= \delta_5^3 a_5 - \delta_6^3 a_6 - \delta_7^3 a_7 + \delta_8^3 a_8, & c_2 &= \delta_5^3 a_5 + \delta_6^3 a_6 - \delta_7^3 a_7 - \delta_8^3 a_8, \\
6c_3 &= \delta_1^4 a_1 + \delta_3^4 a_3 + \sum_{k=5}^8 \delta_k^4 a_k, & 6c_4 &= \delta_2^4 a_2 + \sum_{k=4}^8 \delta_k^4 a_k, \\
3c_5 &= \delta_5^4 a_5 - \delta_6^4 + \delta_7^4 - \delta_8^4 a_8.
\end{aligned}$$

Таким образом, выражение (3.117) с коэффициентами, определяемыми из (3.120), аппроксимирует уравнение (3.116) с погрешностью $O(h^3)$. В простейшем регулярном случае (все $\delta_k = 1$, т. е. точки P_k располагаются в узлах квадратной сетки) рассмотренная аппроксимация совпадает с (3.44), имеющей погрешность аппроксимации $O(h^6)$.

Аналогично из (3.120) можно получить и девятиточечную схему Микеладзе (3.42) на равномерной прямоугольной сетке с шагами h_x, h_y , если положить

$$\begin{aligned}
\delta_1 h &= \delta_3 h = h_x, & \delta_2 h &= \delta_4 h = h_y, \\
\delta_5 &= \delta_6 = \delta_7 = \delta_8 = \sqrt{h_x^2 + h_y^2}/h.
\end{aligned}$$

Главное же достоинство приведенных формул состоит в том, что они дают не только способ построения разностных схем на неравномерной прямоугольной сетке, но и аппроксимации в случае, когда регулярная сетка пересекается криволинейной границей с условием Дирихле (аналог пятиточечной схемы Шортли–Уэллера (3.63)).

Рассмотренный принцип построения разностных уравнений методом неопределенных коэффициентов является наиболее общим и формально применим к аппроксимации высоких порядков для различных типов дифференциальных уравнений с переменными коэффициентами на сеточных шаблонах разной конфигурации и в том числе для трехмерных краевых задач со смешанными граничными условиями.

Главный вопрос здесь заключается, как это видно из предыдущих громоздких выкладок, в высокой технической сложности процесса получения формул. Однако эта проблема может быть успешно решена с помощью программ автоматического проведения аналитических преобразований и положительный опыт в данной области имеется.

Глава 4

Методы конечных объемов

Методы дискретизации исходных краевых задач, рассматриваемые выше, основаны на непосредственной аппроксимации дифференциальных уравнений конечно-разностными. Однако такой подход не является единственным, так как практически для каждой задачи математической физики известны интегральные соотношения, отражающие законы сохранения (баланса) той или иной субстанции (массы, тепла, энергии и т. д.). Эти интегральные равенства могут быть также аппроксимированы путем замены участвующих в них производных конечными разностями и применения квадратурных формул для приближения интегралов.

Поскольку аппроксимируемые при этом балансные соотношения рассматриваются на элементарных ячейках сетки, для такого подхода сложилось название, получившее позднее широкое распространение – *метод конечных объемов*.

Получающиеся сеточные алгебраические уравнения иногда совпадают с “обычными” конечно-разностными, а иногда несколько отличаются. Теоретический анализ для *балансных*, или *консервативных*, *разностных схем*, как их иногда называют, также почти не отличается, и обычно методическое преимущество здесь заключается в меньших требованиях к гладкости исходного решения, так как в законах сохранения содержатся производные более низких порядков, чем в эквивалентных дифференциальных уравнениях.

При общности методологических принципов балансных методов здесь имеются и существенные различия в тактике реализации, и на них мы остановимся отдельно.

Главным принципиальным качеством методов конечных объемов (МКО) является возможность применения поэлементной технологии аппроксимации и формирования алгебраической системы, что приближает их в этом плане к методам конечных элементов, которые рассматриваются в следующей главе. Аналогично локальным матрицам жесткости могут быть определены *локальные матрицы баланса*, из которых собирается (ассемблируется) *глобальная матрица* системы уравнений. Независимое

друг от друга вычисление локальных матриц баланса не только намного упрощает вычислительную сложность реализации многомерных задач, но и значительно повышает эффективность распараллеливания алгоритмов.

Библиография по методам конечных объемов включает огромное количество журнальных статей, однако специализированных монографий на эту тему мало (см. [81], некоторое количество книг также можно найти с помощью поисковиков в Интернете). Однако большое количество материала по МКО фактически можно найти в книгах по методам конечных разностей и их применениям, где исследуются консервативные, или балансные, вычислительные схемы для решения самых широких классов задач. В этой связи можно назвать монографии П.Н.Вабищевича [30, 31], В.И.Васильева [37, 38], В.А.Галкина [41], С.К.Годунова [46, 47], А.М.Гришина [54], Н.Т.Данаева [56], А.Н.Коновалова [90, 91], Л.А.Крукиера [97], В.В.Тишкина [107], Г.И.Марчука [124]–[127], С.Патанкара [142], В.В.Пененко [143], П.Роуча [152], А.А.Самарского [156]–[164], Г.С.Хакимзянова и Ю.И.Шокина с коллегами [195], Н.Н.Яненко [207], [209].

К МКО можно также отнести так называемые методы частиц в ячейках, в течение многих десятилетий успешно развиваемых для моделирования потоков субстанции в задачах электрофизики, гидро-газодинамики и плазмы. Здесь можно отметить изданные в разные годы монографии Ю.А.Березина и В.А.Вшивкова [23], А.Д.Григорьева [52], Ю.Н.Григорьева [53], В.П.Ильина [79], Р.Хокни и Дж.Иствуда [196]. Однако в данной книге эти подходы не рассматриваются.

4.1. Интегральное тождество Г. И. Марчука и его применение

4.1.1. Одномерный случай

Изложение данного подхода наиболее наглядно на примере одномерного уравнения диффузии

$$-\frac{d}{dx}p(x)\frac{du}{dx} + r(x)u = f(x), \quad p(x) > 0, \quad a < x < b. \quad (4.1)$$

Будем считать, что функции p , r и f – кусочно-непрерывные на отрезке $[a, b]$ с возможными точками разрыва первого рода. Решение $u(x)$ при этом в точке ξ разрыва функции p удо-

влетворяет так называемым условиям сопряжения $u|_{\xi^+} = u|_{\xi^-}$, $p^+ u'|_{\xi^+} = p^- u'|_{\xi^-}$, где знаковые индексы относятся к односторонним предельным значениям функций. Для построения аппроксимации возьмем сетку с множеством узлов

$$a = x_0 < x_1 < \dots < x_N < x_{N+1} = b, \quad (4.2)$$

включающим все точки разрыва.

Проинтегрируем уравнение (4.1) по x в пределах $(x_{i-1/2}, x_{i+1/2})$, $x_{i\pm 1/2} = (x_i + x_{i\pm 1})/2$, в результате чего получим соотношение баланса

$$J_{i+1/2} - J_{i-1/2} + \int_{x_{i-1/2}}^{x_{i+1/2}} (ru - f) dx = 0, \quad (4.3)$$

где введены обозначения $J_{i\pm 1/2} = J(x_{i\pm 1/2})$ функции потока

$$J(x) = -p(x) \frac{du}{dx}. \quad (4.4)$$

Для нахождения $J_{i-1/2}$ проинтегрируем исходное уравнение в пределах $(x_{i-1/2}, x)$, что дает соотношение

$$-p \frac{du}{dx} = J_{i-1/2} - \int_{x_{i-1/2}}^x (ru - f) dx'. \quad (4.5)$$

Поделив последнее соотношение на p (напомним, что этот коэффициент диффузии предполагается строго положительным) и произведя затем интегрирование на интервале (x_{i-1}, x_i) , после несложных преобразований находим выражение для потока

$$J_{i-1/2} = \left[-(u_i - u_{i-1}) + \int_{x_{i-1}}^{x_i} \frac{dx}{p} \int_{x_{i-1/2}}^x (ru - f) dx' \right] / \int_{x_{i-1}}^{x_i} \frac{dx}{p}. \quad (4.6)$$

Отсюда заменой индекса i на $i + 1$ получается формула для $J_{i+1/2}$. Подставляя теперь значения потоков в уравнение баланса (4.3), приходим к *интегральному тождеству Г. И. Марчука*

$$\begin{aligned}
& -\frac{u_{i+1}-u_i}{\int_{x_i}^{x_{i+1}} \frac{dx}{p}} + \frac{u_i-u_{i-1}}{\int_{x_{i-1}}^{x_i} \frac{dx}{p}} + \int_{x_{i-1/2}}^{x_{i+1/2}} (ru-f)dx \\
& = \left(\int_{x_i}^{x_{i+1}} \frac{dx}{p} \right)^{-1} \int_{x_i}^{x_{i+1}} \frac{dx}{p} \int_x^{x_{i+1/2}} (ru-f)dx' \\
& + \left(\int_{x_{i-1}}^{x_i} \frac{dx}{p} \right)^{-1} \int_{x_{i-1}}^{x_i} \frac{dx}{p} \int_{x_{i-1/2}}^x (ru-f)dx'. \tag{4.7}
\end{aligned}$$

Это точное соотношение может служить основой для вывода разностных уравнений с погрешностями различных порядков. В частности, если $r(x) \equiv 0$, а интегралы от функций f и $1/p$ вычисляются точно, то из (4.7) следуют трехточечные разностные уравнения с нулевой погрешностью.

Правая часть (4.7) является величиной порядка $O(h^3)$, если только все участвующие в ней подынтегральные функции внутри интервалов (x_i, x_{i+1}) гладкие. Это следует просто из того, что она обращается в нуль после применения квадратурной формулы центральных прямоугольников, которая, как хорошо известно, имеет ошибку интегрирования третьего порядка.

Если этой величиной пренебречь, а оставшиеся интегралы вычислить приближенно по той же формуле центральных прямоугольников, то на равномерной сетке получим "классическое" разностное уравнение диффузии (естественно, после перенормировки, сводящейся к делению на h) с ошибкой второго порядка. На неравномерной сетке (при использовании при-

ближения $\int_{x_{i-1/2}}^{x_{i+1/2}} f dx \approx f_i(h_i + h_{i-1})/2$) получаем трехточечное

уравнение, которое после деления на $\frac{h_i + h_{i-1}}{2}$ также переходит в обычное разностное уравнение, имеющее ошибку $O(h)$.

Естественно, что если функция $1/p$ точно не интегрируется, то приближенное вычисление соответствующих интегралов вносит дополнительную погрешность. Если $p(x)$ задана аналитически, то это, как правило, не приносит существенных трудностей. На практике часто значения функции p известны только в узлах сетки, и тогда требуется дополнительно привлекать интерполяцию (линейную или более высоких порядков, в зави-

симости от требуемой точности).

В принципе все интегралы в тождестве (4.7) при использовании соответствующих квадратурных формул могут быть определены со сколь угодно малой ошибкой (если $r \neq 0$, то придется искать функцию интерполировать предварительно по значениям u_i). В итоге это позволяет построить разностные уравнения с произвольно высоким порядком аппроксимации. Если $r(x) \equiv 0$, то все они будут оставаться трехточечными.

Рассмотренный подход применим и для диффузионно-конвективного уравнения вида (3.27) из п. 3.2.3, в котором член $q(x)u'$ заменяется на $\frac{d}{dx}(p\frac{ds}{dx}u)$. Тогда после обобщения понятия потока (содержащего диффузионную и конвективную составляющие)

$$J(x) = -p\left(\frac{du}{dx} - \frac{ds}{dx}u\right) = -pe^s \frac{d}{dx}(e^{-s}u) \quad (4.8)$$

уравнение (3.27) сводится к форме

$$\frac{dJ}{dx} + ru = f(x). \quad (4.9)$$

Далее выписывается соотношение баланса (4.3) и с небольшими изменениями повторяются предыдущие выкладки. Например, вместо (4.7) после отбрасывания малой правой части получается уравнение

$$-\frac{(e^{-s}u)_{i+1} - (e^{-s}u)_i}{\int_{x_i}^{x_{i+1}} \frac{dx}{p e^s}} + \frac{(e^{-s}u)_i - (e^{-s}u)_{i-1}}{\int_{x_{i-1}}^{x_i} \frac{dx}{p e^s}} + \int_{x_{i-1/2}}^{x_{i+1/2}} (ru - f)dx = 0. \quad (4.10)$$

Как и в предыдущей главе, отметим основные алгебраические свойства получаемых уравнений. Во-первых, если s не зависит от x и $r(x) \equiv 0$, то при любых приближениях оставшихся интегралов уравнения являются симметричными и положительного типа. Если $r \neq 0$, то при использовании простейшей квадратурной формулы прямоугольников эти свойства сохраняются, а при более сложных – теряются. Присутствие конвективного члена в (4.8), т. е. $s(x) \neq \text{const}$, нарушает как свойство положительности типа уравнений (нет диагонального преобладания), так и свойство симметричности. Однако, как легко заметить,

оба эти свойства возвращаются, если ввести новую неизвестную переменную $v_i = e^{-s_i} u_i$.

Из соотношения (4.10) следует одна замечательная аппроксимация, заслуживающая пристального внимания. В данном случае достаточно остановиться на частном случае $r(x) \equiv 0$, $p(x) \equiv 1$ и равномерной сетке. Проведя линейную интерполяцию для $s(x)$ и вычисляя затем интегралы в знаменателях точно, получаем разностное уравнение

$$\frac{1}{h^2} \left[-\frac{s_i - s_{i-1}}{e^{s_i - s_{i-1}} - 1} v_{i-1} + \left(\frac{s_i - s_{i-1}}{1 - e^{s_{i-1} - s_i}} + \frac{s_{i+1} - s_i}{e^{s_{i+1} - s_i} - 1} \right) v_i - \frac{s_{i+1} - s_i}{1 - e^{s_i - s_{i+1}}} v_{i+1} \right] = f_i, \quad (4.11)$$

исследование и применение которого связаны с именами А. М. Ильина, Шарфеттера–Гуммеля и многих других. Главная его особенность заключается в том, что оно имеет погрешность $O(h^2)$ и является абсолютно монотонным (более конкретно, уравнение (4.11) – положительного типа при сколь угодно большой величине конвективного члена). Наибольшее значение такого типа разностные схемы имеют при решении задач с так называемыми пограничными, или внутренними, слоями, характеризующимися наличием локальных подобластей с резко меняющимися свойствами решения (см. § 4.5).

Рассмотренные разностные уравнения являются балансными в том смысле, что сохраняют дискретный аналог свойства баланса субстанции для исходной дифференциальной задачи. Если в (4.9) для простоты взять $r = 0$, то для любых x' , x'' из интервала $[a, b]$ справедливо соотношение

$$J(x'') - J(x') = \int_{x'}^{x''} f(x) dx. \quad (4.12)$$

Вводя обозначение для сеточного потока

$$J_{i+1/2}^h = -[(e^{-s} u)_{i+1} - (e^{-s} u)_i] \left(\int_{x_i}^{x_{i+1}} \frac{dx}{p e^s} \right)^{-1} \quad (4.13)$$

и суммируя уравнения (4.10) в пределах от некоторого индекса i_1 до i_2 , получаем разностный баланс (дискретный эквивалент (4.12))

$$J_{i_1+1/2}^h - J_{i_2-1/2}^h = \sum_{i=i_1}^{i_2} \int_{x_{i-1/2}}^{x_{i+1/2}} f dx. \quad (4.14)$$

Важно отметить, что после замены интегралов в (4.14) на их дискретные аналоги получаемое равенство выполняется точно, какие бы квадратурные формулы для вычисления интегралов в (4.13), (4.14) не применялись.

Сделаем еще одно замечание об учете граничных условий при использовании интегрального тождества Г. И. Марчука. Если на концах отрезка $[a, b]$ заданы условия Дирихле, то вопросов нет, так как уравнения $v_0 = u_a$, $v_{N+1} = u_b$ на сетке (4.2) учитывают их точно. Когда же, например, в точке $x = a$ задан поток J вида (4.4) или (4.8) (наиболее естественное в математической физике условие), то целесообразно сетку выбрать так, чтобы середина первого интервала $x_{1/2}$ совпадала с границей. Тогда в балансном соотношении (4.3) при $i = 1$ поток $J_{1/2}$ учитывается точно и его не надо аппроксимировать. То же самое верно и при задании потока на правом конце отрезка.

В качестве эффективного применения интегрального тождества (4.7) мы покажем, как из него с помощью использования соответствующих квадратурных формул построить на равномерной сетке

$$x_i = a + ih, \quad i = 0, 1, \dots, N+1, \quad h = \frac{b-a}{N+1}, \quad x_{N+1} = b$$

схему четвертого порядка точности, если все участвующие в исходном уравнении (4.1) функции являются достаточно гладкими. Для аппроксимации последнего интеграла в левой части (4.7) с помощью построения отрезка ряда Тейлора мы можем написать

$$\begin{aligned} \int_{x_{i-1/2}}^{x_{i+1/2}} f dx &= \int_{x_{i-1/2}}^{x_{i+1/2}} \left(f_i + (x - x_i) f'_i + \frac{(x - x_i)^2}{2} f''_i + \right. \\ &\quad \left. \frac{(x - x_i)^3}{6} f'''_i \right) dx + O(h^5) \\ &= \frac{h}{24} (f_{i-1} + 22f_i + f_{i+1}) + O(h^5). \end{aligned} \quad (4.15)$$

Двойные интегралы в правой части (4.7) можно вычислить с хорошей точностью путем применения квадратурной формулы Симпсона (или парабол, см., например, [82]) к внешнему интегралу:

$$\begin{aligned}
\int_{x_i}^{x_{i+1}} \frac{dx}{p} \int_x^{x_{i+1/2}} f dx &= \frac{h}{6} \left(\frac{1}{p_i} \int_{x_i}^{x_{i+1/2}} f dx + \frac{1}{p_{i+1}} \int_{x_{i+1/2}}^{x_{i+1}} f dx \right) + O(h^5) = \\
&= \frac{h^2}{12} \left[\frac{1}{p_i} \left(f_i + \frac{f_i + f_{i+1}}{2} \right) - \frac{1}{p_{i+1}} \left(f_{i+1} + \frac{f_i + f_{i+1}}{2} \right) \right] + O(h^5) = \\
&= \frac{h^2}{24} \left(\frac{3f_i + f_{i+1}}{p_i} - \frac{3f_{i+1} + f_i}{p_{i+1}} \right) + O(h^5).
\end{aligned} \tag{4.16}$$

Выражения, аналогичные (4.15), (4.16), легко выписываются и для подынтегральной функции g_i . Вычисление интегралов в знаменателях левой части (4.7) осуществляем также с помощью формулы Симпсона следующим образом:

$$\left(\int_{x_i}^{x_{i+1}} \frac{dx}{p} \right)^{-1} = \bar{p}_{i+1/2}/h + O(h^5), \tag{4.17}$$

где введено обозначение

$$\bar{p}_{i+1/2} = [(p_i^{-1} + 4p_{i+1/2}^{-1} + p_{i+1}^{-1})/6]^{-1}. \tag{4.18}$$

В итоге после подстановки соотношений вида (4.15)–(4.18) в (4.7) и отбрасывания остаточных членов получаем трехточечное уравнение четвертого порядка точности:

$$\begin{aligned}
&\bar{p}_{i-1/2}(u_i - u_{i-1}) + \bar{p}_{i+1/2}(u_i - u_{i+1}) + \\
&+ \frac{h^2}{24}(r_{i-1}u_{i-1} + 22r_i u_i + r_{i+1}u_{i+1}) = \frac{h^2}{24} \left\{ f_{i-1} + 22f_i + f_{i+1} + \right. \\
&+ h \left[\bar{p}_{i+1/2} \left(\frac{3\varphi_i + \varphi_{i+1}}{p_i} - \frac{3\varphi_{i+1} + \varphi_i}{p_{i+1}} \right) + \right. \\
&\left. \left. \bar{p}_{i-1/2} \left(\frac{3\varphi_i + \varphi_{i-1}}{p_i} - \frac{3\varphi_{i-1} + \varphi_i}{p_{i-1}} \right) \right] \right\},
\end{aligned} \tag{4.19}$$

где для краткости используется обозначение $\varphi_i = r_i u_i - f_i$.

Рассмотрим теперь, как для этих уравнений аппроксимировать краевые условия 3-го рода

$$\begin{aligned}
(-s_a u' + t_a u)|_a &= g_a, \quad s_a \geq 0, \quad t_a \geq 0, \quad s_a + t_a > 0, \\
(s_b u' + t_b u)|_b &= g_b, \quad s_b \geq 0, \quad t_b \geq 0, \quad s_b + t_b > 0,
\end{aligned} \tag{4.20}$$

из которых формально следуют условия Дирихле и Неймана как частные случаи (при нулевых значениях s_a или s_b и t_a или t_b соответственно). Ниже условия 1-го рода ($s_a = 0$ или $s_b = 0$) мы

можем из рассмотрения исключить, так как их учет легко производится без внесения какой-либо дополнительной погрешности (при этом уравнения (4.19) строятся только для внутренних узлов сетки $i = 1, \dots, N$, к которым добавляются точные равенства $u_0 = u_a = g_a/t_a$ и/или $u_{N+1} = u_b = g_b/t_b$).

Остановимся подробнее на учете первого из условий (4.20). В силу определения потока (4.4) справедливо равенство

$$-\int_{x_0}^{x_1} u' dx = u_0 - u_1 = \int_{x_0}^{x_1} J \frac{dx}{p},$$

после применения к которому формулы Симпсона получаем соотношение

$$u_0 - u_1 = \frac{h}{6} \left(\frac{J_0}{p_0} + 4 \frac{J_{1/2}}{p_{1/2}} + \frac{J_1}{p_1} \right) + O(h^5).$$

Отсюда с помощью простых преобразований имеем

$$\begin{aligned} u_0 - u_1 - h \bar{p}_{1/2}^{-1} J_0 &= [4p_{1/2}^{-1}(J_{1/2} - J_0) + p_1^{-1}(J_1 - J_0)]h/6 + O(h^5) = \\ &= [4p_{1/2}^{-1} \int_{x_0}^{x_{1/2}} \varphi(x) dx + p_1^{-1} \int_{x_0}^{x_1} \varphi dx]h/6 + O(h^5), \quad \varphi = f - ru, \\ \bar{p}_{1/2}^{-1} &= (p_0^{-1} + 4p_{1/2}^{-1} + p_1^{-1})/6. \end{aligned} \tag{4.21}$$

Для участвующих в данном уравнении интегралов с помощью конечно-разностных аппроксимаций производных

$$\varphi'_0 = \frac{4\varphi_1 - 3\varphi_0 - \varphi_2}{2h} + O(h^2), \quad \varphi''_0 = \frac{\varphi_0 - 2\varphi_1 + \varphi_2}{h^2} + O(h)$$

можно построить следующие квадратурные формулы:

$$\begin{aligned} \int_{x_0}^{x_{1/2}} \varphi dx &= \varphi_0 + \frac{h}{2} \varphi'_0 + \frac{h^2}{2} \varphi''_0 + O(h^4) = \frac{h}{24} (8\varphi_0 + 5\varphi_1 - \varphi_2) + O(h^4), \\ \int_{x_0}^{x_1} \varphi dx &= \frac{h}{12} (5\varphi_0 + 8\varphi_1 - \varphi_2) + O(h^4). \end{aligned}$$

Подставляя данные выражения в (4.21), а также учитывая следующее из (4.20) равенство

$$J_0 = p_0(g_a - t_a u_0)/s_a,$$

приходим к трехточечному уравнению с погрешностью четвертого порядка:

$$\begin{aligned} u_0 - u_1 - h\bar{p}_{1/2}^{-1}p_0(g_a - t_a u_0)/s_a = \\ = \frac{h^2}{36p_{1/2}}(8\varphi_0 - 5\varphi_1 - \varphi_2) + \frac{h^2}{72p_1}(5\varphi_0 + 8\varphi_1 - \varphi_2) + O(h^5). \end{aligned}$$

Совершенно аналогичным образом аппроксимация второго из краевого условия (4.20) приводит к уравнению с погрешностью $O(h^5)$ относительно неизвестных u_{N+1}, u_N, u_{N-1} .

И последний комментарий данного пункта: наличие известных особенностей подынтегральных функций в тождестве (4.7) не препятствует, вообще говоря, построению разностных уравнений сколь угодно высокой точности. Для этого достаточно только применить известные квадратурные формулы для вычисления сингулярных интегралов.

4.1.2. Двумерный случай

Посмотрим теперь, насколько данная идея плодотворна для двумерного уравнения диффузии, которое путем введения проекций вектора потока

$$J^x = -x^\alpha p \frac{\partial u}{\partial x}, \quad J^y = -p \frac{\partial u}{\partial y} \quad (4.22)$$

запишем в виде

$$\frac{1}{x^\alpha} \frac{\partial J^x}{\partial x} + \frac{\partial J^y}{\partial y} = f(x, y). \quad (4.23)$$

Здесь в случае $\alpha = 1$ уравнение соответствует цилиндрической системе координат, а при $\alpha = 0$ – декартовой. Как и ранее, возьмем прямоугольную сетку (необязательно равномерную) с координатными линиями x_i, y_j . Уравнение (4.23) умножим на x^α и проинтегрируем по ячейке $x_{i-1/2} \leq x \leq x_{i+1/2}, y_{j-1/2} \leq y \leq y_{j+1/2}$, в результате чего приходим к соотношению баланса

$$\begin{aligned} \int_{y_{j-1/2}}^{y_{j+1/2}} (J_{i+1/2}^x - J_{i-1/2}^x) dy + \int_{x_{i-1/2}}^{x_{i+1/2}} (J_{j+1/2}^y - J_{j-1/2}^y) x^\alpha dx \\ = \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} f x^\alpha dx dy, \end{aligned} \quad (4.24)$$

где введены обозначения $J_{i\pm 1/2}^x = J^x(x_{i\pm 1/2}, y)$, $J_{j\pm 1/2}^y = J^y(x, y_{j\pm 1/2})$.

Выразим теперь из исходного уравнения (4.23) величину $\frac{\partial J}{\partial x}$ и проинтегрируем полученное выражение в пределах $(x_{i-1/2}, x)$:

$$J^x(x, y) = J_{i-1/2}^x + \int_{x_{i-1/2}}^x \left[f - \frac{\partial J^y}{\partial y} \right] (x')^\alpha dx'. \quad (4.25)$$

Аналогично получаем для второй проекции потока

$$J^y(x, y) = J_{j-1/2}^y + \int_{y_{j-1/2}}^y \left(f - \frac{1}{x^\alpha} \frac{\partial J^x}{\partial x} \right) dy'. \quad (4.26)$$

Подставляя в левые части (4.25) и (4.26) выражения для потоков из (4.22) и производя интегрирование в пределах (x_{i-1}, x_i) , (y_{j-1}, y_j) с предварительным делением на px^α в первом случае и на p — во втором, находим следующие необходимые для дальнейшего выражения:

$$\begin{aligned} J_{i-1/2}^x &= \frac{-u_i + u_{i-1} + \int_{x_{i-1}}^{x_i} \frac{dx}{px^\alpha} \int_{x_{i-1/2}}^x \left(f - \frac{\partial J^y}{\partial y} \right) (x')^\alpha dx'}{\int_{x_{i-1}}^{x_i} \frac{dx}{px^\alpha}}, \\ J_{j-1/2}^y &= \frac{-u_j + u_{j-1} + \int_{y_{j-1}}^y \frac{dy}{p} \int_{y_{j-1/2}}^y \left(f - \frac{1}{x^\alpha} \frac{\partial J^x}{\partial x} \right) dy'}{\int_{y_{j-1}}^{y_j} \frac{dy}{p}}, \end{aligned} \quad (4.27)$$

где используются обозначения $u_i = u(x_i, y)$, $u_j = u(x, y_j)$. Подставляя далее (4.27) и аналогичные им выражения для $J_{i+1/2}^x$, $J_{j+1/2}^y$ в соотношении баланса (4.24), получаем *двумерное интегральное тождество*

$$\int_{y_{j-1/2}}^{y_{j+1/2}} \left(\frac{-u_{i+1} + u_i}{\int_{x_i}^{x_{i+1}} \frac{dx}{px^\alpha}} + \frac{u_i - u_{i-1}}{\int_{x_{i-1}}^{x_i} \frac{dx}{px^\alpha}} \right) dy + \int_{x_{i-1/2}}^{x_{i+1/2}} \left(\frac{-u_{j+1} + u_j}{\int_{y_j}^{y_{j+1}} \frac{dy}{p}} + \frac{u_j - u_{j-1}}{\int_{y_{j-1}}^{y_j} \frac{dy}{p}} \right) x^\alpha dx$$

$$= \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} f x^\alpha dx dy + \phi_{i,j}. \quad (4.28)$$

Здесь $\phi_{i,j}$ обозначает величину

$$\phi_{i,j} = \int_{y_{j-1/2}}^{y_{j+1/2}} (\varphi_{i+1/2}^y - \varphi_{i-1/2}^y) dy + \int_{x_{i-1/2}}^{x_{i+1/2}} (\varphi_{j+1/2}^x - \varphi_{j-1/2}^x) x^\alpha dx, \quad (4.29)$$

где, в свою очередь, переменные $\varphi_{i\pm 1/2}^y$, $\varphi_{j\pm 1/2}^x$ определяются формулами

$$\varphi_{i-1/2}^y = \varphi^y(x_{i-1/2}, y) = \int_{x_{i-1}}^{x_i} \frac{dx}{p x^\alpha} \int_x^{x_{i-1/2}} \left(f - \frac{\partial J^y}{\partial y} \right) (x')^\alpha dx' / \int_{x_{i-1}}^{x_i} \frac{dx}{p x^\alpha}, \quad (4.30)$$

$$\varphi_{j-1/2}^x = \varphi^x(x, y_{j-1/2}) = \int_{y_{j-1}}^{y_j} \frac{\partial y}{p} \int_y^{y_{j-1/2}} \left(f - \frac{1}{x^\alpha} \frac{\partial J^x}{\partial x} \right) dy' / \int_{y_{j-1}}^{y_j} \frac{\partial y}{p}.$$

Тождество (4.28) может служить основой для вывода различного типа разностных уравнений после применения тех или иных формул приближенного интегрирования. Последний член $\phi_{i,j}$ проще всего отбросить, так как величины $\varphi_{i\pm 1/2}^y$ и $\varphi_{j\pm 1/2}^x$ после применения квадратурных формул центральных прямоугольников обращаются в нуль, т. е. они являются величинами порядка $O(h^3)$. Если далее применить простейшие квадратурные формулы прямоугольников и трапеций, то получаем пятиточечную аппроксимацию

$$\begin{aligned} & \frac{h_j^y + h_{j-1}^y}{2} \left(x_{i+1/2}^\alpha p_{i+1/2,j} \frac{-v_{i+1,j} + v_{i,j}}{h_i^x} + x_{i-1/2}^\alpha p_{i-1/2,j} \frac{v_{i,j} - v_{i-1,j}}{h_{i-1}^x} \right) \\ & + \frac{h_i^x + h_{i-1}^x}{2} x_i^\alpha \left(p_{i,j+1/2} \frac{-v_{i,j+1} + v_{i,j}}{h_j^y} + p_{i,j-1/2} \frac{v_{i,j} - v_{i,j-1}}{h_{j-1}^y} \right) \\ & = \frac{(h_i^x + h_{i-1}^x)(h_j^y + h_{j-1}^y)}{4} x_i^\alpha f_{i,j}, \quad (4.31) \end{aligned}$$

которая является симметричной, положительного типа и имеет погрешность $O(h^2)$ на равномерной сетке и, естественно, первого порядка – на неравномерной (после “обычной” перенормировки уравнений).

Из (4.28) несложно также вывести более точные девятиточечные уравнения, что мы предоставляем заинтересованному читателю. Мы не будем останавливаться также на вопросах аппроксимации краевых условий. Если граница области Ω состоит из отрезков прямых, параллельных координатным осям, то их учет производится совершенно аналогично одномерным задачам. В противном же случае картина усложняется, и этому позже будет уделено специальное внимание.

4.2. Интегро-балансные аппроксимации

4.2.1. Одномерные уравнения

Данная группа методов построения балансных (консервативных) разностных уравнений базируется, как и предыдущая, на использовании интегральных законов сохранения субстанции, которые для уравнения диффузии (4.1) записываются в виде (4.3) или подобных ему выражений. Однако дальнейшая идея заключается в конструировании аппроксимации разностей потоков, связанных с решением, вследствие (4.4), интегральным соотношением

$$u(x'') - u(x') = - \int_{x'}^{x''} J(x) \frac{dx}{p(x)}, \quad x', x'' \in [a, b]. \quad (4.32)$$

Если здесь при $x' = x_i$, $x'' = x_{i+1}$ под интегралом поток проинтерполировать (отсюда и название метода – *интегро-интерполяционный*, принадлежащее А. А. Самарскому [156]), то величины $J_{i\pm 1/2}$ легко выражаются через u_i :

$$J_{i+1/2} = -(u_{i+1} - u_i) / \int_{x_i}^{x_{i+1}} \frac{dx}{p} - \phi_{i+1/2}, \quad (4.33)$$

где остаточный член $\phi_{i\pm 1/2}$ записывается в виде

$$\begin{aligned} \phi_{i+1/2} = & \left[J'_{i+1/2} \int_{x_i}^{x_{i+1}} \frac{(x - x_{i+1/2}) dx}{p} \right. \\ & \left. + \frac{1}{2} \int_{x_i}^{x_{i+1}} (x - x_{i+1/2})^2 J''(\xi_{i+1/2}) \frac{dx}{p} \right] / \int_{x_i}^{x_{i+1}} \frac{dx}{p} \quad (4.34) \end{aligned}$$

и имеет порядок $O(h^2)$, если функция $p(x)$ достаточно гладкая на интервале $[x_i, x_{i+1}]$, а $\xi_{i+1/2}$ – некоторая точка из этого интервала.

После подстановки (4.33) в балансное соотношение (4.3) получаем равенство

$$-\frac{u_{i+1}-u_i}{\int_{x_i}^{x_{i+1}} \frac{dx}{p}} + \frac{u_i-u_{i-1}}{\int_{x_{i-1}}^{x_i} \frac{dx}{p}} + \int_{x_{i-1/2}}^{x_{i+1/2}} (ru-f)dx = \phi_{i+1/2} - \phi_{i-1/2}, \quad (4.35)$$

левая часть которого та же, что и в интегральном тождестве (4.7). Отсюда следует, что правые части у них тоже должны совпадать, хотя внешне они существенно отличаются: в (4.7) остаточные члены выписываются через искомое решение, а в (4.35) – через производные от потока.

С точки зрения минимизации ошибки $\phi_{i+1/2}$ применение формул (4.7) и (4.33) не является, вообще говоря, оптимальным. Как известно из теории гауссовских квадратурных формул (см., например, [24]), при заданных пределах интегрирования и числе квадратурных узлов их координаты и коэффициенты однозначно определяются весовой функцией. А именно абсциссы формул Гаусса являются корнями многочленов соответствующих порядков, ортогональных на интервале интегрирования с заданной весовой функцией. Поэтому вместо $x_{i+1/2}$ оптимальной является некоторая точка $\xi_{i+1/2}$, зависящая от функции $p(x)$, и вместо (4.33) можно записать

$$u_{i+1} - u_i = -J_{\xi_{i+1/2}} \int_{x_i}^{x_{i+1}} \frac{dx}{p} + \Phi_{i+1/2}, \quad (4.36)$$

где ошибка $\Phi_{i+1/2}$ имеет вид

$$\Phi_{i+1/2} = \frac{J''(\eta_{i+1/2})}{2} \int_{x_i}^{x_{i+1}} \frac{(x - \xi_{i+1/2})^2}{p(x)} dx, \quad \eta_{i+1/2} \in [x_i, x_{i+1}].$$

Если пользоваться выражением (4.36), то в соотношении баланса в качестве x' , x'' надо брать $\xi_{i-1/2}$, $\xi_{i+1/2}$ и последние значения необходимо использовать в (4.35) вместо $x_{i-1/2}$, $x_{i+1/2}$ соответственно. При этом в определенном смысле “наилучшая” для $r=0$ схема второго порядка на неравномерной сетке имеет вид

$$\begin{aligned} \frac{-p_{i+1/2}(v_{i+1} - v_i)}{h_i} + \frac{p_{i-1/2}(v_i - v_{i-1})}{h_{i-1}} &= \int_{\xi_{i-1/2}}^{\xi_{i+1/2}} f(x) dx, \\ \xi_{i+1/2} &= \int_{x_i}^{x_{i+1}} x \frac{dx}{p} \bigg/ \int_{x_i}^{x_{i+1}} \frac{dx}{p}, \quad p_{i+1/2} = h_i \bigg/ \int_{x_i}^{x_{i+1}} \frac{dx}{p(x)}. \end{aligned} \quad (4.37)$$

Эту схему можно конкретизировать после выбора способа вычисления $\xi_{i+1/2}$, а также линейной или квадратичной интерполяции функции $f(x)$ под интегралом (если он не вычисляется точно).

Рассмотренный подход, который будем называть *интегробалансным*, можно развивать следующим образом. Для аппроксимации интеграла в (4.32) естественно попытаться использовать формулы численного интегрирования более высоких порядков. Обозначая через n_i , $\alpha_{i,k}$ и $J_{i,k}$ число квадратурных узлов, коэффициенты и значения потока в точках $x_{i,k} \in [x_i, x_{i+1}]$, мы можем записать

$$\begin{aligned} \int_{x_i}^{x_{i+1}} J(x) \frac{dx}{p(x)} &= \frac{h_i}{p_{i+1/2}} \sum_{k=1}^{n_i} \alpha_{i,k} J_{i,k} + \phi_{i+1/2}, \\ \sum_{k=1}^{n_i} \alpha_{i,k} &= 1, \quad J_{i,k} = J(x_{i,k}), \\ \phi_{i+1/2} &= C_i h_i^\beta \frac{\partial^\beta J(\eta_i)}{\partial x^\beta}, \quad \eta_i \in [x_i, x_{i+1}]. \end{aligned} \quad (4.38)$$

Здесь не зависящая от h_i константа C_i и показатель степени β определяются конкретной формулой интегрирования. Например, при использовании гауссовых квадратур для $p(x) \equiv 1$ имеем

$$\phi_{i+1/2} = \frac{h_i (n_i!)^4}{[(2n_i)!]^3 (2n_i + 1)} J^{(2n_i)}(\eta_i).$$

Подчеркнем, что в результате выбора n_i , $\alpha_{i,k}$ и $x_{i,k}$ порядок погрешности β в формулах (4.38) может быть сколь угодно большим, техника приближенного интегрирования позволяет это сделать даже при наличии известной особенности функции потока $J(x)$.

Выписывая аналогичные (4.32) балансные соотношения на интервалах $[x_{i-1}, x_i]$ и $[x_i, x_{i+1}]$, с помощью выражений (4.38) после несложных выкладок получаем

$$\begin{aligned}
& -p_{i+1/2} \frac{u_{i+1} - u_i}{h_i} + p_{i-1/2} \frac{u_i - u_{i-1}}{h_{i-1}} \\
& = \sum_{k=1}^{n_i} \alpha_{i,k} J_{i,k} - \sum_{k=1}^{n_{i-1}} \alpha_{i-1,k} J_{i-1,k} - \frac{p_{i+1/2}}{h_i} \phi_{i+1/2} + \frac{p_{i-1/2}}{h_{i-1}} \phi_{i-1/2} \quad (4.39)
\end{aligned}$$

На различных сеточных интервалах значения квадратурных коэффициентов, да и числа n_i могут отличаться. Но в силу их одинаковых нормировок $\sum_{k=1}^{n_i} \alpha_{i,k} = 1$ соотношения (4.39) можно переписать в виде

$$\begin{aligned}
& -p_{i+1/2} \frac{u_{i+1} - u_i}{h_i} + p_{i-1/2} \frac{u_i - u_{i-1}}{h_{i-1}} = \sum_{l=1}^{m_i} \gamma_{i,l} (J_{i,l} - J_{i-1,l}) + \hat{\psi}_i, \\
& m_i = \sup\{n_i, n_{i-1}\}, \quad (4.40) \\
& \hat{\psi}_i = \frac{p_{i-1/2}}{h_{i-1}} \phi_{i-1/2} - \frac{p_{i+1/2}}{h_i} \phi_{i+1/2}, \quad \sum_{l=1}^{m_i} \gamma_{i,l} = 1.
\end{aligned}$$

Здесь значения $J_{i,l} = J(x_{i,l})$ берутся в точках из того же набора квадратурных узлов, что и в формулах (2.32), но конкретные величины $x_{i,l}$ и $x_{i,k}$ могут отличаться даже при совпадении $k = l$ (в частности, в (4.40) одному $J_{i,l}$ могут соответствовать в (4.39) значения потока в разных точках $x_{i,k'}$ и $x_{i,k''}$).

Поскольку из балансного соотношения (4.12) имеем

$$J_{i,l} - J_{i-1,l} = \int_{x_{i-1,l}}^{x_{i,l}} f(x) dx,$$

то выражения (4.40) приводятся к виду

$$-p_{i+1/2} \frac{u_{i+1} - u_i}{h_i} + p_{i-1/2} \frac{u_i - u_{i-1}}{h_{i-1}} = \sum_{l=1}^{m_i} \left(\gamma_{i,l} \int_{x_{i-1,l}}^{x_{i,l}} f(x) dx \right)_i + \hat{\psi}_i. \quad (4.41)$$

Если функция $f(x)$ точно не интегрируется, то приближенное вычисление правой части в (4.41) вносит дополнительную погрешность $\check{\psi}_i$, но она может быть сделана сколь угодно малой (при достаточном числе узловых точек) за счет выбора способа интерполяции и соответствующих квадратурных формул. Предполагая для простоты, что эти формулы строятся для каждого i по одинаковому числу узлов r , приходим в итоге к выражениям

$$-p_{i+1/2} \frac{u_{i+1} - u_i}{h_i} + p_{i-1/2} \frac{u_i - u_{i-1}}{h_{i-1}} = \sum_{l=1}^{m_i} \gamma_{i,l} \sum_{k=1}^r q_{i,l,k} f(x_{i,l,k}) + \hat{\psi}_i + \check{\psi}_i.$$

Отбрасывая теперь погрешность аппроксимации, получаем сеточное уравнение

$$-p_{i+1/2} \frac{v_{i+1} - v_i}{h_i} + p_{i-1/2} \frac{v_i - v_{i-1}}{h_{i-1}} = \sum_{l=1}^{m_i} \gamma_{i,l} \sum_{k=1}^r q_{i,l,k} f_{i,l,k}. \quad (4.42)$$

Рассмотрим теперь, как данный подход можно применить к аппроксимации граничных условий, не привнося никаких дополнительных погрешностей. Запишем краевое условие на левом конце отрезка $x = a$ в формально обобщенном виде

$$\alpha_a u(a) + \beta_a J(a) = g_a, \quad (4.43)$$

где при обращении чисел α_a или β_a в нуль получаем частные случаи условий Неймана или Дирихле.

Сетку выбираем с $x_0 = a$ и вместо соотношения (4.39) при $i = 0$ выписываем точное равенство, вытекающее из (4.32), (4.38):

$$-p_{1/2} \frac{u_1 - u_0}{h_0} = \sum_{k=1}^{n_0} \alpha_{0,k} J_{0,k} + \frac{p_{1/2}}{h_0} \psi_{1/2}, \quad (4.44)$$

в котором все обозначения соответствуют (4.38). Умножая равенство (4.44) на β_a и прибавляя к нему почленно (4.43), получаем

$$\alpha_a u_0 - \beta_a p_{1/2} \frac{u_1 - u_0}{h_0} = \beta_a \sum_{k=1}^{n_0} \alpha_{0,k} [J_{0,k} - J(a)] + \beta_a \frac{p_{1/2}}{h_0} \psi_{1/2} + g_a. \quad (4.45)$$

Заменяя далее разности потоков с помощью балансных соотношений вида (4.12), имеем

$$\begin{aligned} \alpha_a u_0 - \beta_a p_{1/2} \frac{u_1 - u_0}{h_0} &= \beta_a \sum_{k=1}^{n_0} \alpha_{0,k} \int_a^{x_{0,k}} f(x) dx + \hat{\psi}_0 + g_a \\ &= \beta_a \sum_{l=1}^{m_0} \gamma_{0,l} \sum_{k=1}^r q_{0,l,k} f(x_{0,l,k}) + \hat{\psi}_0 + \check{\psi}_0 + g_a, \end{aligned} \quad (4.46)$$

где в данном случае $m_0 = n_0$, $\gamma_{0,l} = \alpha_{0,l}$, $\hat{\psi}_0 = \beta_a \frac{p_{1/2}}{h_0} \psi_{1/2}$, а $\check{\psi}_0$, как и в (4.41), есть суммарная погрешность вычисления m_0 интегралов от функции $f(x)$.

В итоге для $i = 0$ получаем аналогичное по обозначениям с (4.42) уравнение

$$\alpha_a v_0 - \beta_a p_{1/2} \frac{v_1 - v_0}{h_0} = \beta_a \sum_{l=1}^{m_0} \gamma_{0,l} \sum_{k=1}^r q_{0,l,k} f_{0,l,k} + g_a. \quad (4.47)$$

Очевидно, что при $\beta_a \neq 0$ для сохранения симметричности системы последнее уравнение надо поделить на β_a (если же $\beta_a = 0$, то проведенные после (4.43) выкладки не нужны, и к уравнениям (4.42) достаточно добавить из условия Дирихле тривиальное равенство $v_0 = g_a/\alpha_a$).

Если на правом конце задано также граничное условие $\alpha_b u(b) + \beta_b J(b) = g_b$, то сетка выбирается с правым узлом $x_{N+1} = x_b$ и из аналогичных построений получаем уравнение, замыкающее систему (4.42), (4.47):

$$\alpha_b v_{N+1} + \beta_b p_{N+1/2} \frac{v_{N+1} - v_N}{h_N} = \sum_{l=1}^{m_N} \gamma_{N,l} \sum_{k=1}^r q_{N,l,k} f_{N,l,k} + g_b. \quad (4.48)$$

Легко видеть, что если $\alpha_a \beta_a > 0$, $\alpha_b \beta_b > 0$, то уравнения (4.47), (4.48) являются абсолютно монотонными, конкретнее – при любых шагах сетки являются уравнениями положительного типа (напомним, что функция $p(x)$ по изначальному предположению строго положительна). Если же одна из пар граничных коэффициентов имеет разные знаки, то для положительности типа соответствующего разностного граничного уравнения требуется выполнение неравенства типа $|\beta_a p_{1/2}| > \alpha_a h_0$.

Построенные сеточные уравнения являются балансными в двойном смысле. Во-первых, они выведены из соотношений “детального” баланса (4.32) (кстати, при наличии граничного условия Неймана, т. е. при $\alpha_a \alpha_b = 0$, значения потоков $J_{i,l}$ легко вычисляются рекуррентным образом). Во-вторых, для численных усредненных потоков, определяемых как

$$J_{i+1/2} = \sum_{k=1}^{n_i} \alpha_{i,k} J_{i,k} = \sum_{l=1}^{m_i} \gamma_{i,l} J_{i,l} = -p_{i+1/2} \frac{v_{i+1} - v_i}{h_i},$$

выполняются (точно!) суммарные соотношения баланса

$$J_{i''+1/2} - J_{i'-1/2} = \sum_{i=i'}^{i''} \sum_{l=1}^{m_i} \sum_{k=1}^r q_{i,l,k} f_{i,l,k},$$

$$0 \leq i' < i'' \leq N. \quad (4.49)$$

Аналогичные разностные аппроксимации произвольного порядка точности могут быть построены и для диффузионно-конвективного уравнения (3.27). Вводя определение потока вида (4.8), (4.9) и повторяя предыдущие выкладки, мы вместо (4.42) при $r = 0$ получим уравнение

$$\begin{aligned} -p_{i+1/2} \frac{(e^{-s}v)_{i+1} - (e^{-s}v)_i}{h_i} + p_{i-1/2} \frac{(e^{-s}v)_i - (e^{-s}v)_{i-1}}{h_{i-1}} \\ = \sum_{l=1}^{m_i} \gamma_{i,l} \sum_{k=1}^r q_{i,l,k} f(x_{i,l,k}), \end{aligned} \quad (4.50)$$

где значения $x_{i,l,k}$, $\gamma_{i,l}$ и $q_{i,l,k}$ будут, вообще говоря, отличаться от предыдущих, а величины $p_{i+1/2}$ имеют вид

$$p_{i+1/2} = h_i \int_{x_i}^{x_{i+1}} \frac{dx}{pe^s}.$$

4.2.2. Двумерные уравнения

Описанный подход представляется исчерпывающим проблему конструирования разностных аппроксимаций для одномерных краевых задач. Но насколько эти идеи универсальны и применимы для многомерных задач? Успехи здесь гораздо скромнее, и чтобы проиллюстрировать имеющиеся трудности, остановимся на двумерном уравнении Пуассона в декартовой системе координат

$$\frac{\partial J^x}{\partial x} + \frac{\partial J^y}{\partial y} = f(x, y), \quad J^x = -\frac{\partial u}{\partial x}, \quad J^y = -\frac{\partial u}{\partial y}.$$

Используя для простоты равномерную прямоугольную сетку с шагами h_x , h_y , из определения проекций потока и применения квадратурной формулы Симпсона мы можем записать

$$\begin{aligned} u_{i+1,j} - u_{i,j} &= - \int_{x_i}^{x_{i+1}} J^x(x, y_j) dx \\ &= -\frac{h_x}{6} (J_{i,j}^x + 4J_{i+1/2,j}^x + J_{i+1,j}^x) - \psi_{i+1/2,j}^x, \end{aligned} \quad (4.51)$$

$$\begin{aligned}
u_{i,j+1} - u_{i,j} &= - \int_{y_j}^{y_{j+1}} J^y(x_i, y) dy \\
&= - \frac{h_y}{6} (J_{i,j}^y + 4J_{i,j+1/2}^y + J_{i,j+1}^y) - \psi_{i,j+1/2}^y,
\end{aligned}$$

где погрешности численного интегрирования равны (см. [3]):

$$\begin{aligned}
\psi_{i+1/2,j}^x &= \frac{h_x^5}{90} \frac{\partial^4 J^x(\xi_{i+1/2}, y_j)}{\partial x^4}, & \xi_{i+1/2} &\in [x_i, x_{i+1}], \\
\psi_{i,j+1/2}^y &= \frac{h_y^5}{90} \frac{\partial^4 J^y(x_i, \eta_{j+1/2})}{\partial y^4}, & \eta_{i+1/2} &\in [y_i, y_{j+1}].
\end{aligned}$$

Запишем теперь соотношения (4.51) с меньшими на единицу индексами i, j и после несложного комбинирования (почленное вычитание полученных равенств из исходных) получим

$$\begin{aligned}
-\Delta_x u_{i,j} &= \frac{-u_{i-1,j} + 2u_{i,j} - u_{i+1,j}}{h_x^2} = \frac{1}{6h_x} \left[J_{i+1,j}^x - J_{i-1,j}^x + 4(J_{i+1/2,j}^x - J_{i-1/2,j}^x) \right] \\
&\quad + h_x^{-2} (\psi_{i+1/2,j}^x - \psi_{i-1/2,j}^x), \\
-\Delta_y u_{i,j} &= \frac{-u_{i,j-1} + 2u_{i,j} - u_{i,j+1}}{h_y^2} = \frac{1}{6h_y} \left[J_{i,j+1}^y - J_{i,j-1}^y + 4(J_{i,j+1/2}^y - J_{i,j-1/2}^y) \right] \\
&\quad + h_y^{-2} (\psi_{i,j+1/2}^y - \psi_{i,j-1/2}^y).
\end{aligned} \tag{4.52}$$

Рассмотрим выражение, получаемое суммированием линейных комбинаций (4.52): вторые разности по x берем на $(j-1)$, j и $(j+1)$ -й горизонтальных линиях сетки (соответственно с коэффициентами $\theta h_x h_y$, $(6-2\theta)h_x h_y$, $\theta h_x h_y$), а вторые разности по y – на вертикальных линиях с номерами $i-1$, i , $i+1$ (с такими же весовыми множителями). В итоге получаем девятиточечное конечно-разностное выражение

$$\begin{aligned}
&-h_x h_y \left[\theta \Delta_x u_{i,j-1} + (6-2\theta) \Delta_x u_{i,j} + \theta \Delta_x u_{i,j+1} + \theta \Delta_y u_{i-1,j} \right. \\
&\quad \left. + (6-2\theta) \Delta_y u_{i,j} + \theta \Delta_y u_{i+1,j} \right] \\
&= \frac{h_y}{6} \left\{ \theta [J_{i+1,j+1}^x + J_{i+1,j-1}^x - J_{i-1,j-1}^x - J_{i-1,j+1}^x \right. \\
&\quad + 4(J_{i+1/2,j-1}^x + J_{i+1/2,j+1}^x - J_{i-1/2,j-1}^x - J_{i-1/2,j+1}^x)] \\
&\quad \left. + (6-2\theta) [J_{i+1,j}^x - J_{i-1,j}^x + 4(J_{i+1/2,j}^x - J_{i-1/2,j}^x)] \right\}
\end{aligned}$$

$$\begin{aligned}
& + \frac{h_x}{6} \left\{ \theta [J_{i-1,j+1}^y + J_{i+1,j+1}^y - J_{i-1,j-1}^y - J_{i+1,j-1}^y \right. \\
& + 4(J_{i-1,j+1/2}^y + J_{i+1,j+1/2}^y - J_{i-1,j-1/2}^y - J_{i+1,j-1/2}^y)] \quad (4.53) \\
& \left. + (6 - 2\theta)[J_{i,j+1}^y - J_{i,j-1}^y + 4(J_{i,j+1/2}^y - J_{i,j-1/2}^y)] \right\} + \psi_{i,j}, \\
\psi_{i,j} = & \left[\theta(\psi_{i+1/2,j-1}^x + \psi_{i+1/2,j+1}^x - \psi_{i-1/2,j-1}^x - \psi_{i-1/2,j+1}^x) \right. \\
& + (6 - 2\theta)(\psi_{i+1/2,j}^x - \psi_{i-1/2,j}^x) \left. \right] \frac{h_y}{h_x} \\
& + \left[\theta(\psi_{i-1,j+1/2}^y + \psi_{i+1,j+1/2}^y - \psi_{i-1,j-1/2}^y - \psi_{i+1,j-1/2}^y) \right. \\
& \left. + (6 - 2\theta)(\psi_{i,j+1/2}^y - \psi_{i,j-1/2}^y) \right] \frac{h_x}{h_y},
\end{aligned}$$

где θ – пока произвольный весовой множитель.

Обозначим через $V_{i,j} = \{x_{i-1} \leq x \leq x_{i+1}, y_{j-1} \leq y \leq y_{j+1}\}$ и $V'_{i,j} = \{x_{i-1/2} \leq x \leq x_{i+1/2}, y_{j-1/2} \leq y \leq y_{j+1/2}\}$ прямоугольные ячейки со сторонами $S_1 = \{x = x_{i-1}\}$, $S_2 = \{y = y_{j-1}\}$, $S_3 = \{x = x_{i+1}\}$, $S_4 = \{y = y_{j+1}\}$ и $S'_1 = \{x = x_{i-1/2}\}$, $S'_2 = \{y = y_{j-1/2}\}$, $S'_3 = \{x = x_{i+1/2}\}$, $S'_4 = \{y = y_{j+1/2}\}$ соответственно (здесь и ниже в обозначениях сторон мы опускаем очевидные из контекста индексы i, j).

Выпишем теперь линейную комбинацию интегральных соотношений баланса по объемам $V_{i,j}$ и $V'_{i,j}$ с весовыми множителями θ_1 и θ_2 :

$$\begin{aligned}
\theta_1 \sum_{k=1}^4 \int_S J^n dS_k + \theta_2 \sum_{k=1}^4 \int_{S'} J^n dS'_k \\
= \theta_1 \int_{V_{i,j}} f(x, y) dx dy + \theta_2 \int_{V'_{i,j}} f(x, y) dx dy, \quad (4.54)
\end{aligned}$$

где $S = \bigcup_{k=1}^4 S_k$, $S' = \bigcup_{k=1}^4 S'_k$, а J^n есть поток по направлению внешней нормали к стороне S_k или S'_k , который берется по контексту как $\pm J^x$ или $\pm J^y$.

Дальнейшая задача состоит в том, чтобы оценить, как поверхностные интегралы в (4.54) могут быть аппроксимированы членами из правой части равенства (4.53). Остановимся для иллюстрации на стороне S_1 , для которой с помощью формулы Симпсона получаем

$$\int_{S_1} J^x dy = \frac{h_y}{3} (J_{i-1,j-1}^x + 4J_{i-1,j}^x + J_{i-1,j+1}^x) - \psi_{i-1,j}^y,$$

$$\psi_{i-1,j}^y = \frac{h_y^5}{90} \left[\frac{\partial^4 J^x(x_{i-1}, y_j)}{\partial y^4} + O(h) \right].$$

Аналогично выражаются интегралы по остальным сторонам большой ячейки $V_{i,j}$. Например, для S_2 имеем

$$\int_{S_2} J^y dx = \frac{h_x}{3} (J_{i-1,j-1}^y + 4J_{i,j-1}^y + J_{i+1,j-1}^y) - \psi_{i,j-1}^x,$$

$$\psi_{i,j-1}^x = \frac{h_x^5}{90} \left[\frac{\partial^4 J^y(x_i, y_{j-1})}{\partial x^4} + O(h) \right].$$

В то же время можно установить связь с формулой Симпсона и некоей другой квадратурной формулой с параметром θ :

$$\begin{aligned} & \frac{h_y}{3} \left[\theta J_{i-1,j-1}^x + (6-2\theta) J_{i-1,j}^x + \theta J_{i-1,j+1}^x \right] \\ &= \frac{h_y}{3} (J_{i-1,j-1}^x + 4J_{i-1,j}^x + J_{i-1,j+1}^x) - \bar{\psi}_{i-1,j}^y, \\ \bar{\psi}_{i-1,j}^y &= (1-\theta) \frac{h_y}{3} (J_{i-1,j-1}^x - 2J_{i-1,j}^x + J_{i-1,j+1}^x) \\ &= (1-\theta) \frac{h_y^3}{3} \left[\frac{\partial^2 J^x(x_{i-1}, y_j)}{\partial y^2} + \frac{h_y^2}{12} \frac{\partial^4 J^x(x_{i-1}, \eta'_j)}{\partial y^4} \right], \quad \eta'_j \in [y_{j-1}, y_{j+1}], \\ & \frac{h_x}{3} \left[\theta J_{i-1,j-1}^y + (6-2\theta) J_{i,j-1}^y + \theta J_{i+1,j-1}^y \right] \\ &= \frac{h_x}{3} (J_{i-1,j-1}^y + 4J_{i,j-1}^y + J_{i+1,j-1}^y) - \bar{\varphi}_{i,j-1}^x, \\ \bar{\varphi}_{i,j-1}^x &= (1-\theta) \frac{h_x}{3} (J_{i-1,j-1}^y - 2J_{i,j-1}^y + J_{i+1,j-1}^y) \\ &= (1-\theta) \frac{h_x^3}{3} \left[\frac{\partial^2 J^y(x_i, y_{j-1})}{\partial x^2} + \frac{h_x^2}{12} \frac{\partial^4 J^y(\xi'_i, y_{j-1})}{\partial x^4} \right], \quad \xi'_i \in [x_{i-1}, x_{i+1}]. \end{aligned} \quad (4.55)$$

Теперь аналогично рассмотрим, как члены с полуцелыми индексами в правой части (4.53) аппроксимируют интегралы от потоков по сторонам малой ячейки $V'_{i,j}$. Для сторон S'_1 и S'_2 соответственно имеем

$$\begin{aligned} & \frac{2h_y}{3} \left[\theta J_{i-1/2,j-1}^x + (6-2\theta) J_{i-1/2,j}^x + \theta J_{i-1/2,j+1}^x \right] \\ &= \frac{2h_y}{3} (J_{i-1/2,j-1/2}^x + 4J_{i-1/2,j}^x + J_{i-1/2,j+1/2}^x) \\ &+ \frac{2h_y}{3} \left[\theta J_{i-1/2,j-1}^x - J_{i-1/2,j-1/2}^x + (2-2\theta) J_{i-1/2,j}^x \right] \end{aligned}$$

$$\begin{aligned}
& -J_{i-1/2,j+1/2}^x + \theta J_{i-1/2,j+1}^x \Big] = 4 \int_{S'_1} J^x dy - \psi'_{i-1/2,j} + \psi_{i-1/2,j}^x, \\
& \psi'_{i-1/2,j} = \frac{h_y^5}{16 \cdot 90} \left[\frac{\partial^4 J^x(x_{i-1/2}, y_j)}{\partial y^4} + O(h) \right], \\
& \psi_{i-1/2,j}^x = \frac{2h_y^3}{3} \left[(\theta - 1/4) \frac{\partial^2 J^x(x_{i-1/2}, y_j)}{\partial y^2} + \theta \frac{h_y^2}{12} \frac{\partial^4 J^x(x_{i-1/2}, \hat{\eta}_j)}{\partial y^4} \right. \\
& \quad \left. - \frac{h_y^2}{192} \frac{\partial^4 J^x(x_{i-1/2}, \check{\eta}_j)}{\partial y^4}, \hat{\eta}_j, \check{\eta}_j \in [y_{j-1}, y_{j+1}] \right], \\
& \frac{2h_x}{3} \left[\theta J_{i-1,j-1/2}^y + (6 - 2\theta) J_{i,j-1/2}^y + \theta J_{i+1,j-1/2}^y \right] \\
& = \frac{2h_x}{3} (J_{i-1/2,j-1/2}^y + 4J_{i,j-1/2}^y + J_{i+1/2,j-1/2}^y) \\
& + \frac{2h_x}{3} \left[\theta J_{i-1,j-1/2}^y - J_{i-1/2,j-1/2}^y + (2 - 2\theta) J_{i,j-1/2}^y \right. \\
& \quad \left. - J_{i+1/2,j-1/2}^y + \theta J_{i+1,j-1/2}^y \right] = 4 \int_{S'_2} J^y dx - \psi'_{i,j-1/2} + \psi_{i,j-1/2}^y, \\
& \psi'_{i,j-1/2} = \frac{h_x^5}{16 \cdot 90} \left[\partial^4 J^y(x_i, y_{j-1/2}) + O(h) \right], \\
& \psi_{i,j-1/2}^y = \frac{2h_x^3}{3} \left[\left(\theta - \frac{1}{4} \right) \frac{\partial^2 J^y(x_i, y_{j-1/2})}{\partial x^2} + \theta \frac{h_x^2}{12} \frac{\partial^4 J^y(\xi_i, y_{j-1/2})}{\partial x^4} \right], \\
& \quad \hat{\xi}_i, \check{\xi}_i \in [x_{i-1}, x_{i+1}].
\end{aligned} \tag{4.56}$$

Анализируя полученные выражения, несложно заметить, что если в (4.53)–(4.56) мы положим $\theta = \frac{1}{2}$, то правая часть (4.53) будет аппроксимировать (на любой достаточно гладкой функции) левую часть балансного соотношения (4.54) при $\theta_1 = \frac{1}{2}$ и $\theta_2 = 4$, т. е. справедливо равенство

$$\begin{aligned}
& -\frac{1}{6} \left[5(\Delta_x u + \Delta_y u)_{i,j} + \frac{1}{12} (\Delta_x u_{i,j-1} + \Delta_x u_{i,j+1} + \Delta_y u_{i-1,j} + \Delta_y u_{i+1,j}) \right] \\
& = \frac{1}{6h_x h_y} \left(\frac{1}{2} \int_S J^n dS' + 4 \int_{S'} J^n dS' + \psi_{i,j} + \hat{\psi}_{i,j} \right) \\
& = \frac{1}{6h_x h_y} \left[\frac{1}{2} \int_{V_{i,j}} f(x, y) dx dy + 4 \int_{V'_{i,j}} f(x, y) dx dy + \psi_{i,j} + \hat{\psi}_{i,j} \right] \\
& = \frac{1}{6h_x h_y} \left(\sum_{i'} \sum_{j'} q_{i'j'} f_{i',j'} + \psi_{i,j} + \hat{\psi}_{i,j} + \check{\psi}_{i,j} \right).
\end{aligned} \tag{4.57}$$

В последней скобке двойная сумма представляет "объединенную" формулу численного интегрирования функции $f(x, y)$ (если она не интегрируется точно) по ячейкам $V_{i,j}$ и $V'_{i,j}$, ее коэффициенты $q_{i',j'}$, квадратурные узлы и пределы изменения индексов мы не конкретизируем, поскольку здесь возможен различный выбор, а их частный вид нас в данном случае не интересует. Величина $\hat{\psi}_{i,j}$ есть соответствующая погрешность интегрирования, и мы считаем $\hat{\psi}_{i,j} = O(h^6)$, что всегда возможно при достаточной гладкости $f(x, y)$.

Переменная $\psi_{i,j}$ в (4.57) определена выражением (4.53), и ее отдельные компоненты представляют ошибки квадратурных формул Симпсона (4.51); нетрудно проверить, что $\psi_{i,j} = O(h^6)$ при достаточной гладкости функции потока.

Наконец, величина $\hat{\psi}_{i,j}$ включает все описанные выше в (4.55), (4.56) погрешности интегрирования по сторонам S_k и S'_k :

$$\begin{aligned} \hat{\psi}_{i,j} = & \psi_{i,j-1}^x - \psi_{i,j+1}^x + \psi_{i-1,j}^y - \psi_{i+1,j}^y - \psi_{i,j-1/2}^x + \psi_{i,j+1/2}^x \\ & - \psi_{i-1/2,j}^y + \psi_{i+1/2,j}^y - \psi'_{i-1/2,j} + \psi'_{i+1/2,j} - \psi'_{i,j-1/2} + \psi'_{i,j+1/2}. \end{aligned}$$

Из вида отдельных слагаемых следует, что при выбранном значении весового множителя все члены со вторыми производными от потоков после несложных преобразований взаимно сокращаются, а при непрерывности четвертых производных оказывается $\hat{\psi}_{i,j} = O(h^6)$.

Таким образом, отбрасывая в (4.57) все остаточные члены, мы можем сделать вывод о существовании балансной разностной схемы с погрешностью $O(h^4)$, которую запишем в форме

$$-\left(\Delta_x + \Delta_y + \frac{h_x^2}{12}\Delta_x\Delta_y + \frac{h_y^2}{12}\Delta_y\Delta_x\right)v_{i,j} = \frac{1}{6h_xh_y} \sum_{i'N'} q_{i',j'} f'_{i',j'}. \quad (4.58)$$

Сравнивая это уравнение с конечно-разностной аппроксимацией Микеладзе (3.41), легко заметить, что они совпадают при перестановочности операторов Δ_x и Δ_y и выборе соответствующих формул в правой части в (4.58).

Надо отметить, что при переходе к уравнениям с переменными коэффициентами реализация рассмотренных балансных аппроксимаций повышенной точности технически существенно усложняется, и несколько другой подход мы опишем в п. 4.5.

4.3. Аппроксимации на элементно-ориентированных сетках

В этом параграфе будут рассмотрены фактически такие же интегро-балансные аппроксимации, как и в предыдущем. Главное отличие носит (в некотором смысле) философско-методологический оттенок: что при построении сетки считать первичным – точки (узлы) или ячейки (конечные объемы, элементы).

В первом случае (*точечно-ориентированные сетки*, рассмотренные выше) конструирование сетки предполагает сначала построение множества узлов, а затем – при использовании балансных аппроксимаций – формирование около каждого из них соответствующей ячейки. Альтернативный вариант (*элементно-ориентированные сетки*) предполагает первоначальное разбиение расчетной области на конечные объемы, после чего внутри каждого из них определяется сеточный узел. Иногда эти различающиеся изначально подходы дают одинаковые итоговые результаты, но могут наблюдаться и существенные отличия.

Сеточные методы на элементно-ориентированных сетках сейчас получают все большее распространение и обнаруживают необычные свойства: при ухудшении аппроксимационных свойств в “классическом” смысле они обеспечивают точность $O(h^2)$ даже на неравномерных сетках. Впервые этот факт доказан для одномерных краевых задач в монографии В. В. Смелова [167].

4.3.1. Аппроксимация одномерного уравнения диффузии

В качестве простейшего примера рассмотрим балансную аппроксимацию одномерного уравнения диффузии (4.1) при $r(x) = 0$ на элементно-ориентированной сетке. Пусть h_i (рис. 4.1) означает конечный отрезок $[x_{i-1/2}, x_{i+1/2}]$, длина которого имеет то же обозначение, а узел x_i располагается в центре этого отрезка.

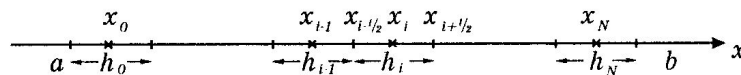


Рис. 4.1. Одномерная элементно-ориентированная сетка

Функцию $p(x)$ мы допускаем на интервале $[a, b]$ кусочно-непрерывной, но точки ее разрыва предполагаются расположенными только на границах $x_{i+1/2} = x_i + \frac{h_i}{2} = x_{i+1} - \frac{h_{i+1}}{2}$, $x_{i-1/2} = x_i - \frac{h_i}{2} = x_{i-1} + \frac{h_{i-1}}{2}$ отрезков h_i . При этом в таких точках должны выполняться условия сопряжения

$$p^+ \frac{du^+}{dx} \Big|_{x_{i\pm 1/2}} = p^- \frac{du^-}{dx} \Big|_{x_{i\pm 1/2}}, \quad u_{i\pm 1/2}^+ = u_{i\pm 1/2}^-, \quad (4.59)$$

где верхние значки “ $\pm$ ” означают предельные значения функций справа и слева при стремлении к точкам $x_{i\pm 1/2}$.

Расстояние между соседними узлами определяется очевидным равенством $x_i - x_{i-1} = \frac{1}{2}(h_i + h_{i-1})$. Точное соотношение баланса на конечном объеме h_i принимает вид

$$-p^- \frac{du^-}{dx} \Big|_{x_i + \frac{h_i}{2}} + p^+ \frac{du^+}{dx} \Big|_{x_i - \frac{h_i}{2}} = \int_{x_{i-1/2}}^{x_{i+1/2}} f(x) dx. \quad (4.60)$$

Для получения разностных уравнений в правой части (4.60) проще всего применить квадратурную формулу центральных прямоугольников, которая имеет достаточно высокий – третий – порядок погрешности на функциях $f(x) \in C^2(h_i)$, даже при наличии разрывов в точках $x_{i\pm 1/2}$:

$$\int_{x_{i-1/2}}^{x_{i+1/2}} f(x) dx = h_i f_i + \psi_i^f, \quad |\psi_i^f| \leq \frac{h_i^3}{24} M_2^f, \\ M_2^f = \max_{x \in h_i} \left\{ \left| \frac{d^2 f}{dx^2} \right| \right\}. \quad (4.61)$$

С аппроксимацией левой части балансного соотношения (4.60) дело обстоит несколько сложнее, и для начала мы выпишем приближения для правой и левой производных в точке $x_{i-1/2}$:

$$\frac{du^+}{dx} \Big|_{i-1/2} = 2 \frac{u_i - u_{i-1/2}}{h_i} + \frac{h_i}{4} \frac{d^2 u}{dx^2}(\xi'_i), \quad \xi'_i \in [x_{i-1/2}, x_i], \\ \frac{du^-}{dx} \Big|_{i-1/2} = 2 \frac{u_{i-1/2} - u_{i-1}}{h_{i-1}} + \frac{h_{i-1}}{4} \frac{d^2 u}{dx^2}(\xi''_{i-1}),$$

$$\xi_{i-1}'' \in [x_{i-1}, x_{i-1/2}]. \quad (4.62)$$

Чтобы построить уравнения относительно значений решения в узлах, из последних равенств необходимо исключить величину $u_{i-1/2}$, что мы и сделаем с помощью условий сопряжения (4.59). После подстановки в них (4.62) и несложных преобразований получаем

$$\begin{aligned} u_{i-1/2} &= \frac{1}{\alpha_i + \beta_{i-1}} (\alpha_i u_i + \beta_{i-1} u_{i-1} - \psi_{i-1/2}), \\ \alpha_i &= \frac{2p_{i-1/2}^+}{h_i}, \quad \beta_{i-1} = \frac{2p_{i-1/2}^-}{h_{i-1}}, \\ \psi_{i-1/2} &= -p_{i-1/2}^+ \frac{h_i}{4} \frac{d^2 u}{dx^2}(\xi_i') + p_{i-1/2}^- \frac{h_{i-1}}{4} \frac{d^2 u}{dx^2}(\xi_{i-1}''). \end{aligned} \quad (4.63)$$

Отсюда после подстановки $u_{i-1/2}$ в (4.62) следует выражение

$$\begin{aligned} p^+ \frac{du^+}{dx} \Big|_{i-1/2} &= \alpha_i (u_i - u_{i-1/2}) - p_{i-1/2}^+ \frac{h_i}{4} \frac{d^2 u}{dx^2}(\xi_i') \\ &= \frac{\alpha_i \beta_{i-1}}{\alpha_i + \beta_{i-1}} (u_i - u_{i-1}) + \hat{\psi}_i, \\ \hat{\psi}_i &= \frac{\alpha_i}{\alpha_i + \beta_{i-1}} p_{i-1/2}^- \frac{h_{i-1}}{4} \frac{d^2 u}{dx^2}(\xi_{i-1}'') - \frac{\beta_{i-1}}{\alpha_i + \beta_{i-1}} p_{i-1/2}^+ \frac{h_i}{4} \frac{d^2 u}{dx^2}(\xi_i'). \end{aligned} \quad (4.64)$$

Таким же образом с помощью условий сопряжения в точке $x_{i+1/2}$ и аналогичных (4.62) аппроксимаций односторонних производных на стыке отрезков h_i, h_{i+1} получаем

$$\begin{aligned} p^- \frac{du^-}{dx} \Big|_{i+1/2} &= \beta_i (u_{i+1/2} - u_i) + p_{i+1/2}^- \frac{h_i}{4} \frac{d^2 u}{dx^2}(\xi_i'') \\ &= \frac{\alpha_{i+1} \beta_i}{\alpha_{i+1} + \beta_i} (u_{i+1} - u_i) + \check{\psi}_i, \\ \check{\psi}_i &= \frac{\alpha_{i+1}}{\alpha_{i+1} + \beta_i} p_{i+1/2}^- \frac{h_i}{4} \frac{d^2 u}{dx^2}(\xi_i'') - \frac{\beta_i}{\alpha_{i+1} + \beta_i} p_{i+1/2}^+ \frac{h_{i+1}}{4} \frac{d^2 u}{dx^2}(\xi_{i+1}'). \end{aligned} \quad (4.65)$$

Наконец, используя (4.61), (4.64) и (4.65), формируем дискретное представление балансного равенства (4.60):

$$\begin{aligned} -a_{i+1}(u_{i+1} - u_i) + a_i(u_i - u_{i-1}) &= h_i f_i + \psi_i, \\ a_i &= \frac{\alpha_i \beta_{i-1}}{\alpha_i + \beta_{i-1}}, \quad \psi_i = \hat{\psi}_i - \check{\psi}_{i-1} + \psi_i^f. \end{aligned} \quad (4.66)$$

Отсюда после отбрасывания погрешности аппроксимации получаем итоговое разностное уравнение

$$-a_{i+1}(v_{i+1} - v_i) + a_i(v_i - v_{i-1}) = h_i f_i. \quad (4.67)$$

Интересно сравнить полученный результат с “обычным” разностным уравнением, которое получается из балансного соотношения (4.3) (при $r \equiv 0$) на точечно-ориентированной элементной сетке, если производные в потоках $J_{i\pm 1/2}$ аппроксимировать центральными разностными, а интеграл от $f(x)$ – простейшей квадратурной формулой (на основе кусочно-постоянного представления функции по значениям в точках x_i):

$$-\frac{p(x_{i+1/2})}{x_{i+1} - x_i}(v_{i+1} - v_i) + \frac{p(x_{i-1/2})}{x_i - x_{i-1}}(v_i - v_{i-1}) = \frac{x_{i+1} - x_{i-1}}{2} f_i, \quad (4.61)$$

имеющей при данной нормировке погрешность $O(h^2)$.

Останавливаясь для простоты на непрерывных коэффициентах диффузии, сразу замечаем, что уравнения (4.67) и (4.61) при равенстве шагов сетки совпадают (с точностью до масштабирования). В противном же случае левые части обоих разностных уравнений по-прежнему одинаковы, но коэффициент при f_i в “классическом” уравнении (4.61), в обозначениях элементарно-ориентированной сетки на рис. 4.1, равен $(h_{i-1} + 2h_i + h_{i+1})/4$ (вместо h_i в (4.67)).

Главное же отличие сравниваемых разностных схем заключается в погрешностях аппроксимаций. Чтобы в этом убедиться, перепишем ψ_i из (4.66) в более наглядном виде

$$\psi_i = -\frac{1}{4} \left\{ a_{i+1} [h_{i+1}^2 u''(\xi'_i) - h_i^2 u''(\xi''_i)] - a_i [h_i^2 u''(\xi'_i) - h_{i-1}^2 u''(\xi''_{i-1})] \right\} + O(h^3). \quad (4.69)$$

Поскольку величины a_i имеют порядок $O(h^{-1})$, то нетрудно убедиться, что в общем случае $\psi_i = O(h)$ на неравномерной сетке, даже при гладких коэффициентах $p(x)$. А это означает, что если уравнение (4.66) поделить на h_i , т. е. отмасштабировать “обычным” образом, когда разностные коэффициенты становятся порядка $O(h^{-2})$, то оказывается, что полученная схема вообще не аппроксимирует исходное уравнение диффузии! Однако величина ψ_i/h_i при $h \rightarrow 0$ оказывается порядка $O(1)$, т. е. ограниченной. Это является очень существенным при анализе погрешности разностного решения (имеет значение и специальный вид погрешности ψ_i в (4.69)).

Остановимся теперь на особенностях аппроксимации краевых условий при использовании одномерных конечных объемов. Проще всего дело обстоит с условиями Неймана на границе. Если $u' = g$ в точках a и (или) b , то в балансных соотношениях (4.59) при $i = 0$ и (или) N надо соответственно значения $\frac{du}{dx} \Big|_{x_0 - \frac{h_0}{2}}$, и (или) $\frac{du}{dx} \Big|_{x_N + \frac{h_N}{2}}$ взять точно из краевых условий и перенести их в правую часть. В случае же условий 1-го или 3-го рода применяется подход с фиктивными точками: решение предполагается продолженным на полшага за границы расчетного отрезка и вводятся вспомогательные узлы $x_{-1} = a - \frac{h_0}{2}$, $x_{N+1} = b + \frac{h_{N+1}}{2}$. Значения решения в граничных точках затем линейно аппроксимируются: $u_a = \frac{u_{-1} + u_0}{2} + O(h^2)$, $u_b = \frac{u_N + u_{N+1}}{2} + O(h^2)$, и “лишние” переменные u_{-1} , u_{N+1} с помощью этих равенств исключаются из нулевого и N -го разностных уравнений.

4.3.2. Аппроксимация уравнения Пуассона на прямоугольной элементарно-ориентированной сетке

В случае двумерных областей с границами, параллельными координатным осям, техника остается аналогичной, что мы для простоты проиллюстрируем на уравнении Пуассона (3.37) в прямоугольнике. Расчетная область разбивается на прямоугольные ячейки (боксы) $V_{i,j}$ со сторонами длиной h_i^x , h_j^y , а в качестве узлов выбираются центры этих конечных объемов (рис. 4.2), координаты которых связаны соотношениями $x_i - x_{i-1} = \frac{1}{2}(h_i^x + h_{i-1}^x)$, $y_j - y_{j-1} = \frac{1}{2}(h_j^y + h_{j-1}^y)$.

Обозначая через

$$x_{i\pm 1/2} = x_i \pm \frac{h_i^x}{2}, \quad y_{j\pm 1/2} = y_j \pm \frac{h_j^y}{2} \quad (4.70)$$

координаты сторон бокса $V_{i,j}$, дискретную форму балансного соотношения (4.60) записываем в виде

$$\int_{y_{j-1/2}}^{y_{j+1/2}} \left(-\frac{du}{dx} \Big|_{x_{i+1/2}} + \frac{du}{dx} \Big|_{x_{i-1/2}} \right) dy$$

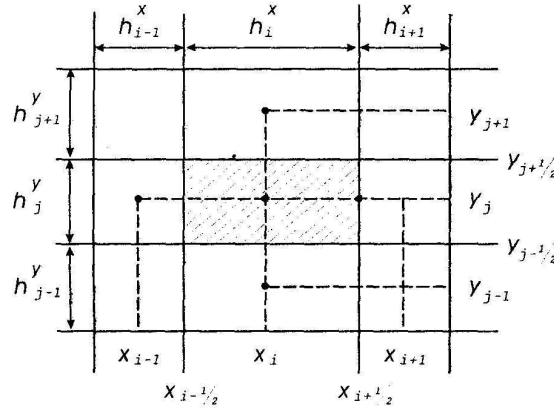


Рис. 4.2. Прямоугольная элементарно-ориентированная сетка

$$\begin{aligned}
 & + \int_{x_{i-1/2}}^{x_{i+1/2}} \left(-\frac{du}{dy} \Big|_{y_{j+1/2}} + \frac{du}{dy} \Big|_{y_{j-1/2}} \right) dx = \\
 & = \int_{x_{i-1/2}}^{x_{i+1/2}} \int_{y_{j-1/2}}^{y_{j+1/2}} f(x, y) dx dy. \quad (4.71)
 \end{aligned}$$

Сравнивая полученное равенство с его аналогом на точечно-ориентированной сетке в предыдущих параграфах, можно заметить их формальное совпадение. Различие заключается "только" в смысле обозначений $x_{i\pm 1/2}$, $y_{j\pm 1/2}$: если ранее это были точки посередине между парами x_i , x_{i+1} и y_j , y_{j+1} , то теперь в (4.71) они определяются из (4.63) – на равномерной сетке это одно и то же, но в общем случае имеется отличие.

Дискретизация интегрального соотношения баланса (4.71) может быть проведена аналогично одномерному случаю, но в силу сделанных здесь упрощений касательно гладкости решения $u(x, y)$ процедура может быть сокращена без изменения конечного результата.

Все интегралы аппроксимируются формулами центральных прямоугольников, а производные заменяются конечными разностями по близлежащим узлам, в результате чего получаем равенство

$$h_j^y \left[-\frac{2(u_{i+1,j} - u_{i,j})}{h_i^x + h_{i+1}^x} + \frac{2(u_{i,j} - u_{i-1,j})}{h_i^x + h_{i-1}^x} \right]$$

$$\begin{aligned}
& + h_i^x \left[-\frac{2(u_{i,j+1} - u_{i,j})}{h_j^y + h_{j+1}^y} + \frac{2(u_{i,j} - u_{i,j-1})}{h_j^y + h_{j-1}^y} \right] \\
& = h_i^x h_j^y f_{i,j} + \psi_{i,j},
\end{aligned} \tag{4.72}$$

в котором погрешность аппроксимации записывается в форме, подсказываемой выражением (4.69), а также несложно выводимой из (4.72) путем разложения в ряд Тейлора функций в окрестности узла (x_i, y_i) :

$$\begin{aligned}
\psi_{i,j} &= \hat{\psi}_{i-1/2} + \check{\psi}_{i+1/2} + \hat{\psi}_{i+1/2} + \check{\psi}_{i-1/2} + \hat{\psi}_{j-1/2} \\
&\quad + \check{\psi}_{j-1/2} + \hat{\psi}_{j+1/2} + \check{\psi}_{j+1/2} + \psi_{i,j}^f, \\
\hat{\psi}_{i\pm 1/2} &= \frac{(h_j^y)^3}{24} \frac{\partial^3 u}{\partial x \partial y^2}(x_{i\pm 1/2}, y_j) + O(h^4), \\
\hat{\psi}_{j\pm 1/2} &= \frac{(h_i^x)^3}{24} \frac{\partial^3 u}{\partial x^2 \partial y}(x_i, y_{j\pm 1/2}) + O(h^4), \\
\check{\psi}_{i\pm 1/2} &= \frac{h_j^y}{4} \left\{ \pm (h_{i\pm 1}^x - h_i^x) \frac{\partial^2 u}{\partial x^2}(x_{i\pm 1/2}, y_j) \right. \\
&\quad \left. + \frac{1}{6} \left[(h_i^x)^2 - h_i^x h_{i\pm 1}^x + (h_{i\pm 1}^x)^2 \right] \frac{\partial^3 u}{\partial x^3}(x_{i\pm 1/2}, y_j) \right\} + O(h^4), \\
\check{\psi}_{j\pm 1/2} &= \frac{h_i^x}{4} \left\{ \pm (h_{j\pm 1}^y - h_j^y) \frac{\partial^2 u}{\partial y^2}(x_i, y_{j\pm 1/2}) \right. \\
&\quad \left. + \frac{1}{6} \left[(h_j^y)^2 - h_j^y h_{j\pm 1}^y + (h_{j\pm 1}^y)^2 \right] \frac{\partial^3 u}{\partial y^3}(x_i, y_{j\pm 1/2}) \right\} + O(h^4), \\
\psi_{i,j}^f &= \frac{1}{24} h_j^x h_j^y \left[(h_i^x)^2 \frac{\partial^2 f}{\partial x^2} + (h_j^y)^2 \frac{\partial^2 f}{\partial y^2} \right]_{i,j} + O(h^5).
\end{aligned} \tag{4.73}$$

Здесь величины $\psi_{i\pm 1/2}, \psi_{j\pm 1/2}$ суть погрешности формул центральных прямоугольников для вычисления интегралов по сторонам бокса с координатами $x = x_{i\pm 1/2}, y = y_{j\pm 1/2}$, переменные $\check{\psi}_{i\pm 1/2}, \check{\psi}_{j\pm 1/2}$ — ошибки приближения производных, стоящих в (4.71) под знаком интеграла, а $\psi_{i,j}^f$ — ошибка квадратурной формулы для интеграла от функции $f(x, y)$ по конечному объему $v_{i,j}$.

Если обе части уравнения (4.72) поделить на объем ячейки $V_{i,j} = h_i^x h_j^y$, т. е. отмасштабировать к обычному для разностных уравнений виду, то величина $\psi_{i,j}/h_i^x h_j^y$ окажется в общем случае на неравномерной сетке порядка $0(1)$, т. е. исходное уравнение Пуассона формально не аппроксимируется, как и уравнение диффузии в предыдущем пункте.

Получаемое из (4.72) отбрасыванием погрешности $\psi_{i,j}$ дискретное уравнение

$$h_j^y \left[-\frac{2(v_{i+1,j} - v_{i,j})}{h_i^x + h_{i+1}^x} + \frac{2(v_{i,j} - v_{i-1,j})}{h_i^x + h_{i-1}^x} \right] + h_i^x \left[-\frac{2(v_{i,j+1} - v_{i,j})}{h_j^y + h_{j+1}^y} + \frac{2(v_{i,j} + v_{i,j-1})}{h_j^y + h_{j-1}^y} \right] = h_i^x h_j^y f_{i,j} \quad (4.74)$$

в случае равномерной сетки (не обязательно квадратной) совпадает, как легко убедиться, с обычным конечно-разностным и, следовательно, имеет ошибку аппроксимации (при обычной нормировке) порядка $O(h^2)$.

Построение уравнений для конечных объемов вблизи границы с учетом краевых условий делается по аналогии с одномерным случаем (для условий Дирихле используются фиктивные точки), если граница расчетной области состоит из прямолинейных отрезков, параллельных координатным осям. Если же граница содержит наклонные прямые или криволинейные участки, то надо переходить от прямоугольных сеточных ячеек к более универсальным (в геометрическом смысле) треугольным, что мы и сделаем в следующем параграфе.

4.4. Треугольные конечные объемы

Построение *треугольной сетки* в двумерной области, или, другими словами, разбиение ее на треугольные ячейки, получило название *триангуляции области*. Это сама по себе достаточно содержательная математическая задача возникла и активно решалась в методах конечных элементов. По этой теме имеется обширная специальная литература (см., например, [?], [26]), но мы такими вопросами заниматься не будем, ограничившись минимально необходимыми по ходу изложения сведениями.

Предполагается, что вся область Ω разбита на множество непересекающихся треугольников. Полученная сетка определяется как триединая совокупность треугольных элементов, их вершин и ребер. Каждое ребро или принадлежит двум смежным элементам, или лежит на границе Γ расчетной области. Граница Γ может быть гладкой или кусочно-гладкой – в последнем случае ее угловые точки обязаны являться узлами сетки. Если какой-то участок границы является криволинейным, то он ап-

проксимируется ломаной (мы рассматриваем простейший подход). На границу наносятся сначала точки (узлы сетки), которые затем соединяются последовательно прямолинейными отрезками. По этой причине в дальнейшем можно рассматривать только полигональные области, составленные из многогранников.

Среди всевозможных треугольных сеток своими положительными качествами выделяется так называемая *триангуляция Делоне*. Это такое разбиение области на треугольники, что описанная вокруг каждого из них окружность не содержит внутри себя никакого другого узла сетки. Триангуляция Делоне обладает таким важным свойством, что каждый получаемый треугольник полностью покрывается непересекающимися *ячейками Дирихле–Вороного*, каждая из которых строится вокруг узла сетки путем проведения перпендикуляров через середины ребер (на рис. 4.3 она представлена заштрихованным шестигранником).

Хотя в принципе триангуляция Делоне допускает и наличие тупых углов, мы ограничимся более простым случаем – с точки зрения геометрических построений – остроугольных треугольников. Существенным в данном случае оказывается то, что точка пересечения перпендикуляров к серединам сторон (центр описанной окружности) лежит внутри треугольника и почти все необходимые преобразования могут быть выполнены поэлементно, т. е. независимо на каждом треугольнике, с заключительной “сборкой” локальных величин на последнем этапе, когда непосредственно вычисляются коэффициенты разностных уравнений для узлов сетки.

Отметим, что в данном случае рассматриваются точечно-ориентированные сетки, когда конечные объемы формируются около первоначально найденных узлов. Ячейки же Дирихле будут играть роль конечных элементов, или боксов.

Пусть индекс k означает номер внутреннего узла сетки с координатами (x_k, y_k) , V_k – площадь D_k , т. е. соответствующей ему ячейки Дирихле (бокса), i_k – количество примыкающих к k -му узлу треугольников, S_i , $i = 1, 2, \dots, i_k$, – стороны ячейки D_k , а $\vec{n}$ – внешнюю нормаль к ячейке. Рассмотрим опять уравнение диффузии вида (4.23), для которого при $\alpha = 0$ (декартова система координат) точное интегральное соотношение баланса на объеме D_k записывается в форме

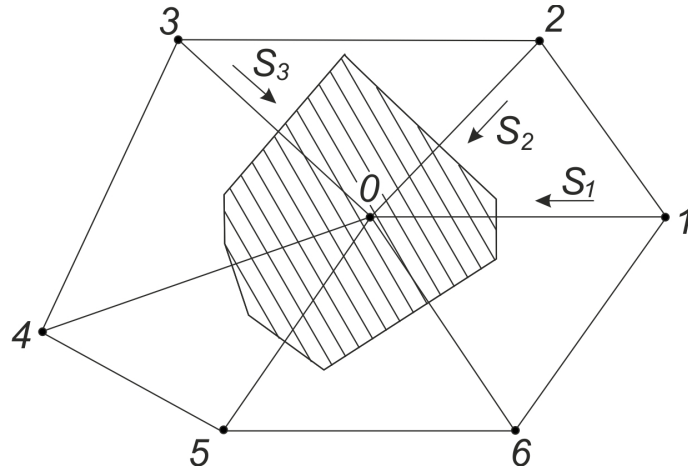


Рис. 4.3. Ячейка Дирихле на треугольной сетке

$$-\sum_{i=1}^{i,k} \int_{S_i} \frac{\partial u}{\partial \vec{n}} p ds = \int_{D_k} f(x, y) dv. \quad (4.75)$$

Обозначим далее через T_j , $j = 1, 2, \dots, i_k$, треугольники, имеющие своей вершиной k -й узел, и одного из их представителей с локальными номерами вершин 1, 2, 3 изобразим на рис. 4.4.

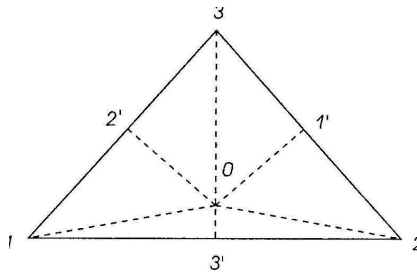


Рис. 4.4. К построению локальных балансов

Обозначая через $J_i = - \int_{S_i} p \frac{\partial u}{\partial \vec{n}} ds$ поток через сторону S_i и через $V_{k,j}$ — часть ячейки V_k , находящуюся в треугольнике T_j , мы можем переписать соотношение (4.75) в виде

$$\sum_{i=1}^{i_k} J_i = \sum_{i=1}^{i_k} \int_{V_{k,i}} f(x, y) dv = \sum_{i=1}^{i_k} f_{k,i}. \quad (4.76)$$

Как видно из рисунков, каждая из сторон S_i и каждый из объемчиков $V_{k,i}$ разбиваются на две части ($S_i = S'_i + S''_i$, $V_{k,i} = V'_{k,i} + V''_{k,i}$), принадлежащие разным треугольникам, и соответственно этому разбиваются в (4.76) величины J_i и $f_{k,i}$:

$$J_i = J'_i + J''_i, \quad f_{k,i} = f'_{k,i} + f''_{k,i}. \quad (4.77)$$

Слагаемые J'_i и J''_i уже относятся к отрезкам перпендикуляров к серединам ребер, находящихся в каждом смежном треугольнике T_j , а $V'_{k,i}$ и $V''_{k,i}$ – вспомогательные треугольники, на которые делится T_j отрезками, соединяющими центр описанной окружности с его вершинами и серединами сторон.

Задачу дискретизации интегрального баланса мы фактически разложили на модули: необходимо аппроксимировать элементарные величины J'_i , J''_i и $f'_{k,i}$, $f''_{k,i}$, а потом произвести сборку частичных результатов, т. е. суммирование в соответствии с (4.75)–(4.77).

В процессе аппроксимации локальных балансов типа (4.75) для всех ячеек V_k , очевидно, к каждому индивидуальному треугольнику T_j будет трехкратное “обращение последовательно при обработке каждой его вершины.

Такой вычислительный процесс целесообразно организовать наоборот, взяв за основу полный анализ каждого треугольного элемента и соотнося при этом вычисляемые величины к узлам сетки, являющимся его вершинами.

При следовании такой поэлементной идеологии для примера на рис. 4.4 нам необходимо аппроксимировать величины

$$J_{j,k'} = - \int_{0,k'} p \frac{\partial u}{\partial n} ds, \quad f_{j,k,k'} = \int_{\Delta_{i,k,k'}} f(x, y) dv, \quad (4.78)$$

где индексы j и k' означают номера треугольника и точки – середины ребра, противоположного k -й вершине элемента T_j , 0 , k' – отрезок длины $l_{0,k'}$, соединяющий точки с номерами 0 и k' , а $\Delta_{j,k,k'}$ – “маленький” треугольник, имеющий вершинами центр описанной окружности, один из узлов сетки и середину одного из ребер.

Проведем элементарные вычисления, которые не будем чрезмерно формализовать, воспользовавшись обозначениями на рис. 4.4:

$$J_{j3'} = l_{o,3'} \frac{u_2 - u_1}{l_{12}} p_{3'} + \psi_{j,3'}, \quad \psi_{j,2'} = \hat{\psi}_{j,3'} + \check{\psi}_{j,3'}. \quad (4.79)$$

Погрешности аппроксимации $\hat{\psi}_{j,3'}$ и $\check{\psi}_{j,3'}$ отвечают соответственно за применение квадратурной формулы односторонних прямоугольников ($\hat{\psi}_{j,3'} = O(h^2)$) и замену нормальной производной центральной разностью ($\check{\psi}_{j,3'} = O(h^3)$) с учетом наличия множителя $l_{o,3'} = O(h)$. Естественно, считаем, как и везде, что при сгущении сетки все характерные расстояния между соседними узлами являются величинами одного порядка $O(h)$.

Что касается приближенного вычисления интеграла от функции $f(x, y)$ по площади треугольника $\Delta_{j,k,k'}$, то здесь можно поступить двояким образом. Во-первых, если подынтегральную функцию линейно проинтерполировать по ее значениям в вершинах треугольника T_j с погрешностью $O(h^2)$, а затем точно проинтегрировать билинейную функцию, то ошибка будет порядка $O(h^4)$. Соответствующая формула будет иметь вид $f_{j,k,k'} = \sum_{i=1}^3 \alpha_i f_i + O(h^4)$, где коэффициенты α_i несложно выражаются через координаты вершин треугольников. Во-вторых, можно поступить и проще: положить функцию равной постоянной и тогда будем иметь

$$\int_{\Delta_{j,k,k'}} f(x, y) dv = \Delta_{j,k,k'} f_k + \psi_j^f, \quad \psi_j^f = O(h^3), \quad (4.80)$$

что тоже допустимо, поскольку величина ψ_j^f оказывается по порядку не хуже чем $\check{\psi}_{j,3'}$.

Суммируя теперь все частичные интегралы в балансном равенстве (4.71) и отбрасывая получаемую полную погрешность аппроксимации ψ_k , получаем в итоге дискретное балансное уравнение вида

$$(Au)_k = a_{k,k} u_k - \sum_{k,i}^{i_k} a_{i,k} u_i = V_k f_k, \quad (4.81)$$

где $A = \{a_{k,i}\}$, $u = \{u_k\}$, причем коэффициенты $a_{k,i}$ выражаются простыми формулами

$$a_{k,i} = p_{k,i} S_i / l_{k,i}, \quad a_{k,k} = \sum_{i=1}^{i,k} a_{k,i} \quad (4.82)$$

($p_{k,i}$ – значение $p(x, y)$ в точке, расположенной в середине отрезка $l_{k,i}$, а правая часть вычислена для простоты в соответствии с (4.80).

Относительно погрешности аппроксимации ψ_k уравнения (4.81) полезно сделать несколько замечаний. Ее величина складывается из компонент, относящихся к разным треугольникам и имеющих зачастую разные знаки, что при суммировании может значительно улучшить конечный результат (до сих пор возможность взаимного сокращения слагаемых мы не учитывали). Если сетка равномерная и все треугольники одинаковые, то поверхностный интеграл по стороне S_i будет вычисляться по формуле центральных прямоугольников с ошибкой $O(h^3)$ – при этом в (4.70) поток J_i не надо разбивать на части J'_i и J''_i , что приводит к формуле (4.79), а вместо нее использовать выражение $J_{i,k} = p_{k,i} S_i \frac{u_i - u_k}{l_{k,i}} + O(h^3)$.

Кроме того, в этом случае повышается и порядок погрешности приближенного интегрирования по площади ячейки: если ячейка симметрична относительно ее центра, то легко проверить справедливость соотношения

$$\int_{V_k} f dv = V_k f_k + O(h^4).$$

Однако в общем случае $\psi_k = O(h^2)$, а это означает, что получаемое из (4.81) разностное уравнение

$$a_{k,k} v_k - \sum_{i=1}^{i,k} a_{k,i} v_i = V_k f_k \quad (4.83)$$

после масштабирования его к “обычному” виду, т. е. деления на площадь $V_k = O(h^2)$, после чего коэффициенты $a_{k,i}$ становятся порядка $O(h^{-2})$, перестает формально аппроксимировать исходное дифференциальное уравнение диффузии, так как $\psi_k / V_k = O(1)$. Обнадеживающим при этом является тот факт, что “наихудшие” компоненты погрешности в соседние узлы входят с разными знаками, поскольку на общей стороне S_i двух примыкающих боксов поток J_i вычисляется одинаково (и с идентичной

ошибкой), а в соответствующие балансные соотношения он входит с разными знаками. При решении же разностных уравнений можно надеяться, что осциллирующие вклады погрешностей взаимно "почти" аннулируются.

Краевые условия 1-го и 2-го рода в описанном подходе реализуются особенно просто ввиду того, что пересечения ребер сетки с границей всегда являются узлами. Если в граничном узле задано условие Дирихле, то для него аппроксимацию строить совсем не надо, а просто в уравнениях для соседних околограничных узлов заданные значения решения учитываются точно. Для узла, лежащего на стыке двух граничных ребер, на которых заданы условия Неймана, расчетные ячейки формируются только по лежащим внутри области треугольникам, а балансные соотношения включают интеграл от потока на границе, который из краевых условий известен точно.

Что касается алгебраических свойств рассматриваемых уравнений вида (4.82), то они легко просматриваются: уравнения являются симметричными ($a_{k,i} = a_{i,k}$) и имеют положительный тип.

Подчеркнем, что описанный принцип построения интегробалансных разностных уравнений обладает большой универсальностью, так как применим для широкого класса дифференциальных уравнений, краевых условий и конфигураций границы (в том числе для трехмерных областей – в последнем случае роль треугольников играют тетраэдры).

Остановимся теперь на таком важном вопросе математической технологии, как формирование алгебраической системы (4.81) на основе упомянутого выше поэлементного подхода.

Введем в рассмотрение векторы третьего порядка

$$u^{(j)} = \begin{bmatrix} u_1^{(j)} \\ u_2^{(j)} \\ u_3^{(j)} \end{bmatrix}, \quad J^{(j)} = \begin{bmatrix} J_1^{(j)} \\ J_2^{(j)} \\ J_3^{(j)} \end{bmatrix},$$

где $u_k^{(j)}$, $k = 1, 2, 3$, – значения неизвестных в вершинах треугольника T_j , а $J_k^{(j)}$ – поток, соответствующий k -му узлу и вычисляемый через интегралы по отрезкам $0k'$, $0k''$, $k' \neq k$, составляющим часть границы ячейки Дирихле, находящейся в T_j .

Очевидно, что после вычисления локальных потоков по формулам вида (4.69)–(4.72) связь между векторами $u^{(j)}$, $J^{(j)}$ может

быть записана в виде

$$J^{(j)} = A^{(j)} u^{(j)}, \quad A^{(j)} = \begin{bmatrix} a_{1,1}^{(j)} & a_{1,2}^{(j)} & a_{1,3}^{(j)} \\ a_{2,1}^{(j)} & a_{2,2}^{(j)} & a_{2,3}^{(j)} \\ a_{3,1}^{(j)} & a_{3,2}^{(j)} & a_{3,3}^{(j)} \end{bmatrix}, \quad (4.77)$$

где $A^{(j)}$ назовем *локальной матрицей баланса*.

Если A и u – “глобальные” матрица и вектор порядка N из (4.81), то мы можем записать

$$P_j u^{(j)} = u, \quad \bar{A}^{(j)} = P_j' A^{(j)} P_j, \quad (4.78)$$

где P_j – прямоугольная матрица размера $3 \times N$ с одной ненулевой (единичной) компонентой в каждом из трех столбцов, а $\bar{A}^{(j)}$ – “растянутая” до полного размера $N \times N$ локальная матрица баланса. Очевидно, что из последних и составляется полная матрица системы

$$A = \sum_j \bar{A}^{(j)}, \quad (4.79)$$

где сумма берется по всем треугольникам сетки.

Подчеркнем здесь ту очень простую математически, но чрезвычайно важную технологически сторону представлений (4.77)–(4.78), что они позволяют поэлементно, независимо друг от друга, вычислять коэффициенты локальных матриц баланса и собирать (ассемблировать) из них *глобальную матрицу системы*. Аналогично по каждому из треугольников T_j можно организовать и накопление величин правых частей сеточных уравнений, в том числе с учетом неоднородных граничных условий 2-го или 3-го рода. В совокупности это значительно облегчает программную реализацию алгоритмов, особенно для сложных краевых задач.

4.5. Поэлементная аппроксимация на прямоугольной сетке

Напомним, что в п. 4.2.2 мы построили на прямоугольной равномерной сетке интегро-балансную аппроксимацию уравнения Пуассона с порядком погрешности $O(h^4)$. Понятно, что такой результат можно получить только для равномерной сетки и постоянных коэффициентов. Однако интересно рассмотреть, как используемая ранее идея линейной комбинации балансов на боль-

ших и малых ячейках может быть эффективной в более общих случаях.

4.5.1. Постановка задачи и обозначения

Рассматривается следующая задача: найти решение уравнения

$$-\frac{\partial}{\partial x}\left(\lambda(x, y)\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(\lambda(x, y)\frac{\partial u}{\partial y}\right) + \mu u = f(x, y), \quad (x, y) \in \Omega, \quad (4.87)$$

в ограниченной области Ω , состоящей из подобластей Ω_k , $k = 1, \dots, N_s$, в каждой из которых заданы свои положительная константа λ_k и гладкие функции $\mu \geq 0$, f . Граница области состоит из внешней и внутренней границ: $\partial\Omega = \Gamma \cup G$. На внешней границе области $\Gamma = \Gamma_d \cup \Gamma_n$ заданы граничные условия в форме

$$u|_{\Gamma_d} = g, \quad \varkappa u + \frac{\partial u}{\partial n}\Big|_{\Gamma_n} = \gamma, \quad (4.88)$$

где g , $\varkappa$, γ – некоторые заданные функции, n – внешняя нормаль к Γ .

Внутренняя граница состоит из частей: $G = \bigcup G_{kl}$. Каждая G_{kl} есть линия раздела подобластей Ω_k и Ω_l , т. е. линия разрыва функции λ . Обозначим две стороны границы G_{kl} через G_k и G_l . На внутренних границах выполняются условия сопряжения

$$u|_{G_k} = u|_{G_l}, \quad \lambda_k \frac{\partial u}{\partial n}\Big|_{G_k} = \lambda_l \frac{\partial u}{\partial n}\Big|_{G_l}. \quad (4.89)$$

Будем рассматривать такую геометрию области Ω , когда ее граница многосвязна и все участки границы параллельны осям координат.

Предположим, что входные данные обеспечивают такую гладкость решения, какая будет требоваться в дальнейшем для существования производных требуемого порядка.

Для дискретизации задачи строится неравномерная прямоугольная объемлющая расчетную область Ω сетка

$$x_{i+1} = x_i + h_i^x, \quad y_{j+1} = y_j + h_j^y, \quad i = 1, \dots, M+1, \quad j = 1, \dots, N+1,$$

таким образом, что граница области пересекает сеточные линии только в узлах сетки.

Введем вокруг сеточного узла (i, j) два вида элементарных объемов – маленький и большой:

$$V_{i,j} = \{x_i - h_{i-1}^x/2 \leq x \leq x_i + h_i^x/2, \quad y_j - h_{j-1}^y/2 \leq y \leq y_j + h_j^y/2\},$$

$$\bar{V}_{i,j} = \{x_{i-1} \leq x \leq x_{i+1}, \quad y_{j-1} \leq y \leq y_{j+1}\}.$$

Рассмотрим линейную комбинацию с весовым параметром ω балансных соотношений по большому и маленькому объемам (для краткости опустим здесь и далее индексы (i, j)):

$$\begin{aligned} J &\equiv \omega \int_S -\lambda \frac{\partial u}{\partial n} dS + (1 - \omega) \int_{\bar{S}} -\lambda \frac{\partial u}{\partial n} dS \\ &= \omega \int_V (f - \mu u) dV + (1 - \omega) \int_{\bar{V}} (f - \mu u) dV, \end{aligned} \quad (4.90)$$

где S и $\bar{S}$ – поверхности объемов V и $\bar{V}$ соответственно.

Каждый из объемов $V_{i,j}$ и $\bar{V}_{i,j}$ для узла (i, j) , не лежащего на границе области, делится координатными линиями x_i и y_j на четыре подобъема V_k и $\bar{V}_k$ соответственно, в каждом из которых функция λ – постоянная. Обозначим ее значение в k -м подобъеме через λ_k . Тогда каждый интеграл в (4.90) можно представить в аддитивной форме, например:

$$J = J_1 + J_2 + J_3 + J_4. \quad (4.91)$$

В случае, когда узел лежит на границе области, в рассмотрении участвуют только те части объемов, которые находятся внутри области.

Далее рассмотрим аппроксимацию интегралов в одном подобъеме:

$$\bar{V}_3 = \{x_i < x < x_{i+1}, \quad y_j < y < y_{j+1}\},$$

$$V_3 = \{x_i \leq x \leq x_i + h_i^x/2, \quad y_j \leq y \leq y_j + h_j^y/2\},$$

имеющем локальные номера вершин 0, 3, 4, 8 (см. рис. 4.5).

4.5.2. Методы аппроксимации интегралов

Для построения различных аппроксимаций используем квадратурные формулы центральных прямоугольников и трапеций:

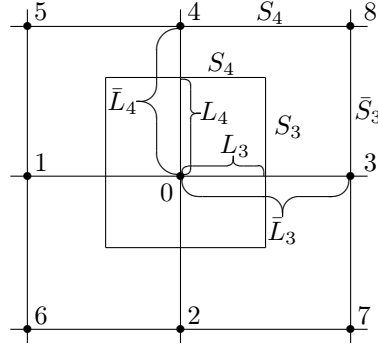


Рис. 4.5. Элементарные ячейки прямоугольной сетки

$$\begin{aligned}\lambda(u_{ij} - u_{i+1,j}) &= \int_{x_i}^{x_{i+1}} J^x(x, y_j) dx = h_i J_{i+1/2,j}^x + \psi_{i+1/2,j}^1 \\ &= \frac{h_i}{2} (J_{i,j}^x + J_{i+1,j}^x) + \psi_{i+1/2,j}^2,\end{aligned}$$

где остаточные члены представлены следующими соотношениями:

$$\psi_{i+1/2,j}^1 = \lambda \frac{h_i^3}{24} J_{xx}^x(\xi', y_j), \quad \psi_{i+1/2,j}^2 = -\lambda \frac{h_i^3}{12} J_{xx}^x(\xi'', y_j), \quad \xi', \xi'' \in [x_i, x_{i+1}].$$

Рассмотрим подробнее для примера аппроксимацию x -проекции потока J_3 вокруг узла 0:

$$J_3^x = -\lambda_3 \left(\omega \int_{y_j}^{y_{j+1/2}} u^x(x_{i+1/2}, y) dy + (1-\omega) \left(\int_{y_j}^{y_{j+1}} u^x(x_{i+1}, y) dy \right) \right). \quad (4.92)$$

Аппроксимируя здесь второй интеграл по формуле трапеций, а первый – по формуле центральных прямоугольников, с использованием легко проверяемого по ряду Тейлора представления

$$u_{i+1/2,j+1/4}^x = \frac{1}{4} (3u_{i+1/2,j}^x + u_{i+1/2,j+1}^x) - \frac{3}{32} h_j^2 u_{i+1/2,j+1/4}^{xyy} + O(h_j^3)$$

получаем следующее выражение:

$$\begin{aligned}J_3^x &= -\lambda_3 \frac{h_j}{8} \left[\omega (3u_{i+1/2,j}^x + u_{i+1/2,j+1}^x) + 4(1-\omega)(u_{i+1,j}^x + u_{i+1,j+1}^x) \right] \\ &\quad + \psi_3^x + O(h^4), \quad (4.93)\end{aligned}$$

где величина ψ_3^x имеет третий порядок:

$$\psi_3^x = -\lambda_3 \frac{h_j^3}{24} u_{i+1/2,j+1/4}^{xyy}. \quad (4.94)$$

Для повышения порядка погрешности на единицу вместо аппроксимации J_3^x будем приближать выражение

$$\begin{aligned} \bar{J}_3^x = & -\lambda_3 \left\{ \omega \int_{y_j}^{y_{j+1/2}} \left[u^x(x_{i+1/2}, y) - u^x(x_i, y) \right] dy \right. \\ & \left. + (1 - \omega) \int_{y_j}^{y_{j+1}} \left[u^x(x_{i+1}, y) - u^x(x_i, y) \right] dy \right\}. \end{aligned} \quad (4.95)$$

Очевидно, что если мы аналогично определим величину

$$\begin{aligned} \bar{J}_3^y = & -\lambda \left\{ \omega \int_{x_i}^{x_{i+1/2}} \left[u^y(x, y_{j+1/2}) - u^y(x, y_j) \right] dx \right. \\ & \left. + (1 - \omega) \int_{x_i}^{x_{i+1}} \left[u^y(x, y_{j+1}) - u^y(x, y_j) \right] dx \right\}, \end{aligned} \quad (4.96)$$

то сумма $\bar{J}_3 = \bar{J}_3^x + \bar{J}_3^y$ будет представлять линейную комбинацию полных потоков через границы подобъемов V_3 и $\bar{V}_3$. Отметим также, что в силу условий сопряжения (4.89) суммы потоков J_k и $\bar{J}_k$ равны между собой, т. е. в дополнение к (4.91) справедливо равенство

$$J = \bar{J}_1 + \bar{J}_2 + \bar{J}_3 + \bar{J}_4. \quad (4.97)$$

Легко заметить теперь, что если мы будем аппроксимировать $\bar{J}_3^x$, $\bar{J}_3^y$ аналогично (2.84), то члены третьего порядка малости вида (2.86) взаимно уничтожатся и полученные выражения $\bar{J}_3^{xh}$, $\bar{J}_3^{yh}$ будут иметь погрешности $O(h^4)$.

Далее легко установить, что если мы определим величины

$$\begin{aligned} \tilde{J}_3^{xh} = & -\lambda_3 \frac{h_j^y}{8} \left[\omega (3u_{i+1/2,j}^x + u_{i+1/2,j+1}^x) \right. \\ & \left. + 4(1 - \omega)(u_{i+1,j}^x + u_{i+1,j+1}^x + u_{i,j}^x + u_{i,j+1}^x) \right], \end{aligned} \quad (4.98)$$

$$\begin{aligned} \tilde{J}_3^{yh} = & -\lambda_3 \frac{h_i^x}{8} \left[\omega (3u_{i,j+1/2}^x + u_{i+1,j+1/2}^y) \right. \\ & \left. + 4(1-\omega)(u_{i+1,j}^y + u_{i+1,j+1}^y + u_{i,j}^y + u^y) \right], \end{aligned} \quad (4.99)$$

получаемые формально добавлением к $\bar{J}_3^{xh}$ и $\bar{J}_3^{yh}$ удвоенных членов квадратурных формул для интегралов $\int_{y_j}^{y_{j+1}} u^x(x_i, y) dy$ и $\int_{x_i}^{x_{i+1}} u^y(x, y_j) dx$ соответственно, то для их сумм $\tilde{J}_3^h = \tilde{J}_3^{xh} + \tilde{J}_3^{yh}$ справедливо, опять же вследствие условий сопряжения, соотношение

$$J = \tilde{J}_1^h + \tilde{J}_2^h + \tilde{J}_3^h + \tilde{J}_4^h + O(h^4). \quad (4.100)$$

В выражениях (4.98), (4.99) теперь надо аппроксимировать производные. Используя для этого формулы центральных прямоугольников и трапеций, получаем

$$\begin{aligned} J_3^h = \tilde{J}_3^{xh} + \tilde{J}_3^{yh} + \psi_3 = & \lambda_3 \left\{ \frac{\omega}{8} \left[\frac{h_j^y}{h_i^x} (3(u_{ij} - u_{i+1,j}) + u_{i,j+1} - u_{i+1,j+1}) \right. \right. \\ & + \frac{h_i^x}{h_j^y} (3(u_{ij} - u_{i,j+1}) + u_{i+1,j} - u_{i+1,j+1}) \Big] \\ & + (1-\omega) \left[\frac{h_j^y}{h_i^x} (u_{ij} - u_{i+1,j} + u_{i,j+1} - u_{i+1,j+1}) \right. \\ & \left. \left. + \frac{h_i^x}{h_j^y} (u_{ij} - u_{i,j+1} + u_{i+1,j} - u_{i+1,j+1}) \right] \right\} \\ = & \lambda_3 \left\{ \frac{h_j^y}{h_i^x} \left(\left(1 - \frac{5\omega}{8}\right) (u_{ij} - u_{i+1,j}) + \left(1 - \frac{7\omega}{8}\right) (u_{i,j+1} - u_{i+1,j+1}) \right) \right. \\ & \left. + \frac{h_i^x}{h_j^y} \left(\left(1 - \frac{5\omega}{8}\right) (u_{ij} - u_{i,j+1}) + \left(1 - \frac{7\omega}{8}\right) (u_{i+1,j} - u_{i+1,j+1}) \right) \right\} \end{aligned} \quad (4.101)$$

где погрешность аппроксимации

$$\psi_3 = \frac{\lambda_3}{6} \left(\frac{9}{8}\omega - 1 \right) h_j^y h_i^x (u_{i,j}^{xxx} h_i^x + u_{i,j}^{yyy} h_j^y) + O(h^4). \quad (4.102)$$

Как видно, ошибка аппроксимации будет величиной $O(h^4)$ только при единственном значении весового параметра $\omega = \frac{8}{9}$, которое также обеспечивает четвертый порядок аппроксимации полного потока

$$J = J_1^h + J_2^h + J_3^h + J_4^h + O(h^4).$$

Рассмотрим теперь аппроксимацию правой части балансного соотношения (4.90), которую также представим в виде суммы слагаемых по подобьям $\bar{V}_k$:

$$g^h = f^h - (\mu u)^h = \sum_{k=1}^4 [f_k^h - (\mu u)_k^h].$$

Остановимся подробнее на представлении выражения для g_3^h , которое получаем путем применения формул центральных прямоугольников для интегралов по ячейкам V_3 и $\bar{V}_3$, а также последующей линейной интерполяции значений функций в их центрах по соответствующим узловым величинам:

$$g_3^h = f_3^h - (\mu u)_3^h = \frac{h_i^x h_j^y}{8} \left\{ \omega (f - \mu u)_{i,j} + \frac{8-7\omega}{2} \left[(f - \mu u)_{i+1,j} + (f - \mu u)_{i,j+1} \right] \right\} + \psi_3^f, \quad (4.103)$$

где $\psi_3^g = O(h^4)$ для любого параметра ω . Отсюда для полной погрешности аппроксимации правой части имеем четвертый порядок.

4.5.3. Основные свойства алгебраической системы

Суммируя, наконец, для всех подобьев выражения вида (4.88), (4.99), (4.103), получаем итоговую аппроксимацию для интегрального соотношения (4.90), которую запишем с использованием локальной нумерации узлов на рис. 4.6:

$$(Au)_{i,j} \equiv \left(a_0 u_0 - \sum_{k=1}^8 a_k u_k \right)_{i,j} = f_{i,j}^h + O(h^4), \quad (4.104)$$

где $a_k = a_k^J + a_k^H$, а отдельные коэффициенты описываются следующими формулами:

$$a_0^J = \sum_{k=1}^8 a_k^J = \left(1 - \frac{5}{8} \omega \right) \left(\lambda_1 B_{j-1,i-1}^{y,x} + \lambda_2 B_{i,j-1}^{x,y} + \lambda_3 B_{j,i}^{y,x} + \lambda_4 B_{i-1,j}^{x,y} \right),$$

$$\begin{aligned}
a_1^J &= \lambda_1 H_{j-1,i-1}^{y,x} + \lambda_4 H_{j,i-1}^{y,x}, & a_2^J &= \lambda_1 H_{i-1,j-1}^{x,y} + \lambda_2 H_{i,j-1}^{x,y}, \\
a_3^J &= \lambda_2 H_{j-1,i}^{y,x} + \lambda_3 H_{j,i}^{y,x}, & a_4^J &= \lambda_3 H_{i,j}^{x,y} + \lambda_4 H_{i-1,j}^{x,y}, \\
a_5^J &= \lambda_4 B_{i-1,j}^{x,y} \left(1 - \frac{7}{8}\omega\right), & a_6^J &= \lambda_1 B_{j-1,i-1}^{y,x} \left(1 - \frac{7}{8}\omega\right), \\
a_7^J &= \lambda_2 B_{i,j-1}^{x,y} \left(1 - \frac{7}{8}\omega\right), & a_8^J &= \lambda_3 B_{j,i}^{y,x} \left(1 - \frac{7}{8}\omega\right),
\end{aligned} \tag{4.105}$$

$$\begin{aligned}
a_0^H &= \frac{\omega}{8} \mu_{i,j} (h_j + h_{j-1}) (h_i + h_{i-1}), \\
a_1^H &= -h_{i-1} (h_j + h_{j-1}) \frac{8-7\omega}{8} \mu_{i-1,j}, \\
a_3^H &= -h_i (h_j + h_{j-1}) \frac{8-7\omega}{8} \mu_{i+1,j}, \\
a_2^H &= -h_{j-1} (h_j + h_{j-1}) \frac{8-7\omega}{8} \mu_{i,j-1}, \\
a_4^H &= -h_j (h_i + h_{i-1}) \frac{8-7\omega}{8} \mu_{i,j+1}.
\end{aligned} \tag{4.106}$$

Правую часть уравнения (4.104) можно записать в виде

$$f_{i,j}^h = a_0^H f_0 + \sum_{k=1}^4 a_k^H f_k,$$

если при этом в формуле (4.103) для a_k^H положить предварительно $\mu_{i,j} \equiv 1$.

Отметим, что построенная таким образом система уравнений (4.104) на равномерной сетке совпадает после деления на величину $(4 - 3\omega)h^2$ при $\lambda = 1$ со схемой Микеладзе (3.42), а также с эквивалентной ей интегро-балансной аппроксимацией (4.57), которые обе имеют погрешность аппроксимации $O(h^4)$.

Здесь необходимо подчеркнуть, во избежание путаницы, что система уравнений (4.102) отнормирована так, что ее коэффициенты имеют порядок $O(1)$, в то время как у непосредственных аппроксимаций (3.42), (4.57) для уравнения Пуассона они составляют величины $O(h^{-2})$. Поэтому у системы (4.104), после перенормировки ее к "стандартному" разностному виду путем деления на коэффициент порядка $O(h^2)$, погрешность аппроксимации становится порядка $O(h^2)$.

Кроме того, легко заметить, что при весовом параметре $\omega = \frac{8}{7}$ и $\lambda = 1$, $\mu = 0$ схема (4.104) совпадает, с точностью до множителя, с обычной пятиточечной разностной схемой (3.38) для

уравнения Пуассона, так как в этом случае $a_5 = a_6 = a_7 = a_8 = 0$.

Обратимся теперь к вопросу об учете краевых условий. Если (i, j) -й узел является граничным с условием Дирихле, то для него сеточное уравнение не строится, а известное значение $u_{i,j} = g_{i,j}$ подставляется в уравнения для соседних внутренних узлов и переносится в соответствующие правые части.

Рассмотрим учет граничного условия 3-го рода на примере, когда оно задано на сеточном ребре $(0-4)$, т. е. подобьем $V_4 = \{x_{i-1} < x < x_i, y_j < y < y_{j+1}\}$ является внешним.

При выводе уравнений (4.104) мы использовали условия сопряжения (4.89) для аннулирования "избыточных" членов с производными на смежных сторонах соседних ячеек. В рассматриваемом случае эти члены не аннулируются и имеют вид

$$J_{04} = \lambda_3 \frac{h_j}{8} \left[\omega(3u_{i,j}^x + u_{i,j+1}^x) + 8(1 - \omega)u_{i,j}^x + u_{i,j+1}^x \right].$$

Это выражение в силу условия (4.88) можно привести к форме

$$J_{04} = \lambda_3 \frac{h_j}{8} \left\{ \omega[3(\varkappa u - \gamma)_0 + (\varkappa u - \gamma)_4] + 8(1 - \omega)[(\varkappa u - \gamma)_0 + (\varkappa u - \gamma)_4] \right\}, \quad (4.107)$$

причем без внесения какой-либо дополнительной погрешности. Теперь для учета краевого условия в левую часть уравнения (4.104) надо добавить члены с u_k из (4.107), а в правую – члены с γ_k . Естественно при этом, что в формуле (4.100) член $\tilde{J}_4^h$ должен отсутствовать (формально для этого в выражениях (4.105) можно положить $\lambda_4 = 0$). Получаемую итоговую матрицу системы (4.104) можно представить в форме

$$A = A^J + A^H + A^N, \quad (4.108)$$

где матрицы $A^J + A^H$ состояются из коэффициентов a_k^J и a_k^H из (4.105), (4.106), отвечающих соответственно за аппроксимацию потоков и "гельмгольцовского" члена μu , а матрица A^N обусловлена модификациями при учете граничных условий 3-го рода. Отметим, что в методах конечных элементов аналоги матриц A^J и A^H называются глобальными матрицами жесткости и масс. Мы же будем называть соответственно матрицами потоков и источников, а суммарную матрицу A – глобальной матрицей баланса.

4.5.4. Локальные матрицы и технология сборки алгебраической системы

Остановимся теперь на поэлементной технологии формирования рассмотренных сеточных уравнений.

Введем в подобье $\bar{V}_3$ локальные векторы решения, потока и правой части

$$u^{(3)} = \{u_k\}, \quad J^{(3)} = \{J_k^{(3)}\}, \quad f^{(3)} = \{f_k^{(3)}\}, \quad k = \{0, 3, 4, 8\}.$$

Здесь верхний индекс обозначает номер подобья, компонента $J_k^{(3)}$ есть поток вокруг k -го узла подобья, причем компонента потока $J_0^{(3)}$ определяется при помощи (4.101).

Каждая компонента $J_k^{(3)}$ есть линейная форма неизвестных значений u_k . Например, компонента $J_3^{(3)}$ после аппроксимации представима в следующем виде: $J_3^{(3)} = a_{30}u_0 + a_{33}u_3 + a_{34}u_4 + a_{38}u_8$. Тогда поток в сеточной ячейке можно записать в векторной форме

$$J^{(3)} = A_3^J u^{(3)}, \quad (4.109)$$

где $A_3^J = \{a_{ij}\}$ есть симметричная квадратная матрица четвертого порядка:

$$A_3^J = \begin{pmatrix} a_{00} & a_{03} & a_{04} & a_{08} \\ a_{30} & a_{33} & a_{34} & a_{38} \\ a_{40} & a_{43} & a_{44} & a_{48} \\ a_{80} & a_{83} & a_{84} & a_{88} \end{pmatrix} \quad (4.110)$$

с элементами, определяемыми из (4.101) и представляемыми следующими формулами (с учетом симметричности матрицы):

$$\begin{aligned} a_{00} &= \lambda_3 \left(1 - \frac{5\omega}{8}\right) B_{j,i}^{y,x}, & a_{08} &= \lambda_3 \left(1 - \frac{7\omega}{8}\right) B_{j,i}^{y,x}, \\ a_{03} &= \lambda_3 H_{j,i}^{y,x}, & a_{04} &= \lambda_3 H_{i,j}^{x,y}, \\ a_{33} &= a_{44} = a_{88} = a_{00}, & a_{34} &= a_{08}, \quad a_{38} = a_{04}, \quad a_{48} = a_{03}, \end{aligned} \quad (4.111)$$

где используются обозначения как в (4.105).

Выражения для компонент локального вектора правой части выписываются из представления (2.95) следующим образом:

$$\begin{aligned}
f_0^3 &= \frac{\omega}{8} f_{ij} + \frac{8-7\omega}{16} (f_{i+1,j} + f_{i,j+1}), \\
f_3^3 &= \frac{\omega}{8} f_{i+1,j} + \frac{8-7\omega}{16} (f_{ij} + f_{i+1,j+1}), \\
f_4^3 &= \frac{\omega}{8} f_{i,j+1} + \frac{8-7\omega}{16} (f_{ij} + f_{i+1,j+1}), \\
f_8^3 &= \frac{\omega}{8} f_{i+1,j+1} + \frac{8-7\omega}{16} (f_{i+1,j} + f_{i,j+1}).
\end{aligned} \tag{4.112}$$

Аналогично можно ввести для каждого подобъема $\bar{V}_l$ соответствующие векторы $u^{(l)}$, $J^{(l)}$, $f^{(l)}$ и выписать представление $J^{(l)} = A_l^J u^{(l)}$. Это представление получено после аппроксимации балансного соотношения в одной ячейке сетки, т. е. “локально”, поэтому назовем матрицу A_l^J локальной матрицей потоков.

Очевидно, что аналогично рассмотрению (4.77) для треугольных конечных объемов мы можем связать каждую локальную матрицу потоков A_l^J с соответствующей “глобальной” матрицей A^J порядка N путем введения соответствий

$$u = P_k u^{(k)}, \quad \bar{A}_l^J = P_k' A_l^J P_k, \quad A^J = \sum_l \bar{A}_l^J, \tag{4.113}$$

где P_k – прямоугольная матрица размера $4 \times N$ с одной ненулевой (единичной) компонентой в каждом столбце, $\bar{A}_l^J$ – “растянутая” локальная матрица потоков порядка $N \times N$, а суммирование проводится по всем ячейкам сетки $\bar{V}_l$.

Очень важно, что элементы *локальных матриц потоков* A_l^J могут вычисляться в каждом подобъеме независимо друг от друга (а при желании – и параллельно), после чего сборка глобальной матрицы A^J проводится путем простого последовательного накопительного суммирования с помощью формул

$$\begin{aligned}
(a_0^J)_{i,j} &= (a_{0,0})_{i,j} + (a_{3,3})_{i-1,j} + (a_{4,4})_{i,j-1} + (a_{8,8})_{i-1,j-1}, \\
(a_3^J)_{i,j} &= (a_{0,3})_{i,j} + (a_{4,8})_{i,j-1}, \quad (a_4^J)_{i,j} = (a_{0,4})_{i,j} + (a_{3,8})_{i-1,j}, \\
(a_7^J)_{i,j} &= (a_{0,8})_{i,j-1}, \quad (a_8^J)_{i,j} = (a_{0,8})_{i,j}.
\end{aligned} \tag{4.114}$$

Отметим, что в методах конечных элементов аналог матрицы A_l^J называется локальной матрицей жесткости, а процесс формирования “большой” матрицы A^J по формулам вида (2.99), (2.100) – сборкой, или ассемблированием, глобальной матрицы.

По аналогии с A_l^J можно ввести локальную матрицу A_l^H четвертого порядка, элементы которой в аналогичных с (2.99) обозначениях на основании выражения (2.93) могут быть записаны следующими формулами:

$$a_{k,k}^H = \frac{h_i^x h_j^y}{8} \omega \mu_k, \quad a_{k,k'}^H = \frac{h_i^x h_j^y}{16} (8 - 7\omega) \mu_{k'},$$

$$k, k' = 0, 3, 4, 8, \quad k \neq k', \quad a_{0,8}^H = a_{8,0}^H = a_{3,4}^H = a_{4,3}^H = 0.$$

Сборка элементов a_k^H глобальной матрицы A^H из $a_{k,k'}^H$ проводится совершенно аналогично формулам (4.115).

Поскольку в задачах математической физики член дифференциального уравнения μ_i имеет смысл источника субстанции (тепла или массы – в задачах теплопроводности или диффузии), то матрицы A_l^H и A^H мы будем называть *локальной* и *глобальной матрицами источника*. Напомним, что их прототипами в методах конечных элементов являются матрицы масс.

Наконец, в околограничных объемах $\bar{V}_l$, у которых хотя бы на одном сеточном ребре задано краевое условие 3-го рода, можно определить локальную матрицу A_l^N , ненулевые элементы которой (все пропорциональные значениям ω в каких-то граничных точках) можно определить с помощью формулы (4.107), из совокупности которых аналогично (4.114) формируется глобальная матрица A^N из (4.108).

Суммарную матрицу $A_l = A_l^J + A_l^H + A_l^N$ будем называть локальной матрицей баланса.

Отметим еще, что аналогично поэлементной сборке матрицы системы легко вычисляются и компоненты правых частей уравнений, причем здесь удобно используется матрица источников. А именно, если ввести локальные векторы $f_h^{(3)} = \{f_h^0, f_h^3, f_h^4, f_h^8\}$ (здесь верхний индекс, в отличие от k в (4.103), означает локальный номер узла, т. е. f_h^0 равно f_3^h из (4.103) и $f^{(3)}$ равно f_0, f_3, f_4, f_8), то между ними можно установить связь $f_h^{(3)} = A_l^H f^{(3)}$.

Рассмотренный поэлементный подход совершенно аналогично переносится на трехмерные смешанные краевые задачи. При этом для значения весового параметра $\omega = \frac{16}{17}$ получается схема с погрешностью порядка $O(h^2)$ (при “стандартной” нормировке уравнений с порядком разностных коэффициентов $O(h^{-2})$), при $\omega = \frac{32}{31}$ следует схема $O(h^4)$ вида (1.54), а при остальных значениях ω локальная погрешность есть величина $O(h)$, причем для $\omega = \frac{16}{15}$ имеем обычную семиточечную схему вида (3.51). Естественно, в силу увеличения размерности технические выкладки существенно усложняются по сравнению с двумерным случаем, поэтому в трехмерных задачах особенно эффективно исполь-

зование сборки глобальной матрицы баланса путем вычисления локальных матриц потоков, источников и краевых условий 3-го рода.

4.5.5. Аппроксимация уравнения диффузии с переменными коэффициентами

Рассмотренные выше методы конечных объемов существенно использовали постоянство коэффициента диффузии в пределах одной ячейки (элемента) сетки. В принципе применяемые подходы приближения интегралов могут быть использованы и для аппроксимации дифференциальных уравнений с переменными коэффициентами, в том числе диффузионно-конвективного типа.

В данном случае возникают следующие ключевые вопросы: можно ли построить на поэлементном подходе метода конечных объемов схему четвертого порядка на равномерной сетке и удастся ли при этом для исходной самосопряженной задачи получить симметричную алгебраическую систему сеточных уравнений? Пока эти вопросы требуют дополнительных исследований.

Приведем параметрическое семейство девятиточечных уравнений, которое формально является обобщением схемы (4.104) на случай переменных коэффициентов диффузии в уравнении

$$L_x u + L_y u \equiv -\frac{\partial}{\partial x} \left(p \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial y} \left(q \frac{\partial u}{\partial y} \right) = f(x, y). \quad (4.115)$$

Анализ аппроксимационных свойств получаемой системы мы не проводим в силу его громоздкости при неравномерности шагов сетки h_i^x, h_j^y . В случае же постоянства шагов h^x, h^y предлагаемая система совпадает с компактной схемой (3.78) и ее модификациями, связанными с заменами разностных операторов $L_{1,x}^h, L_{1,y}^h$ на $L_{k,x}^h, L_{k,y}^h$, которые имеют погрешность аппроксимации $O(h^4)$.

Пусть одномерные разностные операторы на равномерной сетке определены в соответствии с формулами (3.76), (3.85) (индекс k для краткости будем опускать) и удовлетворяют аппроксимационному свойству вида (3.72):

$$\begin{aligned} (L_x^h u)_{i,j} &= \\ &= \frac{1}{h_x^2} \left[\hat{p}_{i+1/2,j} (u_{i,j} - u_{i+1,j}) + \hat{p}_{i-1/2,j} (u_{i,j} - u_{i-1,j}) \right], \end{aligned}$$

$$(L_y^h u)_{i,j} = \frac{1}{h_y^2} \left[\hat{q}_{i,j+1/2} (u_{i,j} - u_{i,j+1}) + \hat{q}_{i,j-1/2} (u_{i,j} - u_{i,j-1}) \right]. \quad (4.116)$$

Здесь коэффициенты $\hat{p}_{i\pm 1/2}$, $\hat{q}_{i,j\pm 1/2}$ определяются по формулам вида (3.77) при $a = b = 0$ или же совпадают с $\hat{p}_{i\pm 1/2,j}^{(k)}$, $\hat{q}_{i,j\pm 1/2}^{(k)}$ из (3.84), (3.88), т. е. могут иметь разные модификации. Тогда девятиточечная аппроксимация диффузионного уравнения (4.115) записывается в форме

$$\begin{aligned} L^h u &\equiv L_x^h u + L_y^h u - \frac{h_x^2}{12} L_x^h \frac{1}{p} L_y^h u - \frac{h_y^2}{12} L_y^h \frac{1}{q} L_x^h u \\ &= f - \frac{h_x^2}{12} L_x^h \frac{f}{p} - \frac{h_y^2}{12} L_y^h \frac{f}{q} = \bar{f}^h, \end{aligned} \quad (4.117)$$

где последний член левой части, например, имеет вид

$$\begin{aligned} \frac{h_y^2}{12} \left(L_y^h \frac{1}{q} L_x^h u \right)_{i,j} &= \frac{1}{12h_x^2} \left\{ \frac{\hat{q}_{i,j+1/2} + \hat{q}_{i,j-1/2}}{q_{i,j}} \right. \\ &\times \left[\hat{p}_{i+1/2,j} (u_{i,j} - u_{i+1,j}) + \hat{p}_{i-1/2,j} (u_{i,j} - u_{i-1,j}) \right] \\ &- \frac{\hat{q}_{i,j+1/2}}{q_{i,j+1}} \left[\hat{p}_{i+1/2,j} (u_{i,j+1} - u_{i+1,j+1}) + \hat{p}_{i-1/2,j+1} (u_{i,j+1} - u_{i-1,j+1}) \right] \\ &\left. - \frac{\hat{q}_{i,j-1/2}}{q_{i,j-1}} \left[\hat{p}_{i+1/2,j} (u_{i,j-1} - u_{i+1,j-1}) + \hat{p}_{i-1/2,j-1} (u_{i,j-1} - u_{i-1,j-1}) \right] \right\}, \end{aligned} \quad (4.118)$$

а предпоследний определяется аналогично. Отметим, что в формулах (4.116)–(4.117) величины с целыми индексами вида $p_{i,j}$, $q_{i,j}$ означают $p(x_i, y_j)$, $q(x_i, y_j)$.

В соответствии с локальной нумерацией узлов на рис. 4.5 разностную схему (4.118) можно записать в покомпонентном виде

$$(L^h u)_{i,j} \equiv a_{0,i,j} u_0 - \sum_{k=1}^8 a_{k,i,j} u_k = \bar{f}_{i,j}^h, \quad (4.119)$$

где разностные коэффициенты имеют вид

$$\begin{aligned} a_{1,i,j} &= \frac{\hat{p}_{i-1/2,j}}{h_x^2} \left(1 - \frac{\hat{q}_{i,j+1/2} + \hat{q}_{i,j-1/2}}{12q_{i,j}} - \frac{\hat{q}_{i-1,j+1/2} + \hat{q}_{i-1,j-1/2}}{12p_{i-1,j}} \frac{h_x^2}{h_y^2} \right), \\ a_{2,i,j} &= \frac{\hat{q}_{i,j-1/2}}{h_y^2} \left(1 - \frac{\hat{p}_{i+1/2,j} + \hat{p}_{i-1/2,j}}{12p_{i,j}} - \frac{\hat{p}_{i+1/2,j-1} + \hat{p}_{i-1/2,j-1}}{12q_{i,j-1}} \frac{h_y^2}{h_x^2} \right), \end{aligned}$$

$$\begin{aligned}
a_{3,i,j} &= \frac{\hat{p}_{i+1/2,j}}{h_x^2} \left(1 - \frac{\hat{q}_{i,j+1/2} + \hat{q}_{i,j-1/2}}{12q_{i,j}} - \frac{\hat{q}_{i+1,j+1/2} + \hat{q}_{i+1,j-1/2}}{12p_{i+1,j}} \frac{h_x^2}{h_y^2} \right), \\
a_{4,i,j} &= \frac{\hat{q}_{i,j+1/2}}{h_y^2} \left(1 - \frac{\hat{p}_{i+1/2,j} + \hat{p}_{i-1/2,j}}{12p_{i,j}} - \frac{\hat{p}_{i+1/2,j+1} + \hat{p}_{i-1/2,j+1}}{12q_{i,j+1}} \frac{h_y^2}{h_x^2} \right), \\
a_{5,i,j} &= \frac{1}{12} \left(\frac{\hat{p}_{i-1/2,j} \hat{q}_{i-1,j+1/2}}{p_{i-1,j} h_y^2} + \frac{\hat{q}_{i,j+1/2} \hat{p}_{i-1/2,j+1}}{q_{i,j+1} h_x^2} \right), \\
a_{6,i,j} &= \frac{1}{12} \left(\frac{\hat{p}_{i-1/2,j} \hat{q}_{i-1,j-1/2}}{p_{i-1,j} h_y^2} + \frac{\hat{q}_{i,j-1/2} \hat{p}_{i-1/2,j-1}}{q_{i,j-1} h_x^2} \right), \\
a_{7,i,j} &= \frac{1}{12} \left(\frac{\hat{p}_{i+1/2,j} \hat{q}_{i+1,j-1/2}}{p_{i+1,j} h_y^2} + \frac{\hat{q}_{i,j-1/2} \hat{p}_{i+1/2,j-1}}{q_{i,j-1} h_x^2} \right), \\
a_{8,i,j} &= \frac{1}{12} \left(\frac{\hat{p}_{i+1/2,j} \hat{q}_{i+1,j+1/2}}{p_{i+1,j} h_y^2} + \frac{\hat{p}_{i+1/2,j+1} \hat{q}_{i,j+1/2}}{q_{i,j+1} h_x^2} \right), \\
a_{0,i,j} &= a_{1,i,j} + \dots + a_{8,i,j}.
\end{aligned} \tag{4.120}$$

Возвращаясь теперь к методу конечных объемов, проведем аппроксимацию интегрального соотношения баланса (4.90), в котором полагаем $\mu = 0$, а нормальные компоненты потока на границе $-\lambda \frac{\partial u}{\partial n}$ заменяем на величины $\pm p \frac{\partial u}{\partial x}$, $\pm q \frac{\partial u}{\partial y}$ при интегрировании соответственно по вертикальным или горизонтальным сторонам ячеек. Ввиду громоздкости получаемого выражения в целом выпишем только “парциальную” составляющую потока $J_3^h = J_{0,i,j}^{(3)}$ по подобъемам V_3 и $\bar{V}_3$ в окрестности 0-й точки, являющуюся формальным обобщением формулы (4.101):

$$\begin{aligned}
J_{0,i,j}^{(3)} &= \frac{\omega}{8} \left\{ \frac{h_j^y}{h_i^x} \left[3\hat{p}_{i+1/2,j}(u_{i,j} - u_{i+1,j}) + \hat{p}_{i+1/2,j+1} \frac{\hat{q}_{i,j+1/2}}{q_{i,j+1}} (u_{i,j+1} - u_{i+1,j+1}) \right] \right. \\
&\quad \left. + \frac{h_i^x}{h_j^y} \left[3\hat{q}_{i,j+1/2}(u_{i,j} - u_{i,j+1}) + \hat{q}_{i+1,j+1/2} \frac{\hat{p}_{i+1/2,j}}{p_{i+1,j}} (u_{i+1,j} - u_{i+1,j+1}) \right] \right\} \\
&\quad + (1-\omega) \left\{ \frac{h_j^y}{h_i^x} \hat{q}_{i,j+1/2} \left[\frac{\hat{p}_{i+1/2,j}}{q_{i,j}} (u_{i,j} - u_{i+1,j}) + \frac{\hat{p}_{i+1/2,j+1}}{q_{i,j+1}} (u_{i,j+1} - u_{i+1,j+1}) \right] \right. \\
&\quad \left. + \frac{h_i^x}{h_j^y} \hat{p}_{i+1/2,j} \left[\frac{\hat{q}_{i,j+1/2}}{p_{i,j}} (u_{i,j} - u_{i,j+1}) + \frac{\hat{q}_{i+1,j+1/2}}{p_{i+1,j}} (u_{i+1,j} - u_{i+1,j+1}) \right] \right\}.
\end{aligned}$$

Если таким же образом выписать выражения для $J_{k,i,j}^{(3)}$, $k = 3, 4, 8$, то по аналогии с (??), (4.111) можно определить элементы $a_{k,l}$ локальной матрицы баланса $A_{3,i,j}^J$, которая в данном случае будет несимметричной:

$$\begin{aligned}
a_{0,3}^{i,j} &= \hat{p}_{i+1/2,j} \left[\frac{\omega}{8} \left(3 \frac{h_j^y}{h_i^x} - \frac{h_i^x}{h_j^y} \frac{\hat{q}_{i+1,j+1/2}}{p_{i+1,j}} \frac{h_i^x}{h_j^y} \right) \right. \\
&\quad \left. + (1 - \omega) \left(\frac{h_j^y}{h_i^x} \frac{\hat{q}_{i,j+1/2}}{q_{i,j}} - \frac{h_i^x}{h_j^y} \frac{\hat{q}_{i+1,j+1/2}}{p_{i+1,j}} \right) \right], \\
a_{0,4}^{i,j} &= \hat{q}_{i,j+1/2} \left[\frac{\omega}{8} \left(3 \frac{h_i^x}{h_j^y} - \frac{h_j^y}{h_i^x} \frac{\hat{p}_{i+1/2,j+1}}{q_{i,j+1}} \right) + \right. \\
&\quad \left. + (1 - \omega) \left(\frac{h_i^x}{h_j^y} \frac{\hat{p}_{i+1/2,j}}{p_{i,j}} - \frac{h_j^y}{h_i^x} \frac{\hat{p}_{i+1/2,j+1}}{q_{i,j+1}} \right) \right], \\
a_{0,8}^{i,j} &= \left(\frac{h_j^y}{h_i^x} \hat{p}_{i+1/2,j+1} \frac{\hat{q}_{i,j+1/2}}{q_{i,j+1}} + \frac{h_i^x}{h_j^y} \hat{q}_{i+1,j+1/2} \frac{\hat{p}_{i+1/2,j}}{p_{i+1,j}} \right) \left(1 - \frac{7}{8} \omega \right), \\
a_{3,0}^{i,j} &= \hat{p}_{i+1/2,j} \left[\frac{\omega}{8} \left(3 \frac{h_j^y}{h_i^x} - \frac{h_i^x}{h_j^y} \frac{\hat{q}_{i,j+1/2}}{p_{i,j}} \right) \right. \\
&\quad \left. + (1 - \omega) \left(\frac{h_j^y}{h_i^x} \frac{\hat{q}_{i+1,j+1/2}}{q_{i+1,j}} - \frac{h_i^x}{h_j^y} \frac{\hat{q}_{i,j+1/2}}{p_{i,j}} \right) \right], \\
a_{4,0}^{i,j} &= \hat{q}_{i,j+1/2} \left[\frac{\omega}{8} \left(3 \frac{h_i^x}{h_j^y} - \frac{h_j^y}{h_i^x} \frac{\hat{p}_{i+1/2,j}}{q_{i,j}} \right) \right. \\
&\quad \left. + (1 - \omega) \left(\frac{h_i^x}{h_j^y} \frac{\hat{p}_{i+1/2,j+1}}{p_{i,j+1}} - \frac{h_j^y}{h_i^x} \frac{\hat{p}_{i+1/2,j}}{q_{i,j}} \right) \right], \\
a_{8,0}^{i,j} &= \left(\frac{h_j^y}{h_i^x} \hat{p}_{i+1/2,j} \frac{\hat{q}_{i+1,j+1/2}}{q_{i+1,j}} + \frac{h_i^x}{h_j^y} \hat{q}_{i,j+1/2} \frac{\hat{p}_{i+1/2,j+1}}{p_{i,j+1}} \right) \left(1 - \frac{7}{8} \omega \right), \\
a_{3,4}^{i,j} &= \left(\frac{h_i^x}{h_j^y} \hat{p}_{i+1/2,j} \frac{\hat{q}_{i+1,j+1/2}}{p_{i,j}} + \frac{h_j^y}{h_i^x} \hat{q}_{i+1,j+1/2} \frac{\hat{p}_{i+1/2,j+1}}{q_{i+1,j+1}} \right) \left(1 - \frac{7}{8} \omega \right), \\
a_{4,3}^{i,j} &= \left(\frac{h_i^x}{h_j^y} \hat{p}_{i+1/2,j+1} \frac{\hat{q}_{i+1,j+1/2}}{p_{i+1,j+1}} + \frac{h_j^y}{h_i^x} \hat{q}_{i,j+1/2} \frac{\hat{p}_{i+1/2,j}}{q_{i,j}} \right) \left(1 - \frac{7}{8} \omega \right), \\
a_{3,8}^{i,j} &= \hat{q}_{i+1,j+1/2} \left[\frac{\omega}{8} \left(3 \frac{h_i^x}{h_j^y} - \frac{h_j^y}{h_i^x} \frac{\hat{p}_{i+1/2,j+1}}{q_{i+1,j+1}} \right) \right. \\
&\quad \left. + (1 - \omega) \left(\frac{h_i^x}{h_j^y} \frac{\hat{p}_{i+1/2,j}}{p_{i+1,j}} - \frac{h_j^y}{h_i^x} \frac{\hat{p}_{i+1/2,j+1}}{q_{i+1,j+1}} \right) \right], \\
a_{8,3}^{i,j} &= \hat{q}_{i+1,j+1/2} \left[\frac{\omega}{8} \left(3 \frac{h_i^x}{h_j^y} - \frac{h_j^y}{h_i^x} \frac{\hat{p}_{i+1/2,j}}{q_{i+1,j}} \right) \right. \\
&\quad \left. + (1 - \omega) \left(\frac{h_j^y}{h_i^x} \frac{\hat{p}_{i+1/2,j}}{q_{i+1,j}} - \frac{h_i^x}{h_j^y} \frac{\hat{p}_{i+1/2,j+1}}{p_{i+1,j+1}} \right) \right], \\
a_{4,8}^{i,j} &= \hat{p}_{i+1/2,j+1} \left[\frac{\omega}{8} \left(3 \frac{h_j^y}{h_i^x} - \frac{h_i^x}{h_j^y} \frac{\hat{q}_{i+1,j+1/2}}{p_{i+1,j+1}} \right) \right.
\end{aligned}$$

$$\begin{aligned}
& + (1 - \omega) \left(\frac{h_j^y}{h_i^x} \frac{\hat{q}_{i,j+1/2}}{q_{i,j+1}} - \frac{h_i^x}{h_j^y} \frac{\hat{q}_{i+1,j+1/2}}{p_{i+1,j+1}} \right) \Bigg], \\
a_{8,4}^{i,j} &= \hat{p}_{i+1/2,j} \left[\frac{\omega}{8} \left(3 \frac{h_j^y}{h_i^x} - \frac{h_i^x}{h_j^y} \frac{\hat{q}_{i,j+1/2}}{p_{i,j}} \right) \right. \\
& \quad \left. + (1 - \omega) \left(\frac{h_j^y}{h_i^x} \frac{\hat{q}_{i+1,j+1/2}}{q_{i+1,j}} - \frac{h_i^x}{h_j^y} \frac{\hat{q}_{i+1,j+1/2}}{q_{i+1,j}} \frac{h_i^x}{h_j^y} \frac{\hat{q}_{i,j+1/2}}{p_{i,j}} \right) \right], \\
a_{0,0}^{i,j} &= a_{0,3}^{i,j} + a_{0,4}^{i,j} + a_{0,8}^{i,j}, \quad a_{3,3}^{i,j} = a_{3,0}^{i,j} + a_{3,4}^{i,j} + a_{3,8}^{i,j}, \\
a_{4,4}^{i,j} &= a_{4,0}^{i,j} + a_{4,3}^{i,j} + a_{4,8}^{i,j}, \quad a_{8,8}^{i,j} = a_{8,0}^{i,j} + a_{8,3}^{i,j} + a_{8,4}^{i,j}.
\end{aligned}$$

По аналогии с формулами (2.108) сборка элементов глобальной матрицы в данном несимметричном случае производится с помощью следующих соотношений:

$$\begin{aligned}
\bar{a}_{0,i,j} &= a_{0,0}^{i,j} + a_{3,3}^{i-1,j} + a_{4,4}^{i,j-1} + a_{8,8}^{i-1,j-1}, & \bar{a}_{1,i,j} &= a_{3,0}^{i-1,j} + a_{8,4}^{i-1,j-1}, \\
\bar{a}_{2,i,j} &= a_{4,0}^{i,j-1} + a_{8,3}^{i-1,j-1}, & \bar{a}_{3,i,j} &= a_{0,3}^{i,j} + a_{4,8}^{i,j-1}, \\
\bar{a}_{4,i,j} &= a_{0,4}^{i,j} + a_{3,8}^{i-1,j}, & \bar{a}_{5,i,j} &= a_{3,4}^{i-1,j}, \\
\bar{a}_{6,i,j} &= a_{8,0}^{i-1,j-1}, & \bar{a}_{7,i,j} &= a_{4,3}^{i,j-1}, & \bar{a}_{8,i,j} &= a_{0,8}^{i,j}.
\end{aligned} \tag{4.121}$$

Получаемая при этом матрица $\bar{L}_h = \{\bar{a}_{k,i,j}\}$ отличается масштабирующим множителем $(1 - \frac{3}{4}\omega)(h_i^x + h_{i-1}^x)(h_j^y + h_{j-1}^y)$ от матрицы системы разностных уравнений, получаемых из непосредственной аппроксимации дифференциального уравнения. В частности, коэффициенты $\bar{a}_{k,i,j}$ из (4.119) в случае равномерной сетки после деления на множитель $(4 - 3\omega)h_x^2 h_y^2$ совпадают с коэффициентом $a_{k,i,j}$ компактной разностной схемы (??), имеющей погрешность $O(h^4)$. Однако открытым остается вопрос, будет ли аппроксимация с коэффициентами (??) иметь преимущество перед "традиционной" разностной схемой на неравномерной сетке.

Глава 5

Конечно-элементные методы и технологии

Рождение методов конечных элементов связано с работой Р. Куранта 1943 года, который предложил для решения вариационных задач использовать кусочно-линейные аппроксимации. Позднее эта идея развилась в кусочно-полиномиальные приближения различных порядков, образующие конечно-мерные функциональные пространства с финитными базисными функциями, имеющими носители на конечных элементах, что обеспечивает им свойства “квазиортогональности”, и как следствие, – высокую экономичность получаемых алгоритмов. Интересно, что в теории аппроксимации были открыты такие же подходы, сформировавшие новый математический объект – сплайн, – ставший уникальным предметом активных теоретических исследований и многочисленных приложений, см. [18], [36], [51].

Методы конечных элементов в 50-е–60-е годы XX-го века были востребованы в первую очередь инженерами, разработавшими основы алгоритмов и их программных технологий, что нашло отражение в таких установившихся прикладных терминах, как “матрица жесткости”, “матрица масс”, “степени свободы” и т. д.

Примерно в тот же период активно велись исследования по теоретическому обоснованию МКЭ средствами функционального анализа в пространствах Соболева. Основные полученные здесь результаты нашли воплощение в книгах российских (точнее, ленинградских) математиков Л.А. Оганесяна, В.Я. Ривкинда, Л.А. Руховца и В.Г. Корнеева [138], [139], [92], а также французских и американских математиков Ж.-П. Обэна [137], Ф. Сьярле [173], О. Зенкевича [77], Г. Стрэнга и Дж. Фикса [171]. Дальнейшее развитие методов конечных элементов нашло отражение в монографиях О. Аксельсона и В. Баркера [2], Г.И. Марчука, В.И. Агошкова и В.В. Шайдурова [127], [199], И. Бабушки и Б. Сабо [154], Ю.М. Лаевского [100]–[102], В.И. Лебедева [104], С. Бреннер и Р. Скотта [26], Ф. Бреucci и М. Форгина [27], а также других авторов, см. например, [1], [83], [133], [136], [140], [166], [169], [186]. Среди активно исследуемых в последние десятилетия направлений следует назвать разрывные методы Галер-

кина и многомасштабные подходы, см. книги Б. Ривiere [149] и Э.П. Шуриной с коллегами [203], [204].

За период своего существования МКЭ завоевали ведущие мировые позиции в теории и практике решения задач математической физики. Их успех объясняется двумя факторами, каждый из которых трудно переоценить. Во-первых, на базе современного функционального аппарата создана единая теория, позволяющая разрабатывать, исследовать и обосновывать методы конечных элементов различных порядков точности для широкого класса дифференциальных уравнений и краевых условий, в том числе при наличии сингулярностей в решениях. Во-вторых, усилиями многочисленных авторов здесь создана единообразная математическая технология, обеспечивающая эффективную реализацию алгоритмов практически для любых сложных задач, включая многомерные области со сложными конфигурациями границ. И не случайно МКЭ занимают почти монопольные позиции на коммерческих рынках прикладного программного обеспечения по моделированию процессов и явлений в самых различных областях. В настоящее время в Интернете можно найти обширную информацию по таким продуктам: NASTRAN, ANSYS, CADFEM, STARCD, FEMLAB, FLUID, FENICS и многие другие. Кроме того, имеется огромное количество библиотечных программ и инструментариев, ориентированных на эффективное применение в различных приложениях и которые или могут свободно (бесплатно) использоваться для некоммерческих целей, или представляют собой дорогостоящую интеллектуальную собственность (MATLAB, NETLIB, BLAS, LAPACK, NAG, IMSL, MKL и т.д.). С точки зрения вычислительной математики, главная проблема пользователя таких разработок – обеспечить их квалифицированное применение для адекватности и эффективности численного эксперимента, но здесь уже возникает целый комплекс вопросов, относящихся к широкой области математического моделирования, посвященной использованию вычислительных методов и информационных технологий для решения больших прикладных задач.

Целью данной главы является ознакомление именно с технологическими, а не теоретическими вопросами методов конечных элементов. Под МКЭ понимается реализация метода Рунге или Галеркина приближенного решения краевых задач, заключающаяся в разбиении расчетной области на конечное число

элементов (ячеек) и аппроксимации вариационных уравнений на основе кусочно-полиномиальных интерполяций, представляемых с помощью финитных базисных функций, отличных от нуля только в нескольких соседних элементах, что позволяет получать алгебраические системы с разреженными матрицами. Разнообразие МКЭ определяется видом решаемых краевых задач, конфигурацией используемых конечных элементов и способами аппроксимации, т. е. выбором базисных функций и искомыми величинами (значениями функций или их производных в каких-то точках), относительно которых составляется алгебраическая система уравнений МКЭ.

5.1. Принцип поэлементного формирования алгебраической системы

Как известно, для успешного решения сложной математической проблемы надо в первую очередь ввести удачные обозначения и определения, с чего мы и начнем непосредственное изложение методов конечных элементов.

Пусть расчетная область Ω разбита на K непересекающихся конечных ячеек (элементов) Ω_k^h , $k = 1, \dots, K$, объединение которых $\Omega^h = \bigcup_{k=1}^K \Omega_k^h$ составляет сеточную область, представляющую дискретный вариант (дискретизацию) исходной области Ω . В частности, может быть $\Omega^h = \Omega$, но допускается и приближенное равенство $\Omega = \Omega^h \cup R^h \approx \Omega^h$, где объем «погрешности аппроксимации области» R^h есть величина $O(h^{-d\gamma})$, d — размерность Ω , γ — порядок аппроксимации области, а h — характерный шаг сетки, равный по порядку «усредненному» диаметру элементарной ячейки. Под характерным шагом обычно понимается некоторая величина $h = O(K^{-d})$ и, строго говоря, она имеет смысл для сеток, которые можно назвать *псевдоравномерными*, или регулярными, характеризуемыми тем, что при $K \rightarrow \infty$ ($h \rightarrow 0$) отношение максимального шага сетки к минимальному (*коэффициент неравномерности сетки*) остается величиной ограниченной.

Подчеркнем, что характер приближений в методах конечных элементов представляет интерес в двух отношениях. Во-первых, необходимо знать оценку погрешности для конкретной дискретизации области с конечной величиной шага сетки. Во-вторых, не менее важно, особенно в теоретическом плане, исследовать

асимптотическое поведение ошибки на последовательности сгущающихся сеток при $K \rightarrow \infty$ и $h \rightarrow 0$.

Пусть на выбранной сетке Ω^h приближенное решение u^h ищется в виде линейной комбинации финитных функций φ_i , $i = 1, \dots, N$, составляющих базис конечномерного пространства V_h :

$$u^h = \sum_{i=1}^N v_i^h \varphi_i. \quad (5.1)$$

Тогда алгебраическая система уравнений МКЭ (метода Рунге или Галеркина) записывается в виде

$$A^h v^h = f^h, \quad (5.2)$$

где $v^h = \{v_i^h\}$ — искомый вектор порядка N , а квадратная матрица A^h и известный вектор f^h того же порядка определяются формулами

$$A^h = \left\{ a_{ij} = a \left(\sum_{j=1}^N \varphi_j, \varphi_i \right) \right\}, \quad f^h = \{f_i^h = (f, \varphi_i)\}. \quad (5.3)$$

Соотношение (5.2) формально можно рассматривать как операторное уравнение в конечномерном пространстве, $A^h : V_h \rightarrow V_h$ — *сеточный оператор* дискретизированной краевой задачи, а искомый и заданный векторы v^h, f^h являются элементами пространства V_h , соответствующими выбранным базисным функциям φ_i .

Здесь билинейная и линейная формы выражаются через интегралы по сеточной области $\Omega^h = \bigcup_{k=1}^K \Omega_k^h$ (для общего вида неоднородных краевых условий необходимо еще включить их соответствующие вклады, выражающиеся через поверхностные интегралы по границе, но мы ради простоты эти аспекты пока опускаем). Очевидно, что такие интегралы можно выразить как сумму интегралов по конечным элементам Ω_k^h , каждый из которых будет содержать конечное число членов, соответствующих тем базисным функциям, которые не равны нулю на соответствующем элементе. Отсюда матрица и вектор правой части из (5.2), (5.3) могут быть представлены в виде

$$A^h = \sum_{k=1}^K A_k^h, \quad f^h = \sum_{k=1}^K f_k^h, \quad (5.4)$$

где ненулевые элементы A_k^h и f_k^h записываются в форме

$$a_{ij}^{(k)} = a\left(\sum_{j=1}^N \varphi_j, \varphi_i\right)^{(k)}, \quad f_i^{(k)} = (f, \varphi_i)^{(k)}. \quad (5.5)$$

Здесь верхний индекс (k) у билинейной или линейной формы означает номер соответствующего конечного элемента Ω_k^h , интеграл по которому берется в определении величин $a_{ij}^{(k)}$, $f_i^{(k)}$. Условно такие соотношения можно отобразить в следующей форме:

$$a(\cdot, \cdot)^{(k)} = \int_{\Omega_k^h}, \quad a(\cdot, \cdot) = \int_{\Omega^h} = \sum_k a(\cdot, \cdot)^{(k)}. \quad (5.6)$$

Пусть теперь N_k означает число базисных функций, не равных тождественно нулю на элементе Ω_k^h , а I_k — совокупность соответствующих «глобальных» номеров базисных функций φ_i . Тогда аппроксимацию (5.1) на k -м элементе, т. е. для $x \in \Omega_k^h$, можно записать в виде

$$u^h|_{\Omega_k^h} = \sum_{i \in I_k} v_i^h \varphi_i = \sum_{i'=1}^{N_k} v_{i'}^h \varphi_{i'}. \quad (5.7)$$

Здесь i' означают «локальные» номера ненулевых на Ω_k^h базисных функций. Число N_k можно назвать *порядком k -го конечного элемента*, но исторически за ним закрепилось наименование *число степеней свободы*.

Обозначим далее через $\bar{A}_k$, $\bar{f}_k^h$ квадратную матрицу и вектор порядка N_k , состоящие только из определенных в (5.5) величин, которые равны соответствующим ненулевым элементам «больших» A_k^h и f_k^h . Очевидно, что между введенными матрицами и векторами различных размерностей можно установить такие взаимосвязи:

$$A_k^h = P_k \bar{A}_k P_k^t, \quad f_k^h = P_k \bar{f}_k^h. \quad (5.8)$$

Здесь P_k — прямоугольная матрица (N строк, N_k столбцов) перенумераций, имеющая в каждом столбце один ненулевой элемент, равный единице, а P_k^T означает транспонированную матрицу, имеющую N_k строк и N столбцов.

Определение 5.1. Матрица $\bar{A}_k^h$ порядка N_k называется локальной матрицей метода конечных элементов (соответствующей элементу Ω_k^h), а матрица A^h — глобальной матрицей МКЭ.

Если ввести дополнительно *локальные векторы неизвестных* порядка N_k

$$\bar{v}_k^h = P_k^T v^h = \{v_i^h, i \in I_k\}, \quad (5.9)$$

то связанные с алгебраической системой (5.2) билинейные формы можно записать как

$$\begin{aligned} (A^h v^h, w^h) &= \left(\sum_k A_k^h v^h, w^h \right) \\ &= \left(\sum_k P_k \bar{A}_k^h P_k^T v^h, w^h \right) = \sum_k (\bar{A}_k^h \bar{v}_k^h, \bar{w}_k^h)^{(k)}, \end{aligned} \quad (5.10)$$

где $(\cdot, \cdot)^{(k)}$ означает скалярное произведение в N_k -мерном векторном евклидовом пространстве.

Одним из замечательных достижений метода конечных элементов является создание высокоэффективной в практическом отношении математической технологии формирования алгебраических систем уравнений на основе вычисления локальных матриц МКЭ и сборки (ассемблирования) глобальных матриц, что намного уменьшило вычислительную сложность алгоритмов и обеспечило унифицированный подход к реализации различного класса методов для самого широкого круга краевых задач, включая разные размерности, типы исходных уравнений и граничных условий, сложные конфигурации границ и функциональные зависимости коэффициентов.

Мы остановимся на этой важной проблеме подробнее, ради простоты вначале рассмотрим так называемые *лагранжевы конечные элементы*, когда в представлении (5.1) величины v_i^h соответствуют значениям (приближенным) искомого решения в каких-то точках сеточной области $x_i \in \Omega^h$, а каждая базисная функция обладает свойством $\varphi_i(x_j) = \delta_{i,j}$ ($\delta_{i,j}$ — символ Кронекера). Мы еще более конкретизируем постановку задачи дискретизации, предполагая, что точки x_i (узлы сетки) суть вершины конечных элементов Ω^h , которые являются многоугольниками

(треугольниками или прямоугольниками, например) в двумерном случае и полигонами (тетраэдрами, параллелепипедами) — в трехмерном. В этой ситуации каждому узлу сетки x_i можно сопоставить i -е уравнение системы уравнений МКЭ (5.2).

Пусть при каких-то выбранных упорядоченностях (не зависящих друг от друга) элементов Ω_k^h и узлов x_i через I_k обозначается совокупность узлов сетки, являющихся вершинами k -го элемента, а через K_i — совокупность номеров элементов, имеющих общую вершину — узел с номером i (*инцидентных* с узлом x_i). Предположим, опять же для упрощения изложения, что финитная базисная функция $\varphi_i(x)$ отлична от нуля только в элементах $\Omega_{k'}^h$ с номерами $k' \in K_i$.

В рассмотренных обозначениях и предпосылках существуют два альтернирующих алгоритма формирования алгебраической системы МКЭ: с поузловой и поэлементной обработкой данных.

Первый подход заключается в непосредственном определении всех элементов i -й строки матрицы A^h и i -го элемента вектора f^h элементов $\Omega_{k'}^h$, инцидентных с узлом x_i , и вычислении необходимых интегралов по этим элементам, организацией цикла по узлам сетки. Вторая стратегия — поэлементная — состоит в нахождении локальной матрицы $\bar{A}_k^h$ и *локального вектора правой части* $\bar{f}_k^h$ порядка N_k с помощью расчета всех требуемых интегралов по элементу Ω_k^h :

$$\begin{aligned}\bar{A}_k^h &= \left\{ a_{i',j'}^{(k)} = a \left(\sum_{j \in I_k} \varphi_j, \varphi_{i'} \right)^{(k)}, \quad i' \in I_k \right\}, \\ \bar{f}_k^h &= \{ (f, \varphi_{i'})^{(k)}, \quad i' \in I_k \}.\end{aligned}\quad (5.11)$$

При этом сборка глобальной матрицы сводится к организации цикла по элементам и последовательному суммированию (накоплению) элементов $a_{i,j}^{(k)}$, $f_{i'}^{(k)}$ с соответствующими элементами a_{ij} , f_i^h , что можно записать в форме псевдокода

$$\begin{aligned}f^h &:= 0, \quad A^h := 0, \quad k = 1, \dots, K: \\ A^h &:= A^h + A_k^h, \quad f^h := f^h + f_k^h, \\ a_{i,j} &:= a_{i,j} + a_{i',j'}^{(k)}, \quad f_i^h := f_i^h + f_{i'}^{(k)} \quad \text{для } i = i', j = j'.\end{aligned}\quad (5.12)$$

Огромное преимущество последнего способа заключается в тривиальном устранении дублирующих вычислений и логиче-

ской простоте алгоритма, что особенно ценно при его программировании. Другое важное достоинство, которое нельзя не упомянуть, — это возможность эффективного распараллеливания такого метода при его реализации на многопроцессорных вычислительных системах.

5.2. Построение одномерных уравнений МКЭ

Мы рассмотрим одномерную краевую задачу первого рода для дифференциального уравнения второго порядка достаточно общего вида:

$$\begin{aligned} -(du')' + bu' + cu &= f(x), & \alpha < x < \beta, \\ u|_{\alpha} &= u_{\alpha}, & u|_{\beta} &= u_{\beta}, \end{aligned} \quad (5.13)$$

где требования на свойства функций d , b , c , f конкретизируем несколько позже.

Для дискретизации задачи введем *неравномерную сетку* Ω^h

$$\alpha = x_0 < x_1 < \dots < x_N < x_{N+1} = \beta \quad (5.14)$$

с шагами $h_i = x_{i+1} - x_i$.

Для данной несамосопряженной задачи метод Галеркина нахождения приближенного обобщенного решения u^h в конечномерном подпространстве V_h определяется вариационным уравнением

$$\begin{aligned} a(u^h, v) &= (f, v) \quad \forall v \in V_h, \\ a(u^h, v) &= \int_{\alpha}^{\beta} [d(u^h)'v' + b(u^h)'v + cuv] dx, \quad (f, v) = \int_{\alpha}^{\beta} f v dx. \end{aligned} \quad (5.15)$$

5.2.1. Кусочно-линейная аппроксимация

Выберем сначала для простоты V_h как пространство *кусочно-линейных функций* на сетке Ω^h , линейных на каждом подинтервале (x_i, x_{i+1}) и обращающихся в нуль, естественно, на границах отрезка $[\alpha, \beta]$. Такое пространство имеет размерность N и *базисные финитные функции-«крышки»* $\varphi_i(x)$, каждая из которых отлична от нуля на интервале (x_{i-1}, x_{i+1}) , обладает свойством $\varphi_i(x_j) = \delta_{i,j}$ и описывается формулами

$$\varphi_i(x) = \begin{cases} (x - x_{i-1})/h_{i-1}, & x \in (x_{i-1}, x_i), \\ (x_{i+1} - x)/h_i, & x \in (x_i, x_{i+1}), \\ 0, & x \notin (x_{i-1}, x_{i+1}). \end{cases} \quad (5.16)$$

Графики таких функций изображены на рис. 5.1, а их первые производные, очевидно, равны соответственно $+1$, -1 и нулю.

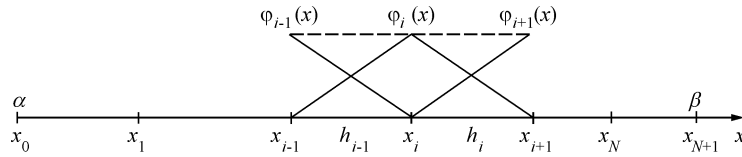


Рис. 5.1. Базисные кусочно-линейные функции-«крышки»

Поскольку искомое решение удовлетворяет неоднородным условиям Дирихле, его приближение будем искать на множестве $\tilde{V}_h$, полученном из пространства V_h сдвигом на линейную функцию $w(x) = u_\alpha + \frac{x - \alpha}{\beta - \alpha}(u_\beta - u_\alpha)$. Это означает, что приближенное решение можно искать в виде разложения по базисным функциям

$$u^h(x) = u_\alpha \varphi_0 + \sum_{i=1}^N v_i \varphi_i + u_\beta \varphi_{N+1}, \quad (5.17)$$

где неизвестные $v_1, \dots, v_N$ должны по смыслу аппроксимировать значения точного решения $u_i = u(x_i)$. Тогда конкретизация уравнений (5.15) метода Галеркина сводится к системе

$$a(u^h, \varphi_i) = (f, \varphi_i), \quad i = 1, 2, \dots, N, \quad (5.18)$$

которая в матричной записи имеет следующую форму:

$$A^h v^h = f^h. \quad (5.19)$$

Здесь $A^h = \{a_{i,j}\}$, $v^h = \{v_i\}$, $f^h = \{f_i^h\}$ — квадратная (глобальная) матрица МКЭ и векторы порядка N , элементы которых равны

$$\begin{aligned} a_{i,j} &= a(\varphi_i, \varphi_j) = \int_{\alpha}^{\beta} (d\varphi_i' \varphi_j' + b\varphi_i \varphi_j' + c\varphi_i \varphi_j) dx, \\ f_i^h &= \int_{\alpha}^{\beta} f \varphi_i dx - g_i, \quad i, j = 1, \dots, N, \end{aligned} \quad (5.20)$$

где величины $g_i = 0$ для $i = 2, \dots, N-1$, а в околограничных узлах описываются формулами

$$g_1 = u_\alpha a_{1,0}, \quad g_N = u_\beta a_{N,N+1}.$$

Здесь $a_{1,0}$, $a_{N,N+1}$ определяются соотношениями (5.20) при $j = 0$ и $j = N + 1$ соответственно. Значения g_1, g_N обусловлены наличием неоднородных граничных условий Дирихле в краевой задаче (5.13) и образуются изначально в (5.18) при интегрировании по крайним сеточным интервалам, после чего они переносятся из левой части уравнений в правую.

Рассмотрим теперь подробнее структуру матрицы алгебраической системы, которую можно представить суммой

$$A^h = A_s^h + A_c^h + A_m^h, \quad (5.21)$$

где каждая из слагаемых матриц определяется соответствующими членами билинейной формы (5.20):

$$\begin{aligned} A_s^h &= \left\{ a_{i,j}^{(s)} = \int_{\alpha}^{\beta} d\varphi'_i \varphi'_j dx \right\}, \\ A_c^h &= \left\{ a_{i,j}^{(c)} = \int_{\alpha}^{\beta} b \varphi_i \varphi'_j dx \right\}, \\ A_m^h &= \left\{ a_{i,j}^{(m)} = \int_{\alpha}^{\beta} c \varphi_i \varphi_j dx \right\}. \end{aligned} \quad (5.22)$$

За матрицами A_s^h и A_m^h исторически закрепились названия *матрица жесткости* и *матрица масс* (введенные, по-видимому, специалистами по приложению МКЭ к задачам теории упругости), а матрицу A_c^h мы назовем *матрицей конвекции*.

Перейдем теперь от рассмотрения глобальных матриц МКЭ к построению *локальных* матриц. Пусть элементарный сеточный отрезок $\Omega_i^h = (x_i, x_{i+1})$, $i = 0, 1, \dots, N$, означает i -й конечный элемент. Поскольку на каждом из них отличны от нуля только две базисные функции из V_h (а на крайних — по одной), то матрица МКЭ может быть представлена следующей суммой:

$$A^h = A_0^h + \sum_{i=1}^{N-1} A_i^h + A_N^h, \quad (5.23)$$

где крайние слагаемые A_0^h , A_N^h содержат по одному ненулевому диагональному элементу: $a_{1,1}^{(0)} = a(\varphi_1, \varphi_1)^{(0)}$, $a_{N,N}^{(N)} = a(\varphi_N, \varphi_N)^{(N)}$

соответственно, а остальные — по четыре элемента вида $a_{i',i''}^{(i)} = a(\varphi_{i'}, \varphi_{i''})^{(i)}$ (здесь $i' = i, i+1$, $i'' = i, i+1$, а верхний индекс означает, что интегралы в билинейной форме берутся по интервалу (x_i, x_{i+1})). Матрицам A_i^h из (5.23) можно сопоставить локальные матрицы $\bar{A}_i^h$ ($\bar{A}_0^h$ и $\bar{A}_N^h$ — первого порядка, а остальные — второго), каждая из которых связана с интегралами только по элементу Ω_i^h .

Естественно, что любая локальная матрица МКЭ разбивается на сумму локальных матриц жесткости, конвекции и масс:

$$\begin{aligned}\bar{A}_i^h &= \bar{A}_{s,i}^h + \bar{A}_{c,i}^h + \bar{A}_{m,i}^h = \{a_{i',i''}^{(i)} = a_{i',i''}^{s,i} + a_{i',i''}^{c,i} + a_{i',i''}^{m,i}\}, \\ \bar{A}_{s,i}^h &= \left\{ a_{i',i''}^{s,i} = \int_{x_i}^{x_{i+1}} p \varphi_{i'}' \varphi_{i''}' dx \right\}, \\ \bar{A}_{c,i}^h &= \left\{ a_{i',i''}^{c,i} = \int_{x_i}^{x_{i+1}} d \varphi_{i'} \varphi_{i''}' dx \right\}, \\ \bar{A}_{m,i}^h &= \left\{ a_{i',i''}^{m,i} = \int_{x_i}^{x_{i+1}} c \varphi_{i'} \varphi_{i''} dx \right\}.\end{aligned}\tag{5.24}$$

Здесь индексы i' , i'' принимают значения или i , или $i+1$. Поскольку производные базисных функций кусочно-постоянны:

$$\varphi_i' = \begin{cases} -h_i^{-1}, & x \in (x_i, x_{i+1}), \\ h_{i-1}^{-1}, & x \in (x_{i-1}, x_i), \\ 0, & x \notin (x_{i-1}, x_{i+1}), \end{cases}\tag{5.25}$$

то способ вычисления элементов матриц МКЭ определяется свойствами коэффициентов d , b , c — в некоторых ситуациях интегралы из (5.24) могут быть найдены точно, но в общем случае для их вычисления необходимо привлекать приближенные квадратурные формулы.

Если эти коэффициенты постоянны на каждом элементе, т. е.

$$\begin{aligned}p(x) &= p_{i+1/2}, & b(x) &= b_{i+1/2}, \\ c(x) &= c_{i+1/2}, & x &\in (x_i, x_{i+1})\end{aligned}\tag{5.26}$$

(при этом даже возможны их разрывы в узлах сетки), то элементы локальных матриц выражаются особенно просто:

$$\begin{aligned}
\bar{A}_{s,i}^h &= \frac{d_{i+1/2}}{h_i} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, & \bar{A}_{c,i}^h &= \frac{b_{i+1/2}}{2} \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}, \\
\bar{A}_{m,i}^h &= c_{i+1/2} \frac{h_i}{6} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, & i &= 1, 2, \dots, N-1, \\
\bar{A}_{s,0}^h &= d_{1/2} h_0^{-1}, & \bar{A}_{c,0}^h &= -b_{1/2}/2, & \bar{A}_{m,0}^h &= 2c_{1/2} h_0, \\
\bar{A}_{s,N}^h &= d_{N+1/2} h_N^{-1}, & \bar{A}_{c,N}^h &= b_{N+1/2}/2, & \bar{A}_{m,N}^h &= 2c_{N+1/2} h_N.
\end{aligned} \tag{5.27}$$

Отсюда получаем, что для i -го узла алгебраическое уравнение метода Галеркина имеет следующий вид:

$$\begin{aligned}
&\frac{d_{i-1/2}}{h_{i-1}}(v_i - v_{i-1}) + \frac{d_{i+1/2}}{h_i}(v_i - v_{i+1}) + \frac{b_{i-1/2}}{2}(v_i + v_{i-1}) - \\
&\quad \frac{b_{i+1/2}}{2}(v_i + v_{i+1}) + \frac{1}{6}[h_{i-1}c_{i-1/2}(2v_i + v_{i-1}) + \\
&\quad h_i c_{i+1/2}(2v_i + v_{i+1})] = f_i^h, \quad i = 1, 2, \dots, N.
\end{aligned} \tag{5.28}$$

Учет краевых условий Дирихле для уравнений, соответствующих околোগраничным узлам, делается здесь тривиально — в первом уравнении надо положить $v_0 = u_a$ и перенести соответствующий член в правую часть, а в последнем уравнении аналогично полагаем $v_{N+1} = u_b$.

Теперь коснемся вопроса о вычислении правых частей f_i^h . Поскольку приближенное решение мы строим в виде кусочно-линейной функции, то напрашивается такое же представление для функции $f(x)$ исходной задачи:

$$f(x) = \sum_{i=0}^{N+1} f_i \varphi_i(x) + \psi(x), \quad f_i = f(x_i), \quad |\psi(x)| = O(h^2).$$

Тогда правая часть уравнения (5.28) имеет вид

$$f_i^h = \frac{1}{6}[h_{i-1}f_{i-1} + 2(h_{i-1} + h_i)f_i + h_i f_{i+1}], \tag{5.29}$$

а ее вычисление, очевидно, также может делаться поэлементно, при этом рассчитываются те же интегралы от произведений базисных функций, которые входят в элементы локальных матриц масс.

Система уравнений (5.28) после учета краевых условий Дирихле записывается в векторно-матричном виде

$$A^h v^h = f^h,$$

где векторы $v^h = (v_1, \dots, v_N)^t$, $f^h = (f_1, \dots, f_N)^t$ имеют размерность по числу внутренних узлов сетки, а матрица A^h является трехдиагональной:

$$\begin{bmatrix} a_{1,1} & a_{1,2} & & & \\ & a_{2,1} & a_{2,2} & & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & & & a_{N-1,N} & \\ & & a_{N,N-1} & a_{N,N} & \end{bmatrix}.$$

Ее элементы определяются из (5.28) следующим образом:

$$\begin{aligned} a_{i,i} &= \frac{d_{i-1/2}}{h_{i-1}} + \frac{d_{i+1/2}}{h_i} + \frac{b_{i-1/2} - b_{i+1/2}}{2} + \frac{h_{i-1}c_{i-1/2} + h_i c_{i+1/2}}{3}, \\ a_{i,i-1} &= \frac{d_{i-1/2}}{h_{i-1}} + \frac{b_{i-1/2}}{2} + \frac{h_{i-1}c_{i+1/2}}{6}, \\ a_{i,i+1} &= \frac{d_{i+1/2}}{h_i} - \frac{b_{i+1/2}}{2} + \frac{h_i c_{i+1/2}}{6}. \end{aligned}$$

Сравнение соотношений (5.28), (5.29) с конечно-разностными уравнениями, непосредственно аппроксимирующими диффузионно-конвективное уравнение (5.13), см., например, [81], показывает, что левая часть (5.28), после его деления на «объем ячейки» $\frac{h_{i-1} + h_i}{2}$, будет та же, что у разностного уравнения, использующего аппроксимацию первой производной центральной разностью (после перехода к равномерной сетке). Правая же часть f_i^h уравнений метода Галеркина содержит необъяснимое (с «разностной точки зрения») использование усредненных узловых значений функции $f(x)$. Отметим также, что в отсутствие конвективного члена ($b(x) = 0$) уравнение (5.28) является симметричным, в отличие от разностного уравнения, даже на неравномерной сетке.

5.2.2. Лангражевые аппроксимации высших порядков

Естественным развитием кусочно-линейной аппроксимации является переход к *интерполяционным многочленам Лагранжа*, которые при использовании $n + 1$ различного узла имеют порядок n и описываются формулой

$$L_n(x) = \sum_{l=0}^n f_l \psi_l(x), \quad (5.30)$$

где $f_l = f(x_l)$ — узловые значения интерполируемой функции, а $\psi_l(x)$ — базисные лагранжевые многочлены n -го порядка (см. [24], [82]):

$$\psi_l(x) = \frac{(x - x_0) \cdots (x - x_{l-1})(x - x_{l+1}) \cdots (x - x_n)}{(x_l - x_0) \cdots (x_l - x_{l-1})(x_l - x_{l+1}) \cdots (x_l - x_n)}, \quad (5.31)$$

обладающие свойством $\psi_l(x_m) = \delta_{lm}$. Аппроксимационные свойства многочлена (5.30) рассмотрим ниже в п. 5.1.

Финитные базисные функции МКЭ на основе лагранжевых интерполянтов строятся следующим образом. На каждом сеточном интервале $[x_i, x_{i+1}]$ выбираются равномерно (для простоты) расположенные интерполяционные узлы

$$x_0^{(i)} = x_i, \quad x_n^{(i)} = x_{i+1}, \quad x_l^{(i)} = x_i + \frac{lh_i}{n}, \quad l = 1, \dots, n-1,$$

и каждому из них сопоставляется базисная лагранжевая функция $\psi_l^{(i)}$, получаемая из (5.31) приписыванием каждому интерполяционному узлу верхнего индекса i . С их помощью формируются финитные базисные функции двух типов: внутренние и узловые. Каждая внутренняя функция соответствует внутренней точке x_l сеточного отрезка и определяется как $\varphi_l^{(i)} = \psi_l^{(i)}$. Базисную функцию, соответствующую узлу x_i , обозначаем с помощью одного индекса — $\varphi^{(i)}$. Носителем каждой внутренней базисной функции $\varphi_l^{(i)}$, $l = 1, \dots, n-1$, является интервал (x_i, x_{i+1}) , а узловой — два смежных интервала:

$$\begin{aligned} \varphi_l^{(i)} &= 0, \quad x \notin (x_i, x_{i+1}), \quad l = 1, \dots, n-1, \\ \varphi^{(i)} &= \begin{cases} 0, & x \notin (x_{i-1}, x_{i+1}), \\ \varphi_0^{(i)}, & x \in (x_i, x_{i+1}), \\ \varphi_n^{(i-1)}, & x \in (x_{i-1}, x_i). \end{cases} \end{aligned}$$

Для сетки (5.14) с $(n-1)$ -й внутренней точкой на каждом из $N+1$ сеточных интервалов общее число базисных функций будет равно

$$\bar{N} = (n-1)(N+1) + N + 2 = n(N+1) + 1,$$

и их можно занумеровать одним индексом:

$$\varphi_k(x) = \varphi_l^{(i)}(x), \quad k = i n + l, \quad k = 0, 1, \dots, \bar{N} - 1.$$

Соответственно можно упорядочить сквозным образом все узлы и внутренние интерполяционные точки:

$$x_k = x_l^{(i)}, \quad k = 0, 1, \dots, \bar{N} - 1.$$

Тогда приближенное решение для задачи Дирихле (5.13) по аналогии с (5.17) можно искать в виде

$$u^h(x) = u_\alpha \varphi_0(x) + \sum_{k=1}^{\bar{N}-2} v_k \varphi_k(x) + u_\beta \varphi_{\bar{N}-1}(x), \quad (5.32)$$

где величины v_k , по своему смыслу, должны аппроксимировать значения точного решения $u(x_k)$.

Алгебраическая система уравнений МКЭ в этом случае имеет также вид (5.19), (5.20), с точностью до определения базисных функций (формально надо заменить целые величины i, j, N на $k, k', \bar{N} - 2$ соответственно), а ее порядок равен $\bar{N} - 2$.

В данном случае локальные матрицы МКЭ имеют порядок $n+1$, по числу ненулевых базисных функций на одном интервале $[x_i, x_{i+1}]$. Приведем для примера *лагранжевые базисные функции* второго порядка, вид которых изображен на рис. 5.2:

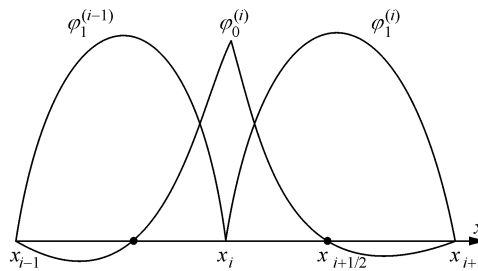


Рис. 5.2. Кусочно-квадратичные базисные функции

$$\varphi^{(i)} = \begin{cases} 1 - 3\frac{x-x_i}{h_i} + 2\left(\frac{x-x_i}{h_i}\right)^2, & x \in (x_i, x_{i+1}), \\ 1 + 3\frac{x-x_i}{h_{i-1}} + 2\left(\frac{x-x_i}{h_{i-1}}\right)^2, & x \in (x_{i-1}, x_i), \end{cases} \quad (5.33)$$

$$\varphi_1^{(i)} = 4\left[\frac{x-x_i}{h_i} - \left(\frac{x-x_i}{h_i}\right)^2\right], \quad x \in (x_i, x_{i+1}).$$

Соответствующая сетка из «основных» и «вспомогательных» узлов приведена на рис. 5.3.

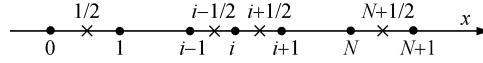


Рис. 5.3. Схема расположения основных и вспомогательных узлов для лагранжевых конечных элементов второго порядка

Локальная матрица жесткости (при постоянном коэффициенте $d(x)$ на интервале (x_i, x_{i+1})) в этом случае имеет следующий вид (соответствующий локальной упорядоченности неизвестных как $v_i, v_{i+1/2}, v_{i+1}$):

$$\bar{A}_{s,i}^h = \frac{p_{i+1/2}}{h_i} \begin{bmatrix} 7 & -8 & 1 \\ -8 & 16 & -8 \\ 1 & -8 & 7 \end{bmatrix}.$$

Если же коэффициент $d(x)$ зависит от x на интервале $[x_i, x_{i+1}]$, то данная матрица является простейшим приближением к точной матрице жесткости

$$\bar{A}_{s,i}^h = \left\{ a_{l,l'}^{(i)} = \int_{x_i}^{x_{i+1}} p(x) \frac{d\varphi_l^{(i)}}{dx} \frac{d\varphi_{l'}^{(i)}}{dx} dx; \quad l, l' = 0, 1, 2 \right\}.$$

Как легко обнаружить, получаемые алгебраические уравнения имеют разную форму для узловых точек сетки и срединных точек (рассмотрим ради краткости исходную задачу с нулевыми коэффициентами $b(x)$ и $c(x)$):

$$\begin{aligned} \frac{d_{i+1/2}}{h_i} (-8v_i + 16v_{i+1/2} - 8v_{i+1}) &= (f, \varphi_1^{(i)}), & i = 0, \dots, N-1, \\ \frac{d_{i-1/2}}{h_{i-1}} (v_{i-1} - 8v_{i-1/2} + 7v_i) + \frac{d_{i+1/2}}{h_i} (7v_i - 8v_{i+1/2} + v_{i+1}) &= (f, \varphi_0^{(i)}), \\ i &= 1, \dots, N. \end{aligned} \quad (5.34)$$

В этих уравнениях для $i = 0, 1, N$ вместо величин v_0 и v_{N+1} надо подставить известные из граничных условий значения u_α и u_β , а затем соответствующие члены перенести в правые части.

Получаемая матрица системы уравнений для $N = 4$ приведена ниже, где точками и звездочками изображены ненулевые элементы, соответствующие основным и вспомогательным узлам:

$$A^h = \begin{bmatrix} \star & \cdot & & & & & & & & \\ \star & \cdot & \star & \cdot & & & & & & \\ & \cdot & \star & \cdot & & & & & & \\ & & \cdot & \star & \cdot & \star & \cdot & & & \\ & & & \cdot & \star & \cdot & & & & \\ & & & & \cdot & \star & \cdot & \star & \cdot & \\ & & & & & \cdot & \star & \cdot & & \\ & & & & & & \cdot & \star & \cdot & \\ & & & & & & & \cdot & \star & \\ & & & & & & & & \cdot & \star \end{bmatrix}.$$

Здесь полезно сделать одно замечание относительно системы уравнений (5.34): из трехчленных уравнений для срединных точек неизвестные $v_{i\pm 1/2}$ могут быть легко выражены через узловые неизвестные и исключены из остальных уравнений, в результате чего мы редуцируем алгебраическую задачу к системе N трехчленных уравнений относительно искомых v_i .

Этот прием получил название *конденсация* и имеет общую применимость для лагранжевых элементов любого порядка. Все «внутренние» неизвестные из одного сеточного интервала могут быть выражены через соответствующие узловые неизвестные (для этого надо обратить матрицы $(n-1)$ -го порядка), и после их исключения мы приходим к трехчленной системе уравнения для v_i .

Остановимся на этом вопросе подробнее. Вводя по аналогии с обозначениями базисных функций «узловые» и «внутренние» неизвестные

$$v_i, v_l^{(i)}, \quad l = 1, \dots, n-1, \quad i = 0, \dots, N,$$

выпишем соответственно «узловые» и «внутренние» уравнения:

$$a_{i,i-1}v_{i-1} + \sum_{l=1}^{n-1} a_{i,l}^{(i-1)}v_l^{(i-1)} + a_{i,i}v_i + \sum_{l=1}^{n-1} a_{i,l}^{(i)}v_l^{(i)} + a_{i,i+1}v_{i+1} = f_i^h, \\ i = 1, \dots, N,$$

$$a_{l,0}^{(i)} v_i + \sum_{l'=1}^{n-1} a_{l,l'}^{(i)} v_{l'}^{(i)} + a_{l,n}^{(i)} v_{i+1} = f_{l,i}^h, \quad i=0, \dots, N, \quad l=1, \dots, n-1.$$

Здесь $a_{l,l'}^{(i)}$ — определяемые в соответствии с (5.24) элементы локальных матриц $\bar{A}_i^h$, а коэффициенты «узловых» уравнений равны (с учетом обозначений в (5.15), (5.24) и равенств $\varphi^{(i)} = \varphi_0^{(i)} = \varphi_n^{(i-1)}$)

$$\begin{aligned} a_{i,i-1} &= a(\varphi^{(i-1)}, \varphi^{(i)}) = a_{0,n}^{(i-1)}, \\ a_{i,i+1} &= a(\varphi^{(i+1)}, \varphi^{(i)}) = a_{n,0}^{(i)}, \\ a_{i,i} &= a(\varphi^{(i)}, \varphi^{(i)}) = a_{n,n}^{(i-1)} + a_{0,0}^{(i)}. \end{aligned}$$

Для небольших значений n «внутренние» неизвестные могут быть относительно легко выражены через «узловые» неизвестные с помощью «своих» $n-1$ уравнений:

$$v_l^{(i)} = p_l^{(i)} v_i + q_l^{(i)} v_{i+1} + g_l^{(i)}, \quad l=1, \dots, n-1.$$

С помощью таких соотношений из «узловых» уравнений исключаются все $v_l^{(i-1)}, v_l^{(i)}, l=1, \dots, n-1$, что и составляет процедуру конденсации, после чего получаем систему трехточечных уравнений

$$\bar{a}_{i,i-1} v_{i-1} + \bar{a}_{i,i} v_i + \bar{a}_{i,i+1} v_{i+1} = \bar{f}_i, \quad i=1, \dots, N.$$

Формулы для вычисляемых по ходу дела коэффициентов $p_l^{(i)}, q_l^{(i)}, g_l^{(i)}$ и $\bar{a}_{i,i\pm 1}, \bar{a}_{i,i}, \bar{f}_i$ для краткости опускаем.

Отметим, что структура рассмотренных соотношений принципиально не меняется, если на каждом сеточном интервале $[x_i, x_{i+1}]$ брать разные порядки n_i интерполяционных многочленов.

Остановимся теперь на особенностях построения уравнений МКЭ при наличии граничных условий, содержащих производные от искомого решения. Пусть, например, в задаче (5.13) на правой границе вместо условия Дирихле задано краевое условие третьего рода

$$\alpha u + \frac{\partial u}{\partial x} = \gamma, \quad x = \beta.$$

В данном случае представление приближенного решения (5.32) вместо известного значения u_β содержит неизвестную величину $v_{\bar{N}-1}$, а порядок алгебраической системы (5.19) становится равным $\bar{N} - 1$. При этом исходные уравнения (5.18) метода Галеркина, вследствие формулировки обобщенного решения, приобретают вид

$$a(u^h, \varphi_k) = (f, \varphi_k) - g_k, \quad k = 0, 1, \dots, \bar{N} - 1,$$

где первые n значений g_k равны (мы полагаем для простоты $b = c = 0$):

$$g_k = u_\alpha \int_{x_0}^{x_1} d(x) \varphi_0'(x) \varphi_k'(x) dx = u_\alpha \int_{x_0}^{x_1} d(x) \varphi_0'(x) (\varphi_l^{(0)}(x))' dx, \\ k, l = 1, 2, \dots, n,$$

вследствие учета условия Дирихле на левой границе, «средние» значения g_k , $k = n+1, \dots, \bar{N}-2$, — нулевые, а последнее, получаемое с учетом условия третьего рода на правой границе, имеет вид

$$g_{\bar{N}-1} = d(x_{N+1}) \frac{du^h}{dx}(x_{N+1}) \cdot \varphi_{N+1}^{(0)}(x_{N+1}) = d_{N+1} \gamma - \varkappa v_{\bar{N}-1}.$$

5.2.3. Эрмитовые базисные функции

Общим свойством лагранжевых конечных элементов любого порядка является непрерывность производных только нулевого порядка для определяемого на (α, β) приближенного решения, в силу самого способа его построения. Естественное развитие такого подхода заключается в переходе к эрмитовой интерполяции, использующей так называемые кратные узлы, в которых задаются значения не только приближаемой функции, но и ее производных.

Очень простой и самой распространенной здесь является *кубическая эрмитова интерполяция*, которая определяется по четырем задаваемым величинам $v_i, v'_i, v_{i+1}, v'_{i+1}$ (см. [82]):

$$v^h(x) = v_i \varphi\left(\frac{x - x_i}{h_i}\right) + h_i v'_i \psi\left(\frac{x - x_i}{h_i}\right) + \\ v_{i+1} \varphi\left(\frac{x_{i+1} - x}{h_i}\right) + h_i v'_{i+1} \psi\left(\frac{x_{i+1} - x}{h_i}\right). \quad (5.35)$$

Здесь φ и ψ — финитные базисные функции с носителем (x_{i-1}, x_{i+1}) , описываемые формулами

$$\begin{aligned}\varphi(\xi) &= (|\xi| - 1)^2(2|\xi| + 1), \\ \psi(\xi) &= \xi(|\xi| - 1)^2, \quad \xi = \frac{2x - x_{i-1} + x_{i+1}}{x_{i+1} - x_i},\end{aligned}$$

и обладающие свойствами

$$\varphi(0) = \psi'(0) = 1, \quad \varphi(\pm 1) = \psi(\pm 1) = \varphi'(\pm 1) = \psi'(0) = \psi'(\pm 1) = 0.$$

Графики этих функций приведены на рис. 5.4, а локальная погрешность аппроксимации многочленом некоторой функции

$$w(x) \in C^4[x_i, x_{i+1}],$$

$$w(x_i) = v_i, \quad w'(x_i) = v'_{i+1}, \quad w(x_{i+1}) = v_{i+1}, \quad w'(x_{i+1}) = v'_{i+1}$$

равна

$$z(x) = w(x) - v^h(x) = \frac{w^{(4)}(\eta)}{4!}(x - x_i)^2(x - x_{i+1})^2, \quad \eta \in [x_i, x_{i+1}].$$

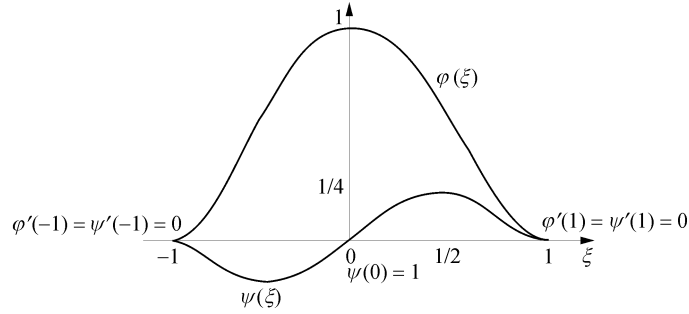


Рис. 5.4. Кусочно-квадратичные базисные функции

Локальные матрицы жесткости и масс (четвертого порядка) в данном случае, при кусочно-постоянных коэффициентах и упорядоченности неизвестных на элементе Ω_i^h в последовательности $v_i, v'_i, v_{i+1}, v'_{i+1}$, имеют следующий вид:

$$\begin{aligned}\bar{A}_{s,i}^h &= \frac{d_{i+1/2}}{30h_i} \begin{bmatrix} 36 & 3h & -36 & 3h \\ \cdot & 4h^2 & -3h & -h \\ \cdot & \cdot & 36 & -3h \\ \cdot & \cdot & \cdot & 4h^2 \end{bmatrix}, \\ \bar{A}_{m,i}^h &= \frac{h_i c_{i+1/2}}{420} \begin{bmatrix} 156 & 22h & 54 & -13h \\ \cdot & 4h^2 & 13h & -3h^2 \\ \cdot & \cdot & 156 & -22h \\ \cdot & \cdot & \cdot & 4h^2 \end{bmatrix}.\end{aligned}\tag{5.36}$$

Матрица $\bar{A}_{m,i}^h$ положительно определена, а локальная матрица жесткости $\bar{A}_{s,i}^h$ положительно полуопределена и имеет нулевое собственное значение, соответствующее функции $v^h(x) \equiv 1$, т. е. локальному собственному вектору $(1, 0, 1, 0)'$.

В случае равномерной сетки с шагом $h_i = h$, постоянных коэффициентов d, c и нулевого конвективного члена ($b(x) = 0$) алгебраические уравнения получаемой системы имеют два характерных вида (их можно назвать уравнениями для искомой функции v_i и ее производной v'_i):

$$\begin{aligned} & \frac{d}{30h} (-36v_{i-1} - 3hv'_{i-1} + 72v_i - 36v_{i+1} + 3hv'_{i+1}) + \\ & \frac{ch}{420} (54v_{i-1} + 13hv'_{i-1} + 312v_i + 54v_{i+1} - 13hv'_{i+1}) \\ & = f_i^h = \int_{\alpha}^{\beta} f \varphi \left(\frac{x}{h} - i \right) dx, \\ & \frac{d}{30h} (3v_{i-1} - hv'_{i-1} + 8hv'_i - 3v_{i+1} - hv'_{i+1}) + \\ & \frac{ch^2}{420} (-13v_{i-1} - 3hv'_{i-1} + 8hv'_i + 13v_{i+1} - 3hv'_{i+1}) \\ & = f'_i = h \int_{\alpha}^{\beta} f \psi \left(\frac{x}{h} - i \right) dx. \end{aligned} \quad (5.37)$$

Интересно рассмотреть эти уравнения с точки зрения разностной аппроксимации исходного дифференциального уравнения. Предположим, что $d = 1, b = c = 0, f = 1$, так что решаемое уравнение Пуассона есть $-u'' = 1$. Тогда из (5.37) мы имеем

$$\begin{aligned} & -\frac{6}{5} \frac{v_{i+1} - 2v_i + v_{i-1}}{h^2} + \frac{1}{5} \frac{v'_{i+1} - v'_{i-1}}{2h} = 1, \\ & -\frac{1}{5} \frac{v_{i+1} - v_{i-1}}{2h} + \frac{1}{5} v'_i - \frac{1}{30} (v'_{i+1} - 2v'_i + v'_{i-1}) = 0. \end{aligned} \quad (5.38)$$

Учет краевых условий в системе (5.37) делается просто. Если заданы условия первого рода, то в уравнениях для $i = 1$ и N вместо v_0 и v_{N+1} подставляются u_{α} и u_{β} соответственно. Если же, например, слева задано значение производной u'_{α} , то для $i = 1$ полагается $v'_0 = u'_{\alpha}$. Во всех возможных случаях мы получаем $2N + 2$ уравнений для такого же числа неизвестных.

Отсюда с помощью тейлоровских разложений получаем, что первое сеточное уравнение аппроксимирует уравнение Пуассона с погрешностью $-\frac{h^2}{15}u^{(4)}$, а второе уравнение асимптотически дает ошибку $-\frac{h^4}{1230}u^{(5)}$.

Рассмотренный подход формально без труда переносится на эрмитовые базисные функции более высоких порядков, однако он связан с такой неприятностью, как повышение порядка локальных матриц МКЭ и усложнением структуры (с уменьшением степени разреженности) глобальной матрицы алгебраической системы уравнений.

5.2.4. Иерархические базисы

Эффективные методы приближенных вычислений основаны, как правило, на последовательном уточнении искомых решений. При этом, если полученная на каком-то этапе погрешность представляется слишком большой, было бы желательно повысить точность за счет дополнительных незначительных вычислений.

Рассмотренные выше системы базисных функций высоких степеней такой возможностью, к сожалению, не обладают. Если, например, после использования лагранжевой интерполяции n -го порядка перейти к базисным функциям степени $n + 1$, то все элементы матриц МКЭ приходится рассчитывать заново.

В этом плане более предпочтительными являются *иерархические базисы*, активно развиваемые в 1970-80 гг. И. Бабушкой и его последователями, см. [154]. Главные свойство и достоинство таких базисов заключается в том, что при увеличении порядка аппроксимирующего пространства предыдущие базисные функции остаются неизменными и к ним только добавляются новые.

Описание таких базисов проведем на *стандартном элементе*

$$\Omega_{st} = \{-1 < \xi < 1\},$$

который отображается в произвольный интервал (x_i, x_{i+1}) с помощью формулы

$$x = Q_i(\xi) \equiv \frac{1-\xi}{2}x_i + \frac{1+\xi}{2}x_{i+1}, \quad \xi \in \Omega_{st}. \quad (5.39)$$

При этом обратное отображение имеет вид

$$\xi = Q_i^{-1}(x) = \frac{2x - x_i - x_{i+1}}{x_{i+1} - x_i}. \quad (5.40)$$

Базис p -го порядка включает $p + 1$ функцию:

$$\begin{aligned} \varphi_1(\xi) = \frac{1-\xi}{2}, \quad \varphi_2(\xi) = \frac{1+\xi}{2}, \quad \varphi_l(\xi) = \psi_{l-1}(\xi), \\ l = 3, \dots, p+1, \end{aligned} \quad (5.41)$$

где вспомогательная функция $\psi_l(\xi)$ определяется через *многочлен Лежандра* P_{l-1} :

$$\psi_l(\xi) = \sqrt{\frac{2l-1}{2}} \int_1^\xi P_{l-1}(t) dt, \quad l = 2, 3, \dots \quad (5.42)$$

Напомним, что многочлены Лежандра ортогональны на рассматриваемом интервале

$$\int_{-1}^1 P_i(x) P_j(x) dx = \frac{2}{2i+1} \delta_{i,j}$$

и могут быть вычислены из трехчленной рекуррентной формулы

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), \quad n = 1, 2, \dots,$$

где первые многочлены имеют вид (см. [24], [82])

$$P_0(x) = 1, \quad P_1(x) = x, \quad P_2(x) = \frac{1}{2}(3x^2 - 1).$$

Многочлены Лежандра также удовлетворяют следующим соотношениям:

$$(2n+1)P_n(x) = P'_{n+1}(x) - P'_{n-1}(x), \quad n = 1, 2, \dots$$

Все корни x_i многочлена Лагранжа $P_n(x)$ находятся в интервале $(-1, 1)$ и являются абсциссами n -точечной гауссовской квадратурной формулы

$$\int_{-1}^1 f dx \approx \sum_{i=1}^n w_i f(x_i).$$

Базисные функции φ_1, φ_2 называются узловыми (они обладают «лагранжевыми» свойствами $\varphi_1(-1) = 1, \varphi_1(1) = 0, \varphi_2(1) = 1, \varphi_2(-1) = 0$), а остальные — внутренними.

На основе свойств многочленов Лежандра показывается, что справедливо соотношение

$$\psi_l(\xi) = \frac{1}{\sqrt{2(2l-1)}} (P_l(\xi) - P_{l-2}(\xi)), \quad (5.43)$$

т. е. функция $\psi_l(\xi)$ является многочленом степени l . Легко также убедиться в том, что

$$\begin{aligned} \psi_i(-1) = \psi_i(1) = 0, \quad i = 2, 3, \dots, \\ \int_{-1}^{+1} \frac{d\varphi_k}{d\xi} \frac{d\psi_i}{d\xi} d\xi = 0, \quad k = 1, 2; \quad \int_{-1}^1 \frac{d\psi_i}{d\xi} \frac{d\psi_j}{d\xi} d\xi = \delta_{i,j}. \end{aligned} \quad (5.44)$$

В силу последних свойств локальные матрицы жесткости при постоянном коэффициенте $d(x)$ на интервале (x_i, x_{i+1}) вычисляются особенно легко, так как нахождение их диагональных коэффициентов не требует даже интегрирования:

$$\bar{A}_{s,i}^h = \frac{2d_{i+1/2}}{h_i} \begin{bmatrix} 1/2 & -1/2 & 0 & \cdot & 0 \\ -1/2 & 1/2 & 0 & \cdot & 0 \\ 0 & 0 & 1 & \cdot & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (5.45)$$

В таком случае система уравнений МКЭ для базиса сколь угодно высокого порядка записывается особенно просто, если в (5.13) $b(x) = c(x) = 0$. Пусть на каждом сеточном отрезке $[x_i, x_{i+1}]$ используется $p_i + 1$ базисная функция вида (5.41), а приближенное решение записывается в виде

$$u^h(x) = v_i \varphi_1^{(i)}(x) + v_{i+1} \varphi_2^{(i)}(x) + \sum_{l=3}^{p_i+1} v_l^{(i)} \varphi_l^{(i)}(x).$$

В силу ортогональных свойств (5.44) производных от «внутренних» базисных функций $\varphi_l^{(i)}(x)$, $l \geq 3$, соответствующие неизвестные $v_l^{(i)}$ определяются независимо от остальных из соотношений

$$v_l^{(i)} = (f, \varphi_l^{(i)}) / a(\varphi_l^{(i)}, \varphi_l^{(i)}).$$

Каждая же пара узловых функций $\varphi_2^{(i-1)}(x)$, $\varphi_1^{(i)}(x)$ образует одну базисную функцию $\varphi_i(x)$ из (5.16) для кусочно-линейной интерполяции, и в итоге для неизвестных v_i получается трехточечная система уравнений вида (5.28), с учетом $b_{i\pm 1/2} = c_{i\pm 1/2} = 0$.

Если же функция $d(x)$ не постоянна, то локальная матрица жесткости $\bar{A}_{s,i}^h$ имеет уже не простой вид (5.45), а является плотной. Но поскольку «внутренние» базисные функции $\varphi_l^{(i)}(x)$ отличны от нуля только на одном сеточном интервале, как и в лагранжевых элементах высших порядков, то получаемые алгебраические системы имеют схожую алгебраическую структуру. К ним, в частности, также может быть эффективно применена процедура конденсации, т. е. исключения «внутренних» неизвестных $v_l^{(i)}$ и получение трехточечной системы для узловых переменных v_i .

5.3. Двумерные конечные элементы

Любую двумерную полигональную расчетную область Ω можно разбить бесчисленным количеством способов на достаточно малые многоугольные конечные элементы, совокупность которых будем обозначать через Ω^h . Если Ω представима совокупностью прямоугольников со сторонами, параллельными осям декартовой системы координат, то удобно брать прямоугольные конечные элементы. Однако более универсальным подходом является *триангуляция* области, т. е. разбиение ее на непересекающиеся треугольники. Способ триангуляции и характеристики получаемых треугольников могут существенно влиять на точность последующей аппроксимации, на обусловленность получаемой алгебраической системы и на вычислительную сложность алгоритма в целом.

По этим вопросам имеется огромное количество литературы, см., например, монографии [110]–[113] и цитируемые там работы. Этим проблемам специально посвящается следующая глава книги. Здесь мы ограничимся только определением «хорошей» триангуляции, так называемой *триангуляции Делоне*.

Определение 5.2. Разбиение полигональной области на треугольники называется триангуляцией Делоне, если любая окружность, описанная около какого-либо треугольника, не содержит внутри себя вершины никакого другого треугольника (расположение вершины на самой окружности допускается).

Если граница области криволинейная, то построение аппроксимаций в получаемых криволинейных элементах усложняется, и мы здесь рассмотрим отдельные простейшие подходы.

5.3.1. Конечно-элементные аппроксимации на треугольниках

Рассмотрим сначала треугольники, образованные делением пополам (параллельными друг другу диагоналями, для определенности изображенным на рис. 5.5) прямоугольных ячеек неравномерной *прямоугольной сетки*

$$\begin{aligned}x &= x_i, & i &= 0, 1, \dots, I+1, \\y &= y_j, & j &= 0, 1, \dots, J+1,\end{aligned}$$

с шагами $h_i^x = x_{i+1} - x_i$, $h_j^y = y_{j+1} - y_j$.

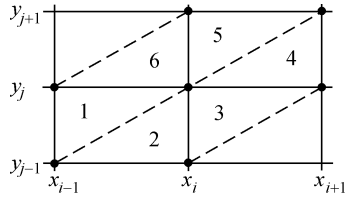


Рис. 5.5. Триангуляция прямоугольной сетки

Каждому узлу сетки с индексами i, j сопоставим финитную кусочно-линейную базисную функцию $\varphi_{i,j}(x, y)$, носителем которой являются шесть треугольников Δ_k с общей вершиной (x_i, y_j) , пронумерованные на рис. 4.5. Будучи линейной по x и y на каждом треугольнике, равная единице

в (i, j) -м узле и нулю в остальных, такая функция имеет форму шестигранной пирамиды. Она определяется однозначно и может быть записана в форме

$$\varphi_{i,j}(x, y) = 1 + a_{i,j}^{(k)}(x - x_i) + b_{i,j}^{(k)}(y - y_j), \quad x, y \in \Delta_k, \quad (5.46)$$

где коэффициенты $a_{i,j}^{(k)}$, $b_{i,j}^{(k)}$ представляют собой значения производных по x, y в разных треугольниках Δ_k и имеют вид, приведенный в табл. 5.1.

Таблица 5.1. Коэффициенты кусочно-линейной базисной функции

k	1	2	3	4	5	6
$a_{i,j}$	$1/h_{i-1}^x$	0	$1/h_i^x$	$1/h_i^x$	0	$1/h_{i+1}^x$
$b_{i,j}$	0	$1/h_{j-1}^y$	$1/h_{j-1}^y$	0	$1/h_j^y$	$1/h_{j+1}^y$

Очевидно, что на данной прямоугольной сетке по значениям $v_{i,j} = v(x_i, y_j)$ однозначно определяется кусочно-линейная лагранжева интерполяция

$$v^h(x, y) = \sum_{i,j} v_{i,j} \varphi_{i,j}(x, y). \quad (5.47)$$

Иногда базисные функции $\varphi_{i,j}$ называются *функциями Куранта*, а представление (5.47) — *кусочно-линейным восполнением* функции $v(x, y)$. Очевидно, что если приближаемая функция принадлежит классу C^2 , то погрешность интерполяционной формулы (5.47) есть величина порядка $O(h^2)$.

Рассмотрим для конкретности применение описанных базисных функций к построению матрицы жесткости метода Рунта для вариационной задачи, соответствующей однородной первой краевой задаче для уравнения диффузии в прямоугольнике $[a, b] \times [c, d]$:

$$\begin{aligned} -\frac{\partial}{\partial x} p \frac{\partial u}{\partial x} - \frac{\partial}{\partial y} p \frac{\partial u}{\partial y} &= f(x, y), & a = x_0 < x < x_{I+1} = b, \\ c = y_0 < y < y_{J+1} = d, & u|_{\Gamma} = 0. \end{aligned} \quad (5.48)$$

Билинейная и линейная формы в данном случае записываются через двойные интегралы

$$\begin{aligned} a(\varphi_{i,j}, \varphi_{i',j'}) &= \int_a^b \int_c^d p(x, y) (\varphi_{i,j}^x \varphi_{i',j'}^x + \varphi_{i,j}^y \varphi_{i',j'}^y) dx dy, \\ f_{i,j}^h &= (f, \varphi_{i,j}) = \int_a^b \int_c^d f \varphi_{i,j} dx dy, \end{aligned} \quad (5.49)$$

которые естественным образом разбиваются на сумму элементарных интегралов по треугольникам. При этом в силу свойств функций $\varphi_{i,j}$ имеем $a(\varphi_{i,j}, \varphi_{i',j'}) = 0$ при $|i - i'| + |j - j'| > 1$.

Поскольку производные базисных функций на каждом треугольнике постоянные, то при $p(x, y) = 1$ матрица жесткости вычисляется просто, и система алгебраических уравнений принимает (с учетом условий $v_{i,j} = 0$ на границе) пятиточечную форму

$$\begin{aligned} & -\frac{h_j^y + h_{j-1}^y}{2h_{i-1}^x} v_{i-1,j} - \frac{h_i^x + h_{i-1}^x}{2h_{j-1}^y} v_{i,j-1} - \frac{h_j^y + h_{j-1}^y}{2h_i^x} v_{i+1,j} - \frac{h_i^x + h_{i-1}^x}{2h_j^y} v_{i,j+1} + \\ & \frac{(h_i^x + h_{i-1}^x)(h_j^y + h_{j-1}^y)}{2} \left(\frac{1}{h_i^x h_{i-1}^x} + \frac{1}{h_j^y h_{j-1}^y} \right) v_{i,j} = f_{i,j}^h, \end{aligned} \quad (5.50)$$

$$i = 1, \dots, I, \quad j = 1, 2, \dots, J,$$

которая имеет стандартный для конечно-разностных аппроксимаций вид после деления на объем сеточной ячейки

$$V_{i,j} = \frac{(h_i^x + h_{i-1}^x)(h_j^y + h_{j-1}^y)}{4}$$

(при этом, правда, в случае неравномерной сетки теряется симметрия конечно-элементной системы).

Если в задаче (5.48) однородные условия Дирихле заменить на неоднородные, т. е. $u|_{\Gamma} = g(x, y)$, то приближенное решение ищется в виде

$$v^h(x, y) = \sum_{i,j \in \Omega} v_{i,j} \varphi_{i,j}(x, y) + \sum_{x,y \in \Gamma} g_{i,j} \varphi_{i,j}(x, y),$$

где первая сумма берется только по внутренним узлам сетки. При этом система уравнений МКЭ формально имеет тот же вид, только теперь в левой части (5.50) для членов, соответствующих граничным узлам, надо заменить $v_{i,j}$ на $g_{i,j}$ и перенести их в правую часть.

Перейдем теперь к описанию конечных элементов на треугольниках более общего вида, которые изображены на рис. 5.6.

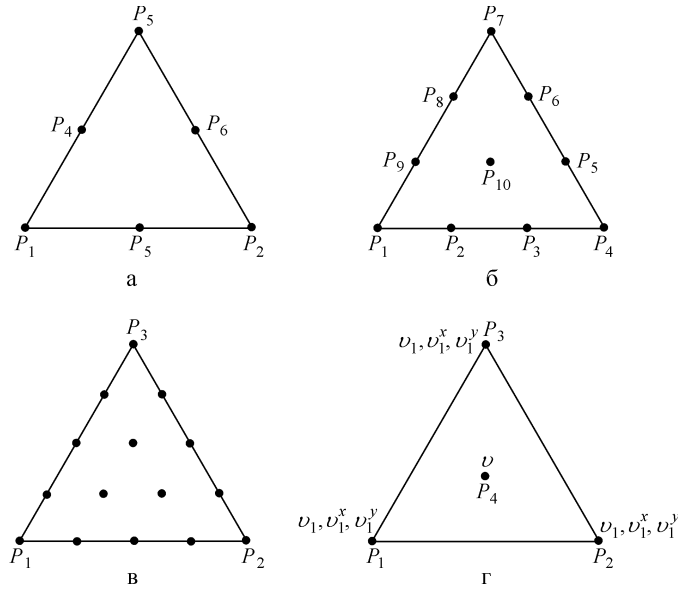


Рис. 5.6. Треугольные конечные элементы

Здесь произвольная линейная интерполяция Лагранжа может быть записана в виде

$$v^h(x, y) = \alpha_1 + \alpha_2 x + \alpha_3 y, \quad (5.51)$$

где три коэффициента α_k однозначно определяются по значениям приближаемой функции в вершинах треугольника:

$$\alpha_1 + \alpha_2 x_l + \alpha_3 y_l = v(x_l, y_l), \quad l = 1, 2, 3,$$

если только он не вырождается в отрезок. Подпространство же линейных функций на треугольнике может быть представлено как линейная оболочка трех базисных функций вида (5.51), все девять коэффициентов которых находятся из интерполяционных условий

$$\varphi_k(x_l, y_l) = \alpha_1^{(k)} + \alpha_2^{(k)} x_l + \alpha_3^{(k)} y_l = \delta_{k,l}, \quad k, l = 1, 2, 3.$$

Лангражевые интерполяции более высоких порядков на треугольнике строятся относительно легко путем добавления новых узловых точек на сторонах и внутри треугольника. Так, при использовании трех срединных точек сторон (кроме вершин, см. рис. 5.6, а) однозначно определяется шесть квадратичных базисных функций:

$$\varphi_k(x, y) = \alpha_1^{(k)} + \alpha_2^{(k)} x_l + \alpha_3^{(k)} y_l + \alpha_4^{(k)} x_l^2 + \alpha_5^{(k)} x_l y_l + \alpha_6^{(k)} y_l^2 = \delta_{k,l}, \quad k, l = 1, 2, \dots, 6. \quad (5.52)$$

Алгебраическая задача нахождения базисных функций (5.52) может быть сформулирована следующим образом. Обозначим через

$$\bar{\alpha}^{(k)} = (\alpha_1^{(k)}, \dots, \alpha_6^{(k)})^t$$

вектор-столбец из шести коэффициентов функции $\varphi_k(x, y)$ и через e_k — вектор-орд с единственной ненулевой k -й единичной компонентой. Тогда мы можем записать

$$S \bar{\alpha}^{(k)} = \begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & x_1 y_1 & y_1^2 \\ 1 & x_2 & y_2 & x_2^2 & x_2 y_2 & y_2^2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_6 & y_6 & x_6^2 & x_6 y_6 & y_6^2 \end{bmatrix} \begin{bmatrix} \alpha_1^{(k)} \\ \alpha_2^{(k)} \\ \vdots \\ \alpha_6^{(k)} \end{bmatrix} = e_k,$$

т. е. вектор $\bar{\alpha}^{(k)}$ является k -м столбцом матрицы, обратной к S . Таким образом, задача нахождения всех 36 коэффициентов базисных функций сводится к вычислению матрицы S^{-1} .

На рис. 5.6, б, в приведены примеры использования дополнительных точек для построения лагранжевых интерполяций, имеющих, соответственно, третий и четвертый порядок по каждой переменной. Подчеркнем, что здесь под порядком p подразумевается максимальная сумма степеней при x и y в одном слагаемом. Общее правило очень простое: каждая пара сторон треугольника разбивается на равные (для простоты) части с помощью отрезков, параллельных третьей стороне, а получаемые точки берутся в качестве интерполяционных узлов. Их количество d , равное числу коэффициентов всех базисных функций треугольника, для различных порядков n определяется приводимой ниже таблицей.

Таблица 5.2. Число степеней свободы лагранжевых интерполяций на треугольнике

n	1	2	3	p
d	3	6	10	$(p+1)(p+2)/2$

Все базисные функции такого типа делятся на три группы: узловые, реберные и внутренние. Их число, соответственно, равно 3, $3(p-1)$ и $(p-1)(p-2)/2$.

Отметим следующие качества финитных базисных функций, важные с точки зрения структуры глобальных матриц МКЭ: внутренние функции имеют носителем только «свой» треугольник, носителем реберных функций являются два соответствующих смежных треугольника, а носитель узловой функции — это все треугольники с данной общей вершиной.

Укажем еще на два свойства кусочно-лагранжевых интерполяций, касающиеся их гладкости. Во-первых, данные аппроксимации непрерывны при переходе из одного треугольника в другой. Это объясняется тем, что на их смежной стороне оба представления суть одномерные многочлены Лагранжа p -го порядка, определяемые по одним и тем же $(p+1)$ точкам. Во-вторых, как бы мы не увеличивали порядок p , при переходе из одного треугольника в соседний даже первые производные терпят скачок.

От указанного недостатка можно частично избавиться путем использования двумерной эрмитовой интерполяции. Таким относительно простым примером является кубическая интерполяция

$$v^h(x, y) = \alpha_1 + \alpha_2 x + \alpha_3 y + \alpha_4 x^2 + \alpha_5 xy + \alpha_6 y^2 + \alpha_7 x^3 + \alpha_8 x^2 y + \alpha_9 xy^2 + \alpha_{10} y^3, \quad (5.53)$$

в которой максимальная суммарная степень $x^k y^l$ равна трем, а десять коэффициентов однозначно находятся по следующим десяти задаваемым величинам:

$$v_1, v_1^x, v_1^y, v_2, v_2^x, v_2^y, v_3, v_3^x, v_3^y, v_4,$$

где $v_4 = v(x_4, y_4)$ есть значение функции в центре тяжести треугольника, схематически изображенного на рис. 5.6, г. Здесь отмеченные кружочками вершины являются «тройными» точками, в которых заданы не только значения функции, но и ее градиенты (в данном случае — значениями производных по x и y). Подчеркнем, что точка P_4 должна быть именно центром тяжести, а никакой другой.

Таким образом, технически проблема в данном случае сводится к определению десяти коэффициентов каждой из десяти бикубической базисной функции вида

$$\varphi_k(x, y) = \alpha_1^{(k)} + \alpha_2^{(k)} x + \alpha_3^{(k)} y + \alpha_4^{(k)} x^2 + \alpha_5^{(k)} xy + \alpha_6^{(k)} y^2 + \alpha_7^{(k)} x^3 + \alpha_8^{(k)} x^2 y + \alpha_9^{(k)} xy^2 + \alpha_{10}^{(k)} y^3,$$

а также к последующему вычислению 55 элементов симметричной матрицы жесткости десятого порядка. Получаемая при этом аппроксимирующая функция $v^h(x, y)$ обладает непрерывными производными в узлах сетки, но при переходе из одного треугольника в другой во внутренних точках их общей стороны величины v^x, v^y терпят разрыв.

Коэффициенты $\alpha_l^{(k)}$, $k, l = 1, \dots, 10$, кубических базисных функций являются, очевидно, элементами матрицы десятого порядка, обратной к матрице

$$S = \begin{bmatrix} 1 & x_1 & y_1 & x_1^2 & x_1 y_1 & y_1^2 & x_1^3 & x_1^2 y_1 & x_1 y_1^2 & y_1^3 \\ 1 & x_2 & y_2 & x_2^2 & x_2 y_2 & y_2^2 & x_2^3 & x_2^2 y_2 & x_2 y_2^2 & y_2^3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & x_{10} & y_{10} & x_{10}^2 & x_{10} y_{10} & y_{10}^2 & x_{10}^3 & x_{10}^2 y_{10} & x_{10} y_{10}^2 & y_{10}^3 \end{bmatrix}.$$

Одним из распространенных приемов для вычисления матриц МКЭ является использование *треугольных*, или *барицентрических*, координат. Пусть в треугольнике $P_1 P_2 P_3$ через l_k ,

$k = 1, 2, 3$, обозначается сторона, противоположная k -й вершине, а P_0 обозначает некоторую внутреннюю точку треугольника. Обозначим через S_k , $k = 1, 2, 3$, площадь треугольника с вершиной P_0 и основанием l_k , а через $S = S_1 + S_2 + S_3$ — площадь всего треугольника $P_1P_2P_3$. Тогда треугольными координатами L_1, L_2, L_3 точки P_0 называются величины

$$L_k = S_k/S, \quad k = 1, 2, 3.$$

Очевидно, что треугольными координатами вершины $P_k(x_k, y_k)$ являются $L_k = 1$ и $L_{k'} = 0$ при $k' \neq k$. Отсюда следует, что связь треугольных координат с декартовыми описывается соотношениями

$$x = L_1x_1 + L_2x_2 + L_3x_3, \quad y = L_1y_1 + L_2y_2 + L_3y_3.$$

С учетом определения L_k имеем равенство $L_1 + L_2 + L_3 = 1$, что в совокупности нам дает векторное уравнение

$$\begin{bmatrix} 1 \\ x \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{bmatrix} \begin{bmatrix} L_1 \\ L_2 \\ L_3 \end{bmatrix}.$$

Обозначая через $b_{k,j}$ элементы обратной матрицы данной системы:

$$b_{k,1} = x_{k+1}y_{k+2} - x_{k+2}y_{k+1}, \quad b_{k,2} = y_{k+1} - y_{k+2}, \quad b_{k,3} = x_{k+2} - x_{k+1},$$

где $k = 1, 2, 3$ и $x_{k'} = x_{k'-3}$, $y_{k'} = y_{k'-3}$ при $k' > 3$, мы можем записать

$$L_k = (b_{k1} + b_{k2}x + b_{k3}y) \frac{1}{2S},$$

$$S = \frac{1}{2}(x_2y_3 + x_3y_1 + x_1y_2 - x_2y_1 - x_3y_2 - x_1y_3).$$

Полный многочлен n -го порядка $P_n(x, y)$ может быть представлен с помощью треугольных координат в виде

$$P_n = \sum_{i=1}^n a_i L_1^p L_2^q L_3^r, \quad p + q + r = n,$$

где в сумму входят все однородные члены степени n . Например, для квадратного разложения имеем

$$P_2 = L_1^2 a_1 + L_2^2 a_2 + L_3^2 a_3 + L_1 L_2 a_4 + L_1 L_3 a_5 + L_2 L_3 a_6.$$

Замечательным свойством треугольных координат является существование явных выражений для интегралов от их степеней по площади произвольного треугольника:

$$\int_S L_1^p L_2^q L_3^r ds = 2S \frac{p!q!r!}{(p+q+r+2)!}.$$

С помощью треугольных координат на треугольном конечном элементе несложно строится иерархический базис И. Бабушки, являющийся обобщением рассмотренного выше базиса его имени для одномерного случая.

Для единичного равностороннего треугольника в локальных координатах ξ , η , изображенных на рис. 5.7, треугольные (барицентрические) координаты описываются формулами

$$L_1 = \frac{1}{2} \left(1 - \xi - \frac{\eta}{\sqrt{3}} \right), \quad L_2 = \frac{1}{2} \left(1 + \xi - \frac{\eta}{\sqrt{3}} \right), \quad L_3 = \frac{\eta}{\sqrt{3}}.$$

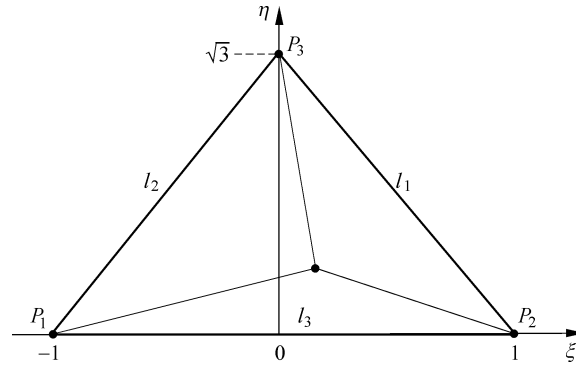


Рис. 5.7. Стандартный треугольный элемент

Иерархический базис порядка p включает три типа базисных функций:

- Три *узловые базисные функции*

$$\varphi_k = L_k, \quad k = 1, 2, 3.$$

- $3(p-1)$ *реберные базисные функции* такие, что $p-1$ функций, соответствующих какой-то стороне треугольника, обращаются в нуль на двух других сторонах; например, 1-й стороне, лежащей напротив вершины P_1 , соответствуют базисные функции

$$\varphi_k^{(1)} = L_2 L_3 \chi_k(2L_3 - 1), \quad k = 2, \dots, p,$$

где функция $\chi_k(z)$ определяются с помощью равенства

$$\psi_k(z) = \frac{1}{4}(1 - z^2)\chi_k(z),$$

а функция $\psi_k(z)$ выражается через полином Лежандра с помощью формулы (4.49) (например, $\chi_2 = -\sqrt{6}$, $\chi_3(z) = -\sqrt{10}z$); базисные функции $\varphi_k^{(2)}$, $\varphi_k^{(3)}$, $k = 2, \dots, p$, определяются аналогично:

$$\begin{aligned} \varphi_k^{(2)} &= L_1 L_3 \chi(2L_3 - 1), \\ \varphi_k^{(3)} &= L_1 L_2 \chi(L_2 - L_1); \end{aligned}$$

поясним, что здесь аргументы функции χ подобраны так, что они меняются от -1 до $+1$ при движении от одной вершины соответствующей стороны треугольника к другой, обращаясь в нуль на середине стороны (например, величина $L_2 - L_1 = \xi$ равна -1 в точке P_1 и $+1$ — в вершине P_2).

- $(p-1)(p-2)/2$ внутренние базисные функции $\varphi_k^{(0)}$, первая из которых имеет вид

$$\varphi_1^{(0)} = L_1 L_2 L_3;$$

остальные внутренние функции образуются умножением $\varphi_1^{(0)}$ на многочлены Лежандра $P_n(z)$ или их парные произведения, причем суммарный порядок этих многочленов не должен превышать $p-3$:

$$\begin{aligned} \varphi_m^{(0)} &= L_1 L_2 L_3 P_i(L_2 - L_1) P_j(2L_3 - 1), \\ i, j &= 0, 1, \dots, p-3, \quad i+j = 0, 1, \dots, p-3, \end{aligned}$$

где $m = m(i, j)$; например, для $p = 8$ общее количество базисных функций равно 45, из которых 3 — узловые, 21 — реберная и 21 — внутренняя.

5.3.2. Прямоугольные конечные элементы

Очевидным достоинством прямоугольной сетки является то, что на ней можно легко конструировать двумерные интерполяции —

лагранжевые или эрмитовые — как тензорное произведение одномерных, что дает нам семейство *биполиномиальных аппроксимаций* различных порядков p . Подчеркнем, что под степенью p здесь понимается максимальный порядок x или y .

Кусочно-линейные базисные функции. Если $\varphi_i(x)$ — одномерная кусочно-линейная базисная функция вида (5.16), то финитная *билинейная функция* с носителем $(x_{i-1}, x_{i+1}) \times (y_{j-1}, y_{j+1})$, соответствующая узлу сетки (x_i, y_j) , определяется как

$$\varphi_{i,j}(x, y) = \varphi_i(x)\varphi_j(y), \quad (5.54)$$

где $\varphi_j(y)$ находится аналогично $\varphi_i(x)$ из (5.16):

$$\varphi_j(y) = \begin{cases} (y_{j+1} - y)/h_j^y, & y \in (y_j, y_{j+1}), \\ (y - y_{j-1})/h_{j-1}^y, & y \in (y_{j-1}, y_j), \\ 0, & y \notin (y_{j-1}, y_{j+1}). \end{cases}$$

В (i, j) -м элементе $\Omega_{i,j}^h = \{x_i < x < x_{i+1}, y_j < y < y_{j+1}\}$ отличны от нуля четыре базисных функции, которые можно обозначить как

$$\begin{aligned} \varphi_{i,j}^{(1)} &= \varphi_i(x)\varphi_j(y), & \varphi_{i,j}^{(2)} &= \varphi_{i+1}(x)\varphi_j(y), \\ \varphi_{i,j}^{(3)} &= \varphi_{i+1}(x)\varphi_{j+1}(y), & \varphi_{i,j}^{(4)} &= \varphi_i(x)\varphi_{j+1}(y). \end{aligned}$$

При этом приближенное решение на данном элементе представляется в виде

$$v_{i,j}^h(x, y) = \sum_{k=1}^4 v_k^{(i,j)} \varphi_{i,j}^{(k)},$$

где $v_1^{(i,j)} = v(x_i, y_j)$, $v_2^{(i,j)} = v(x_{i+1}, y_j)$, $v_3^{(i,j)} = v(x_{i+1}, y_{j+1})$, $v_4^{(i,j)} = v(x_i, y_{j+1})$ суть значения приближенного решения в соответствующих узлах сетки.

Локальная матрица жесткости, соответствующая краевой задаче (5.48), в этом случае имеет четвертый порядок и записывается в форме

$$\bar{A}_s^{(i,j)} = \left\{ \bar{a}_{k,l}^{(i,j)} = \int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} p(x, y) \left(\frac{\partial \varphi_{i,j}^{(k)}}{\partial x} \frac{\partial \varphi_{i,j}^{(l)}}{\partial x} + \frac{\partial \varphi_{i,j}^{(k)}}{\partial y} \frac{\partial \varphi_{i,j}^{(l)}}{\partial y} \right) dx dy, \right.$$

$$k, l = 1, \dots, 4 \}.$$

Лагранжевые базисы высших порядков. На прямоугольнике $[x_i, x_{i+1}] \times [y_j, y_{j+1}]$ может быть также определена биквадратичная аппроксимация

$$v_{i,j}^h(x, y) = \sum_{k,l=0}^2 a_{k,l}^{(i,j)} x^k y^l, \quad (5.55)$$

девять коэффициентов которой однозначно определяются из условий интерполяции по значениям в девяти точках, изображенных на рис. 5.8, а (помимо узлов, здесь используются срединные точки с координатами $x_{i+1/2}, y_{j+1/2}$).

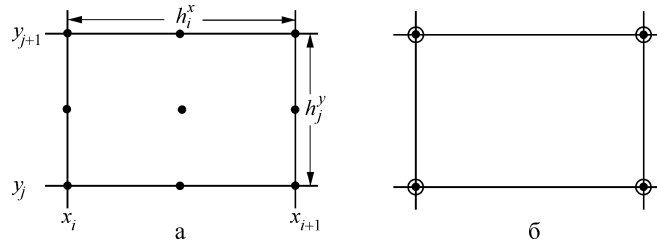


Рис. 5.8. Прямоугольные конечные элементы

Аналогично могут быть построены (с использованием дополнительных реберных и внутренних точек) биполиномиальные аппроксимации порядка p по каждой из переменных:

$$v_{i,j}^h(x, y) = \sum_{k,l=0}^p \alpha_{k,l} x^k y^l. \quad (5.56)$$

При этом проще всего разбить параллельными линиями исходный сеточный прямоугольник на p^2 равных частей, а коэффициенты $a_{k,l}^{(i,j)}$ находить из соответствующих условий двумерных интерполяции. По этим же условиям строятся $(p+1)^2$ базисных функций, отличных от нуля на данном прямоугольнике, каждая из которых равна единице в одной из $(p+1)^2$ точек и нулю — в остальных.

Отметим, что в данном случае двумерные базисные функции выражаются явно через произведения одномерных базисных многочленов Лагранжа:

$$\varphi_{i,j}^{(k,l)} = \Psi_k(x)\Psi_l(y), \quad k, l = 1, \dots, p+1,$$

где $\Psi_k(x)$, $\Psi_l(y)$ определяются из (5.31) при замене x_0 , x_n на x_i , x_{i+1} или y_j , y_{j+1} соответственно.

Соответственно расположению своих максимумов эти базисные функции делятся на четыре узловые, $4(p-1)$ реберные и $(p-1)^2$ внутренние функции. Порядок локальных матриц МКЭ равен, естественно, $(p+1)^2$. Структура глобальных матриц МКЭ формируется с учетом того, что носителем внутренних базисных функций является только «свой» сеточный прямоугольник, носителем реберных функций — два смежных прямоугольника, а узловая базисная функция отлична от нуля в четырех прямоугольниках, имеющих (i, j) -й узел своей общей вершиной.

Эрмитовый кубический базис. Как и в треугольных конечных элементах, кусочно-лагранжевые биполиномиальные интерполяции вида (5.56) являются непрерывными в целом в сеточной области Ω^h , но не более того. То есть аппроксимации в двух смежных прямоугольниках на смежном ребре сводятся к идентичным полиномам одной переменной, но при переходе из одного прямоугольника в другой производные, вообще говоря, терпят разрыв. Этого недостатка легко избежать с помощью введения биполиномиальных эрмитовых аппроксимаций, из которых простейшей является *бикубическая эрмитовая интерполяция*, использующая кратные узлы в вершинах прямоугольника со значениями v , v^x , v^y , v^{xy} (см. рис. 5.8, б, а также [82]). Такое приближение описывается формулой (5.56) при $p = 3$, а ее коэффициенты определяются из вида кубического эрмитового многочлена (5.35). Общее количество ненулевых базисных функций на одном прямоугольнике и порядок локальных матриц МКЭ в данном случае равен 16.

Здесь для каждого (i, j) -го узла можно сопоставить четыре базисные функции, являющиеся произведениями одномерных базисных функций $\varphi_i(x) = \varphi(\frac{x-x_i}{h_i^x})$, $\psi_j(y) = \psi(\frac{y-y_j}{h_j^y})$ из (5.35):

$$\begin{aligned} \Phi_1 &= \varphi_i(x)\varphi_j(y), & \Phi_2 &= \varphi_i\psi_j(y), \\ \Phi_3 &= \psi_i(x)\varphi_j(y), & \Phi_4 &= \psi_i(x)\psi_j(y). \end{aligned}$$

Отметим, что в отличие от эрмитовой кубической интерполяции (5.53) на треугольной сетке бикубическая эрмитовая интерполяция (5.56) при $p = 3$ обеспечивает непрерывность произ-

водных первого порядка (более того, оказывается непрерывной даже смешанная производная $v^{xy!}$).

Иерархические базисы. Аналогично одномерному случаю на прямоугольной сетке легко строится иерархический полиномиальный базис И. Бабушки, который мы выпишем для стандартного квадратного элемента

$$-1 \leq \xi \leq 1, \quad -1 \leq \eta \leq 1,$$

получаемого из произвольного прямоугольника заменой переменных вида (5.39).

В этом случае конструируется следующий набор финитных базисных функций:

- Четыре узловые функции

$$\begin{aligned} \varphi_1(\xi, \eta) &= \frac{1}{4}(1 - \xi)(1 - \eta), \\ \varphi_2(\xi, \eta) &= \frac{1}{4}(1 + \xi)(1 - \eta), \\ \varphi_3(\xi, \eta) &= \frac{1}{4}(1 + \xi)(1 + \eta), \\ \varphi_4(\xi, \eta) &= \frac{1}{4}(1 - \xi)(1 + \eta), \end{aligned} \tag{5.57}$$

каждая из которых равна единице в одном из узлов сетки и нулю в остальных.

- $4(p - 1)$ реберные функции, по $p - 1$ на каждой стороне элемента ($p \geq 2$):

$$\begin{aligned} \varphi_k^{(1)} &= \frac{1}{2}(1 - \eta)\psi_k(\xi), \\ \varphi_k^{(2)} &= \frac{1}{2}(1 + \xi)\psi_k(\eta), \\ \varphi_k^{(3)} &= \frac{1}{2}(1 + \eta)\psi_k(\xi), \\ \varphi_k^{(4)} &= \frac{1}{2}(1 - \xi)\psi_k(\eta), \quad k = 2, \dots, p, \end{aligned} \tag{5.58}$$

где функции $\psi_k(\xi)$ определяются через многочлены Лежандра по формуле (4.49).

- Внутренние функции

$$\varphi_k^{(0)} = \psi_i(\xi)\psi_j(\eta); \quad i, j = 2, \dots, p,$$

у которых индекс k выражается через i, j по какой-либо принятой упорядоченности, а общее количество равно $(p - 1)^2$.

Число всех базисных функций, не обращающихся тождественно в нуль на одном элементе, равно $4 + 4(p-1) + (p-1)^2 = (p+1)^2$. Иерархичность данного базиса заключается в том, что при повышении его порядка на единицу (переход от p к $p+1$) добавляется только $2(p+2)$ новых базисных функций, а функции до p -го порядка включительно остаются неизменными.

О компактных схемах МКЭ. Построение лагранжевых и эрмитовых базисов высших порядков призвано, естественно, повысить точность конечно-элементных аппроксимаций. Однако этому сопутствует значительное увеличение числа степеней свободы и количества ненулевых элементов глобальной матрицы МКЭ. В свою очередь, это приводит к росту трудоемкости процесса решения СЛАУ, который является наиболее критичным по ресурсоемкости в реализации больших практических задач.

С этой точки зрения значительный интерес представляют хорошо известные в теории разностных схем компактные аппроксимации, см. [81], [178], обеспечивающие при небольшом числе используемых узлов сеточного шаблона высокую точность численного решения. Кроме того, получаемые с их помощью СЛАУ обладают хорошими алгебраическими свойствами и допускают применение быстрых итерационных алгоритмов.

В качестве простого примера мы приведем знаменитую конечно-разностную схему Ш.Е. Микеладзе

$$\begin{aligned}
 -\Delta_h^{(9)} v_{i,j} &\equiv \frac{5(h_x^2 + h_y^2)}{3h_x^2 h_y^2} v_{i,j} - \frac{1}{6h_x^2} \left(5 - \frac{h_x^2}{h_y^2}\right) (v_{i-1,j} + v_{i+1,j}) - \\
 &\quad \frac{1}{6h_y^2} \left(5 - \frac{h_y^2}{h_x^2}\right) (v_{i,j-1} + v_{i,j+1}) - \\
 &\quad \frac{h_x^2 + h_y^2}{12h_x^2 h_y^2} (v_{i-1,j-1} + v_{i-1,j+1} + v_{i+1,j-1} + v_{i+1,j+1}) \\
 &= f_{i,j} + \frac{h_x^2}{12} \frac{\partial^2 f_{i,j}}{\partial x^2} + \frac{h_y^2}{12} \frac{\partial^2 f_{i,j}}{\partial y^2}, \tag{5.59}
 \end{aligned}$$

которая аппроксимирует с погрешностью $O(h^4)$ уравнение Пуассона на девятиточечном шаблоне типа «ящик» равномерной прямоугольной сетки, изображенном на рис. 5.9.

Порядок точности этого уравнения остается четвертым, если вторые производные в правой части (5.59) заменить на конечные разности второго порядка.

Представляет интерес вопрос, можно ли получить данную компактную схему с помощью технологии конечных элементов. Ответ на данный вопрос является положительным, а результаты исследований в кратком изложении выглядят следующим образом.

Рассмотрим для простоты однородную задачу Дирихле в прямоугольной области Ω , замыкание которой $\bar{\Omega}$ есть объединение ячеек $\Omega_{i,j}^h = [x_i, x_{i+1}] \times [y_j, y_{j+1}]$, $i = 0, 1, \dots, I$, $j = 0, 1, \dots, J$. Соответствующая обобщенная постановка имеет следующий вид: требуется найти функцию $u \in H_0^1(\Omega)$ такую, что $\forall v \in H_0^1(\Omega)$

$$a(u, v) \equiv \int_{\Omega} \nabla u \nabla v \, dx \, dy = \int_{\Omega} f v \, dx \, dy \equiv f(v), \quad (5.60)$$

где $f(v)$ есть линейный ограниченный в $H_0^1(\Omega)$ функционал, т. е.

$$|f(v)| \leq c|v|_{H^1(\Omega)} = ca^{1/2}(v, v),$$

а число $c > 0$ не зависит от функции v .

Преобразуя произвольную ячейку $\Omega_{i,j}^h$ в единичный квадрат $e_0 = [0, 1]^2$:

$$\xi = (x - x_i)/h_i^x, \quad \eta = (y - y_j)/h_j^y, \quad (x, y) \in \Omega^h,$$

введем в e_0 билинейный базис $\{\bar{\varphi}_{k,l}\}$, $k, l = 1, 2$,

$$\begin{aligned} \bar{\varphi}_{1,1}(\xi, \eta) &= (1 - \xi)(1 - \eta), & \bar{\varphi}_{1,2}(\xi, \eta) &= \xi(1 - \eta), \\ \bar{\varphi}_{1,2}(\xi, \eta) &= (1 - \xi)\eta, & \bar{\varphi}_{2,2}(\xi, \eta) &= \xi\eta, \end{aligned}$$

а также дополнительную внутреннюю «биквадратичную» функцию

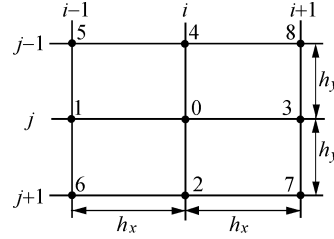


Рис. 5.9. Сеточный шаблон для девятиточечного уравнения типа «ящик»

$$\psi(\xi, \eta) = \xi(1 - \xi)\eta(1 - \eta),$$

равную нулю на границе ячейки ∂e_0 .

Определим теперь в e_0 модифицированный параметрический базис:

$$\varphi_{k,l}(\xi, \eta) = \bar{\varphi}_{k,l}(\xi, \eta) + (-1)^{k+l} \sqrt{\omega} \psi(\xi, \eta), \quad k, l = 1, 2, \quad (5.61)$$

где $\omega > 0$ — некоторый числовой параметр. Новые функции образуют линейно независимую систему в e_0 и обладают свойствами

$$\sum_{k,l=1}^2 \varphi_{k,l}(\xi, \eta) = 1,$$

а в вершинах e_0 принимают те же значения, что и $\bar{\varphi}_{k,l}(\xi, \eta)$. По элементарным функциям стандартным образом построим систему линейно независимых финитных функций $\psi_{i,j}(x, y)$, заданных всюду в $\bar{\Omega}$. Такая система образует базис пространства

$$V_{\omega}^h = \text{span} \{ \psi_{i,j}(x, y) \}, \quad (5.62)$$

где $i = 1, \dots, I$, $j = 1, \dots, J$, а каждая из функций $\psi_{i,j}(x, y)$ (индекс ω по контексту опускаем) отлична от нуля при $(x, y) \in [x_{i-1}, x_{i+1}] \times [y_{j-1}, y_{j+1}]$. Отметим также, что $V_{\omega}^h \subset H_0^1(\Omega)$ и $V_{\omega}^h \subset C(\bar{\Omega})$, а V_0^h есть пространство с билинейным базисом.

Введенный базис для ячейки $\Omega_{i,j}^h$ определяет локальную матрицу жесткости четвертого порядка

$$\begin{aligned} A_{\omega}^{(i,j)} &= \left\{ \text{mes}(\Omega_{i,j}^h) \int_{e_0} \left[(h_i^x)^{-2} \frac{\partial \psi_{k,l}}{\partial \xi} \frac{\partial \psi_{m,n}}{\partial \xi} + (h_j^y)^{-2} \frac{\partial \psi_{k,l}}{\partial \eta} \frac{\partial \psi_{m,n}}{\partial \eta} \right] d\xi d\eta \right\} \\ &= A_{x,\omega}^{(i,j)} + A_{y,\omega}^{(i,j)}, \quad k, l, m, n = 1, 2, \quad (5.63) \\ A_{x,\omega}^{(i,j)} &= \frac{1}{6} \frac{h_j^y}{h_i^x} \begin{bmatrix} 2 + \omega/15 & -2 - \omega/15 & 1 - \omega/15 & -1 + \omega/15 \\ -2 - \omega/15 & 2 + \omega/15 & -1 + \omega/15 & 1 - \omega/15 \\ 1 - \omega/15 & -1 + \omega/15 & 2 + \omega/15 & -2 - \omega/15 \\ -1 + \omega/15 & 1 - \omega/15 & -2 - \omega/15 & 2 + \omega/15 \end{bmatrix}, \\ A_{y,\omega}^{(i,j)} &= \frac{1}{6} \frac{h_i^x}{h_j^y} \begin{bmatrix} 2 + \omega/15 & 1 - \omega/15 & -2 - \omega/15 & -1 + \omega/15 \\ 1 - \omega/15 & 2 + \omega/15 & -1 + \omega/15 & -2 - \omega/15 \\ -2 - \omega/15 & -1 + \omega/15 & 2 + \omega/15 & 1 - \omega/15 \\ -1 + \omega/15 & -2 - \omega/15 & 1 - \omega/15 & 2 + \omega/15 \end{bmatrix}. \end{aligned}$$

Отметим, что матрица $A_{15}^{(i,j)}$, получаемая из (5.63) при $\omega=15$, соответствует стандартной пятиточечной аппроксимации оператора Лапласа вида (5.50).

На конечномерном пространстве (5.62) традиционное применение метода Галеркина дает следующую приближенную вариационную задачу: найти сеточную функцию

$$u^h(x, y) = \sum_{i,j} u_{i,j} \psi_{i,j}(x, y) \in V_\omega^h \quad (5.64)$$

такую, что

$$a(u^h, v^h) = f(v^h), \quad v^h \in V_\omega^h. \quad (5.65)$$

Сформулируем теперь модифицированную вариационную сеточную задачу: найти функцию $u^h \in V_\omega^h$ такую, что

$$a(u^h, v^h) = f(P_h v^h), \quad v^h \in V_\omega^h, \quad (5.66)$$

где P_h есть оператор интерполирования

$$P_h v^h = \sum_{i,j} v_{i,j} \varphi_{i,j}(x, y), \quad (5.67)$$

использующий вышеприведенные билинейные базисные функции.

Получаемая из (5.66) система линейных алгебраических уравнений обладает тем уникальным свойством, что при $\omega = 15/2$ на равномерной сетке ($h_i^x = h^x$, $h_j^y = h^y$) она совпадает со схемой Микеладзе (5.59), имеющей четвертый порядок точности. Более конкретно, в теории конечно-разностных методов оценка $O(h^4)$ характеризует разность точного и приближенного решений только в узлах сетки (при ограниченности производных искомого решения до шестого порядка включительно).

Мы не будем останавливаться на нетривиальном исследовании погрешности обобщенного решения, а только сформулируем итоговый результат (для равномерной сетки):

$$\|u - u^h\|_{H_0^1(\Omega)} \leq c \left\{ \left| \omega - \frac{15}{2} \right| h^2 \|u\|_{C^3(\bar{\Omega})} + h^4 \|u\|_{H^5(\Omega)} \right\}.$$

В частности, непрерывность вложения $H^5(\Omega)$ в $C^3(\bar{\Omega})$ при $u \in H^5(\Omega)$ обеспечивает принадлежность решения пространству $C^3(\bar{\Omega})$. Таким образом, в норме указанного соболевского пространства ошибка оказывается порядка $O(h^2)$ при любом значении ω и $O(h^4)$ — при $\omega = 15/2$.

5.3.3. Изопараметрические преобразования и криволинейные элементы

Рассмотренные выше аппроксимации на треугольных и прямоугольных элементах не охватывают всех возможных геометрических конфигураций, которые часто встречаются в практических задачах. Для областей с криволинейной границей самым простым и самым грубым подходом является замена кривой, например второго порядка, ломаной линией. Другой подход может использовать на околограничных криволинейных элементах специальные аппроксимации с выбором базисных функций в локальных системах координат.

Наконец, одним из развитых математических направлений является конформное (или квазиконформное) отображение областей различных конфигураций в некоторые стандартные. Здесь, как правило, задача сводится, в свою очередь, к решению непростых краевых задач для специальных уравнений в частных производных.

Однако наибольшее распространение получила математическая технология, основанная на локальных преобразованиях «неправильных» элементов в некоторые стандартные, где уже в новых переменных вычисляются элементы матриц МКЭ. Естественно, что после нахождения приближенного решения может потребоваться его представление в исходных переменных и что преобразование координат должно быть по крайней мере невырожденным.

Наиболее типичные преобразования: «неправильный четырехугольник–квадрат», «неправильный треугольник–стандартный (прямоугольный равнобедренный) треугольник», «треугольник с одной криволинейной стороной–стандартный треугольник», примеры которых изображены на рис. 5.10–5.12.

Наиболее естественными являются так называемые *изопараметрические преобразования*, которые строятся в том же классе многочленов, что и сами применяемые аппроксимации. При этом достигается «золотая середина» между сохранением необходимой гладкости кусочно-полиномиальных приближений и вычислительной сложностью алгоритма.

Рассмотрим преобразование прямолинейного четырехугольника с вершинами $P_k(x_k, y_k)$, $k = 1, 2, 3, 4$, в единичный квадрат в новой системе координат $0 \leq \xi, \eta \leq 1$, см. рис. 5.10. Формулы

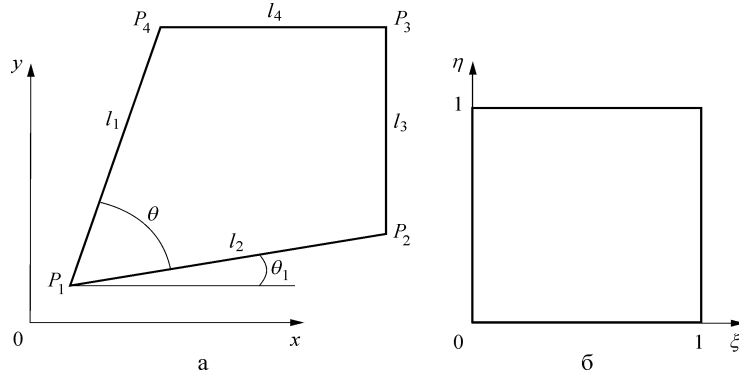


Рис. 5.10. Преобразование четырехугольника в единичный квадрат

такого преобразования записываются в виде

$$\begin{aligned} x &= x(\xi, \eta) = x_1 + (x_2 - x_1)\xi + (x_3 - x_1)\eta + a\xi\eta, \\ y &= y(\xi, \eta) = y_1 + (y_2 - y_1)\xi + (y_3 - y_1)\eta + b\xi\eta, \end{aligned} \quad (5.68)$$

где используются обозначения для коэффициентов

$$a = x_4 - x_3 - x_2 + x_1, \quad b = y_4 - y_3 - y_2 + y_1.$$

При этом, очевидно, точка P_1 переходит в $(0, 0)$, P_2 в $(1, 0)$ и т. д.

Якобиан этого билинейного преобразования имеет форму

$$I(\xi, \eta) = \begin{vmatrix} \frac{\partial x}{\partial \xi} & \frac{\partial x}{\partial \eta} \\ \frac{\partial y}{\partial \xi} & \frac{\partial y}{\partial \eta} \end{vmatrix} = \begin{vmatrix} x_2 - x_1 + a\eta & x_3 - x_1 + a\xi \\ y_2 - y_1 + b\eta & y_3 - y_1 + b\xi \end{vmatrix},$$

которая является линейной относительно ξ, η , так как коэффициент при члене $\xi\eta$ равен нулю.

Проверим, при каких условиях данный якобиан во всех четырех вершинах квадрата имеет одинаковый знак и, следовательно, не обращается в нуль ни в одной из внутренних точек. Рассмотрим $I(\xi, \eta)$ при $\xi = \eta = 0$:

$$\begin{aligned} I(0, 0) &= (x_2 - x_1)(y_3 - y_1) - (y_2 - y_1)(x_3 - x_1) \\ &= l_1 l_2 \left[\frac{(x_2 - x_1)(y_3 - y_1)}{l_1 l_2} - \frac{(y_2 - y_1)(x_3 - x_1)}{l_1 l_2} \right] \\ &= l_1 l_2 [\cos \theta_1 \sin(\theta + \theta_1) - \sin \theta_1 \cos(\theta + \theta_1)] = l_1 l_2 \sin \theta, \end{aligned}$$

где $l_1, l_2, \theta, \theta_1$ — длины сторон исходного четырехугольника и углы, изображенные на рис. 5.10. Отсюда видно, что если угол четырехугольника с вершиной P_1 удовлетворяет условию $0 < \theta < \pi$, то $I(0, 0) > 0$. Аналогично показывается, что если остальные углы тоже меньше π , то значения якобиана во всех вершинах четырехугольника и, следовательно, во внутренних точках будут положительными. Таким образом, якобиан строго больше нуля и преобразование (5.59) невырождено, если четырехугольник выпуклый, т. е. каждый из его углов меньше π . При этом, очевидно, обратное преобразование $(x, y) \rightarrow (\xi, \eta)$ будет также невырождено.

На практике формулы обратного преобразования $\xi = \xi(x, y)$, $\eta = \eta(x, y)$ строить не обязательно, так как для вычисления элементов матриц МКЭ достаточно пользоваться формулами (5.59).

Проиллюстрируем это на вычислении диагонального элемента матрицы жесткости, соответствующего некоторой базисной функции $\varphi(x, y)$:

$$\int p(x, y)(\varphi_x^2 + \varphi_y^2) dx dy = \int p(x(\xi, \eta), y(\xi, \eta)) [(\varphi_\xi \xi_x + \varphi_\eta \eta_x)^2 + (\varphi_\xi \xi_y + \varphi_\eta \eta_y)^2] I(\xi, \eta) d\xi d\eta. \quad (5.69)$$

Если функция $p(x, y)$ постоянна на элементе или имеет полиномиальный вид, то вычисление интеграла (5.60) на стандартном квадрате в переменных ξ, η легко осуществляется точно. В противном случае ее можно предварительно аппроксимировать, например, с помощью тех же базисных функций.

Рассмотрим теперь преобразование неправильного треугольника в «стандартный» прямоугольный равнобедренный треугольник, см. рис. 5.11, что также позволяет существенно упростить вычисления матриц МКЭ.

Здесь достаточно применить линейное преобразование

$$\begin{aligned} x &= x_1 + (x_2 - x_1)\xi + (x_3 - x_1)\eta, \\ y &= y_1 + (y_2 - y_1)\xi + (y_3 - y_1)\eta, \end{aligned} \quad (5.70)$$

при котором точки P_1, P_2, P_3 переходят, соответственно, в $(0, 0)$, $(1, 0)$, $(0, 1)$, см. обозначения на рис. 5.7. Якобиан этого преобразования есть

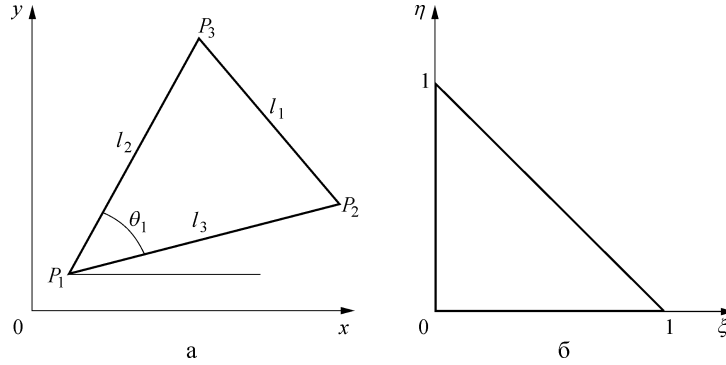


Рис. 5.11. Преобразование неправильного треугольника в стандартный

$$\begin{aligned} I(\xi, \eta) &= (x_2 - x_1)(y_3 - y_1) - (y_2 - y_1)(x_3 - x_1) \\ &= l_2 l_3 \sin \theta_1 = 2S, \end{aligned} \quad (5.71)$$

т. е. равен удвоенной площади исходного треугольника. Обратное к (5.70) преобразование является также линейным:

$$\begin{aligned} \xi &= [(y_3 - y_1)(x - x_1) - (x_3 - x_1)(y - y_1)] \frac{1}{2S}, \\ \eta &= [(x_2 - x_1)(y - y_1) - (y_2 - y_1)(x - x_1)] \frac{1}{2S}. \end{aligned} \quad (5.72)$$

Вычисления элементов матриц МКЭ очевидным образом могут проводиться по формулам типа (5.69).

Заметим, что в случае постоянного на элементе коэффициента диффузии p и кусочно-линейных базисных функций нахождение элементов матрицы жесткости легко делается и без преобразования прямолинейных треугольников.

Как известно (см., например [82]), в двумерном случае по трем точкам P_k , $k = 1, 2, 3$, не лежащим на одной прямой, интерполяционный многочлен Лежандра (первого порядка по каждой из переменных) определяется однозначно:

$$v^h(x, y) = v_1 \varphi_1(x, y) + v_2 \varphi_2(x, y) + v_3 \varphi_3(x, y),$$

где базисные функции обладают свойством $\varphi_i(P_j) = \delta_{i,j}$ и имеют следующий вид (S означает площадь треугольника $P_1 P_2 P_3$):

$$\varphi_1(x, y) = [(x - x_2)(y_2 - y_3) - (y - y_2)(x_2 - x_3)] \frac{1}{2S},$$

$$\begin{aligned}\varphi_2(x, y) &= [(x - x_3)(y_3 - y_1) - (y - y_3)(x_3 - x_1)] \frac{1}{2S}, \\ \varphi_3(x, y) &= [(x - x_1)(y_1 - y_2) - (y - y_1)(x_1 - x_2)] \frac{1}{2S}.\end{aligned}$$

Производные этих функций вычисляются элементарно:

$$\begin{aligned}\varphi_1^x &= \frac{y_2 - y_3}{2S}, & \varphi_1^y &= \frac{x_2 - x_3}{2S}, & \varphi_2^x &= \frac{y_3 - y_1}{2S}, \\ \varphi_2^y &= \frac{x_3 - x_1}{2S}, & \varphi_3^x &= \frac{y_1 - y_2}{2S}, & \varphi_3^y &= \frac{x_1 - x_2}{2S}.\end{aligned}$$

А отсюда при $p(x, y) = 1$ элементы матрицы жесткости (третьего порядка) определяются следующими элегантными формулами:

$$\begin{aligned}a_{ij} = a(\varphi_i, \varphi_j) &= \int_S (\varphi_i^x \varphi_j^x + \varphi_i^y \varphi_j^y) dx dy = \frac{-l_i l_j}{4S} \cos \theta_{i,j}, \quad i \neq j, \\ a_{ii} &= \frac{l_i^2}{4S}, \quad i, j = 1, 2, 3.\end{aligned}$$

Здесь l_i — длина стороны треугольника, лежащей напротив вершины P_i , а $\theta_{i,j}$ — угол треугольника между i -й и j -й сторонами.

Легко видеть, что у данной матрицы диагональные элементы положительны, внедиагональные — отрицательны и сумма элементов каждой строки равна нулю. Это справедливо для любого треугольника, не вырождающегося в отрезок. Рассмотрим теперь технологические вопросы, связанные с *криволинейными конечными элементами*.

Более сложным преобразованием, по сравнению с (5.70), но возможным, является такое, в котором стандартный треугольник в новых переменных ξ, η получается из исходного треугольника с одной криволинейной границей, описываемой полиномом второй степени, см. рис. 5.12.

В данном случае преобразование удобно провести в два этапа. На первом исходный треугольник переводим в прямоугольный равнобедренный треугольник с «криволинейной гипотенузой», в котором в новых координатах x', y' вершинами являются точки $(0, 0)$, $(1, 0)$, $(0, 1)$, причем в первой из них угол прямой. Такое преобразование, очевидно, имеет вид (5.70) (переменные ξ, η заменяются соответственно на x', y'), а криволинейная сторона в новых независимых переменных по-прежнему, описывается многочленом второго порядка. Второй этап преобразования заключается в аппроксимации полученного криволинейного треугольника «стандартным» прямолинейным треугольником, см.,

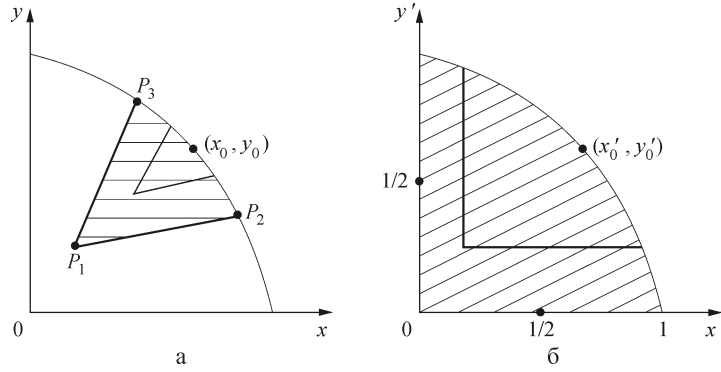


Рис. 5.12. Преобразование треугольника с одной криволинейной стороной

соответственно, рис. 5.11 и 5.10. Для этого можно применить отображение вида

$$x' = \xi + (4x'_0 - 2)\xi\eta, \quad y' = \eta + (4y'_0 - 2)\xi\eta, \quad (5.73)$$

где x'_0, y'_0 — координаты точки на криволинейной гипотенузе треугольника, в которую переходит точка $\xi = \eta = 1/2$. Якобиан преобразования (5.73) линеен и имеет вид

$$I(\xi, \eta) = \begin{vmatrix} x'_\xi & x'_\eta \\ y'_\xi & y'_\eta \end{vmatrix} = 1 + (4y'_0 - 2)\xi + (4x'_0 - 2)\eta.$$

Он является строго положительным в вершинах и, следовательно, во всех внутренних точках треугольника, если выполняются неравенства

$$x'_0 > 1/4, \quad y'_0 > 1/4, \quad (5.74)$$

определяющие обозначенный на рис. 5.12, б штриховой линией сектор, в котором должна находиться точка (x'_0, y'_0) . В свою очередь, этой точке соответствует точка с координатами $x_0(x'_0, y'_0)$, $y_0(x'_0, y'_0)$ на криволинейной границе исходного треугольника, а условия (5.74) определяют в переменных x, y также некоторый сектор, в котором должна находиться точка x_0, y_0 для обеспечения рассматриваемого преобразования.

Приведем также пример более общего преобразования $(\xi, \eta) \rightarrow (x, y)$, переводящего треугольник с тремя криволи-

нейными сторонами в стандартный прямоугольный равнобедренный треугольник и основанного на представлении функций $x(\xi, \eta)$, $y(\xi, \eta)$ многочленами второго порядка:

$$\begin{aligned} x &= x_3 + (4x_6 - x_1 - 3x_3)\xi + 4(x_5 - x_2 - 3x_3)\eta + \\ &\quad 2(x_1 + x_3 - 2x_6)\xi^2 + 4(x_3 + x_4 - x_5 - x_6)\xi\eta + \\ &\quad 2(x_2 + x_3 - 2x_5)\eta^2, \\ y &= y_3 + (4y_6 - y_1 - 3y_3)\xi + (4y_5 - y_2 - 3y_3)\eta + \\ &\quad 2(y_1 + y_3 - 2y_6)\xi^2 + 4(y_3 + y_4 - y_5 - y_6)\xi\eta + \\ &\quad 2(y_2 + y_3 - 2y_5)\eta^2. \end{aligned} \quad (5.75)$$

Здесь x_k , y_k , $k = 1, \dots, 6$, — координаты вершин и некоторых срединных точек сторон криволинейного треугольника, которые переводятся в соответствующие вершины и середины сторон

$$(1, 0), (0, 1)(0, 0), \left(\frac{1}{2}, \frac{1}{2}\right), \left(0, \frac{1}{2}\right)\left(\frac{1}{2}, 0\right)$$

прямолинейного треугольника, см. рис. 5.13.

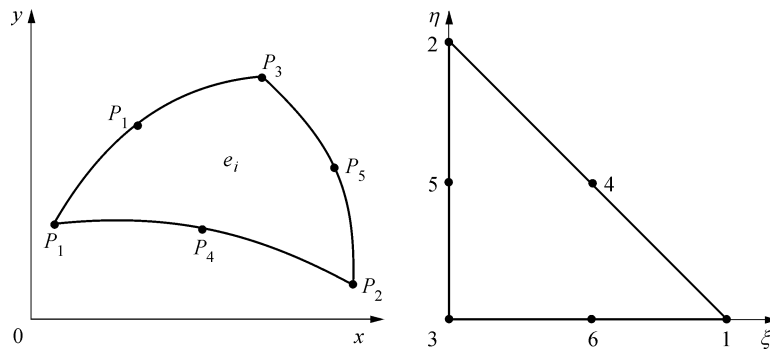


Рис. 5.13. Преобразование криволинейного треугольника в прямоугольный равнобедренный

Аналогичное преобразование с помощью полиномов второго порядка можно построить для перевода криволинейного четырехугольника в квадрат $-1 \leq \xi, \eta \leq 1$. Рассмотрим, например, изображенный на рис. 5.14 четырехугольник, у которого вторая сторона (между вершинами P_2 и P_3) записывается в параметрической форме $x = x_2(\eta)$, $y = y_2(\eta)$, причем

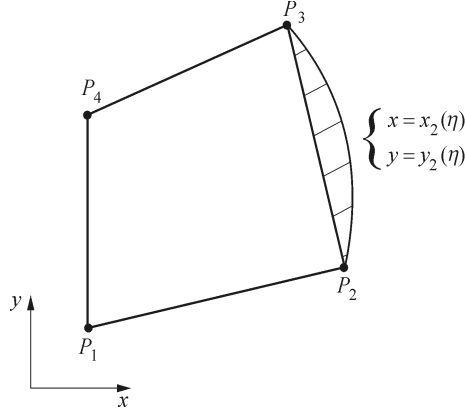


Рис. 5.14. Преобразование четырехугольника с одной криволинейной стороной

$$x(-1) = X_2, \quad y(-1) = Y_2, \quad x(1) = X_3, \quad y(1) = Y_3.$$

В этом случае указанное преобразование для x -координаты описывается наглядной формулой

$$x = \frac{1}{4}(1 - \xi)(1 - \eta)X_1 + \frac{1}{4}(1 + \xi)(1 - \eta)X_2 + \frac{1}{4}(1 + \xi)(1 + \eta)X_3 + \frac{1}{4}(1 - \xi)(1 + \eta)X_4 + \left[x_2(\eta) - \frac{1 - \eta}{2}X_2 - \frac{1 + \eta}{2}X_3 \right] \frac{1 + \xi}{2}. \quad (5.76)$$

Здесь первые четыре члена правой части осуществляют линейное отображение прямолинейного четырехугольника P_1, P_2, P_3, P_4 в квадрат, а пятый член состоит из произведения двух функций. Первый множитель (в квадратных скобках) представляет разность между $x_2(\eta)$ и соответствующей x -координатой точки отрезка P_2, P_3 , а второй множитель есть линейная функция, меняющая свое значение от единицы на второй стороне до нуля на четвертой стороне четырехугольника.

После сокращения взаимноуничтожающихся членов соотношение (5.76), а также аналогичное выражение для y -координаты принимают следующий вид:

$$\begin{aligned} x &= \frac{1}{4}(1 - \xi)(1 - \eta)X_1 + \frac{1}{4}(1 - \xi)(1 + \eta)X_4 + x_2(\eta)\frac{1 + \xi}{2}, \\ y &= \frac{1}{4}(1 - \xi)(1 - \eta)Y_1 + \frac{1}{4}(1 - \xi)(1 + \eta)Y_4 + y_2(\eta)\frac{1 + \xi}{2}. \end{aligned} \quad (5.77)$$

В общем случае все четыре стороны четырехугольника могут быть криволинейными. Записывая их в параметрической форме

$$\begin{aligned} x &= x_i(\xi), & y &= y_i(\xi), & -1 \leq \xi \leq 1, & i = 1, 3, \\ x &= x_i(\eta), & y &= y_i(\eta), & -1 \leq \eta \leq 1, & i = 2, 4, \end{aligned}$$

и обозначая через $X_k, Y_k, k = 1, 2, 3, 4$, глобальные координаты четырех вершин, получаем следующие формулы отображения:

$$\begin{aligned} x &= \frac{1}{2}(1 - \eta)x_1(\eta) + \frac{1}{2}(1 + \xi)x_2(\eta) + \frac{1}{2}(1 + \eta)x_3(\xi) + \\ &\quad \frac{1}{2}(1 - \xi)x_4(\eta) - \frac{1}{4}(1 - \xi)(1 - \eta)X_1 - \frac{1}{4}(1 + \xi)(1 - \eta)X_2 \\ &\quad \frac{1}{4}(1 + \xi)(1 + \eta)X_3 - \frac{1}{4}(1 - \xi)(1 + \eta)X_4, \\ y &= \frac{1}{2}(1 - \eta)y_1(\xi) + \frac{1}{2}(1 + \xi)y_2(\eta) + \frac{1}{2}(1 + \eta)y_3(\xi) + \\ &\quad \frac{1}{2}(1 - \eta)y_4(\eta) - \frac{1}{4}(1 - \xi)(1 - \eta)Y_1 - \frac{1}{4}(1 + \xi)(1 - \eta)Y_2 \\ &\quad \frac{1}{4}(1 + \xi)(1 + \eta)Y_3 - \frac{1}{4}(1 - \xi)(1 + \eta)Y_4. \end{aligned} \quad (5.78)$$

Обратное отображение $\xi = \xi(x, y), \eta = \eta(x, y)$ в общем случае в явном виде не представляется, но обычно для его реализации эффективно используются численные итерационные алгоритмы.

5.3.4. Элементы Равьяра–Тома

Для актуальных во многих приложениях смешанных МКЭ разработаны различные эффективные аппроксимации в соответствующих соболевских пространствах, достаточно полно изложенные в монографии Ф.Бреции и М.Фортина [27]. Мы в данном пункте остановимся на конечных элементах Равьяра–Тома, широко применяемых для построения базисных d -мерных векторных функций $\vec{q} = (q_1, \dots, q_d)$ в пространствах

$$H(\operatorname{div}, \Omega) = \{\vec{q} | \vec{q} \in (L^2(\Omega))^\alpha, \operatorname{div} \vec{q} \in (L^2(\Omega))\}, \quad (5.79)$$

для которых нормы определяются следующим образом (см. также обозначения в пп. 1.1.2, 1.1.4):

$$\begin{aligned} \|\vec{q}\|_{div, \Omega}^2 &= |\vec{q}|_{0, \Omega}^2 + |\operatorname{div} \vec{q}|_{0, \Omega}^2, \\ |v|_{m, \Omega}^2 &= \sum_{|\alpha|=m} \int_{\Omega} |D^\alpha v|^2 d\vec{x}, \quad |\alpha| = \alpha_1 + \dots + \alpha_m. \end{aligned} \quad (5.80)$$

Поскольку обобщенные решения многих интересных задач моделирования (теплопереноса, упругости, электромагнетизма и т.д.) лежат в $H(\operatorname{div}, \Omega)$, мы сначала коротко остановимся на свойствах этих пространств (в дополнении к материалу из первой главы), а затем рассмотрим некоторые особенности их аппроксимаций с помощью специальных конечно-элементных базисных функций.

Пусть $H^m(\Omega)$ есть пространство соболевских функций, в области Ω с липшицевой границей Γ , см. § 1.1. Соответствующую норму обозначим как

$$\|v\|_{m,\Omega} = \left(\sum_{|\alpha| \leq m} \int_{\Omega} |D^\alpha u|^2 d\vec{x} \right)^{1/2}.$$

Известно, что для функций $v \in H^1(\Omega)$ с нормой $\|v\|_{1,\Omega}^2 = |v|_{0,\Omega}^2 + |v|_{1,\Omega}^2$ их следы на границе Γ , образуют гильбертово пространство $H^{1/2}(\Gamma)$, являющееся плотным подпространством в $L^2(\Omega)$, и для оператора следа $\gamma : H^1(\Omega) \rightarrow H^{1/2}(\Gamma)$ норма определяется следующим образом:

$$\|\gamma v\|_{1/2,\Gamma} = \inf \|w\|_{1,\Omega}, \quad w \in H^1(\Omega), \quad \gamma w = v. \quad (5.81)$$

Если $\bar{v}$ есть единственное решение в $H^1(\Omega)$ задачи Дирихле для уравнения Гельмгольца

$$-\Delta \bar{v} + \bar{v} = 0, \quad \bar{v}|_{\Gamma} = v, \quad (5.82)$$

то справедливо соотношение

$$\|v\|_{1/2,\Gamma} = \|\bar{v}\|_{1,\Omega}.$$

Кроме того, на границе можно определить пространство функций v^* в пространстве $H^{-1/2}(\Gamma)$, являющимся двойственным к $H^{1/2}(\Gamma)$ и имеющим норму

$$\|v^*\|_{-1/2,\Gamma} = \sup_{v \in H^{1/2}(\Gamma)} \frac{\langle v, v^* \rangle}{\|v\|_{1/2,\Gamma}} = \|\bar{v}^*\|_{1,\Omega},$$

где $\bar{v}^*$ есть решение вариационной задачи Неймана

$$\int_{\Omega} \operatorname{grad} \bar{v}^* \operatorname{grad} v d\vec{x} + \int_{\Omega} \bar{v}^* \bar{v} d\vec{x} = \langle v^*, v \rangle, \quad \forall v \in H^1(\Omega),$$

а скобка $\langle \cdot, \cdot \rangle$ означает отношение двойственности между $H^{-1/2}(\Gamma)$ и $H^{1/2}(\Gamma)$.

Для векторной функции $\vec{q} \in H(\operatorname{div}; \Omega)$ определяется скалярная величина – *след нормали* на границе

$$\vec{q} \cdot \vec{n}|_{\Gamma} \in H^{-1/2}(\Gamma),$$

а также справедлива формула Грина

$$\int_{\Omega} \operatorname{div} \vec{q} v d\vec{x} + \int_{\Omega} \vec{q} \cdot \operatorname{grad} v d\vec{x} = \langle \vec{q} \cdot \vec{n}, v \rangle, \quad \forall v \in H^1(\Omega).$$

Если границу области разбить на две непересекающиеся части ($\Gamma = \Gamma_D \cup \Gamma_N$), то можно определить пространство

$$H_{0,\Gamma_D}^1 = \{v | v \in H^1(\Omega), v|_{\Gamma_D} = 0\},$$

которое при $\Gamma_D = \Gamma$ и $\Gamma_N = \Gamma$ совпадает с пространствами $H_0'(\Omega)$ и $H^1(\Omega)$ из пп. 1.1.4 – 1.1.7 соответственно.

По аналогии можно ввести пространство

$$H_{0,\Gamma_N}(\operatorname{div}; \Omega) = \{\vec{q} | \vec{q} \in H(\operatorname{div}; \Omega), (\vec{q} \cdot \vec{n}, v) = 0, \forall v \in H_{0,\Gamma_D}^1(\Omega)\},$$

содержащее функции из $H(\operatorname{div}; \Omega)$, у которых следы нормали обращаются в нуль на Γ_N . В частности, мы будем обозначать $H_0(\operatorname{div}; \Omega) = H_{0,\Gamma_N}(\operatorname{div}; \Omega)$, если $\Gamma_N = \Gamma$. Наконец, можно отметить еще одно важное подпространство в $H(\operatorname{div}; \Omega)$:

$$N^0(\operatorname{div}; \Omega) = \{\vec{q} | \vec{q} \in H(\operatorname{div}; \Omega), \operatorname{div} \vec{q} = 0\},$$

состоящее из бездивергентных векторных функций $\vec{q}$.

Рассмотрим теперь конечно-элементные аппроксимации функций из $H(\operatorname{div}; \Omega)$ в предположении, что $\Omega^h = \bigcup_{k=1}^K \Omega_k^h$ есть регулярная декомпозиция, или триангуляция, области Ω с границей Γ на K непересекающихся конечных элементов Ω_k^h , которые, например, в двумерном случае могут быть треугольниками или четырехугольниками, а в трехмерном – четырехгранниками (тетраэдрами) или шестигранниками. Граница k -го элемента

$\Gamma_k^h = \bigcup_{k'} \Gamma_{k,k'}^h$ состоит из граней $\Gamma_{k,k'}^h = \bar{\Omega}_k^h \cap \bar{\Omega}_{k'}^h$, каждая из которых k' может представляться сегментом плоскости или криволинейной поверхностью. Формально элементом с номером $k = 0$ можно считать дополнение к области Ω^h , т.е. $\Omega_0^h = \mathcal{R}^d \setminus \Omega^h$, и при этом грани $\Gamma_{k,0}^h$ будут принадлежать внешней границы Γ^h области Ω^h .

Частным случаем конечного элемента является симплекс, т.е. d -мерный выпуклый многогранник, имеющий вершинами $d + 1$ точку и для $d = 0, 1, 2, 3$ представляющий собой точку, отрезок, треугольник и тетраэдр соответственно, причем каждая его грань есть $(d - 1)$ -мерный симплекс.

Для каждого типа элемента T_h с границей Γ_h определяем соответствующий *стандартный элемент* $\hat{T}$ с границей $\hat{\Gamma}$ (единичные равнобедренные прямоугольные треугольник и тетраэдр, единичный квадрат или куб). Мы будем предполагать, что $T = \mathcal{F}(\hat{T})$, где $\mathcal{F}$ есть некоторое достаточно гладкое (по крайней мере из C^1) отображение из $\mathcal{R}^d$ в $\mathcal{R}^d$, причем якобиан $D\mathcal{F}(\hat{x})$ при $\hat{x} \in \hat{T}$ является обратимым и $D\mathcal{F}^{-1}(x) = (D\mathcal{F}(\hat{x}))^{-1}$ при $x \in T$. Более того, иногда конечные элементы T будут предполагаться *аффинными*, а конкретнее, когда отображение $\mathcal{F}$ имеет вид

$$\mathcal{F}(\hat{x}) = x_0 + B\hat{x}, \quad \hat{x} \in \hat{T}, \quad x_0 \in T,$$

где $B = D\mathcal{F}(\hat{x})$ является постоянной невырожденной матрицей. В общем случае определяем $J(\hat{x}) = |\det D\mathcal{F}(\hat{x})|$ и для $\hat{x} \in \hat{\Gamma}$ имеем

$$J_{\hat{n}}(\hat{x}) = J(\hat{x})|(D\mathcal{F}^{-1})^T \hat{n}|.$$

Если $\hat{v}(\hat{x})$ есть функция в $\hat{T}$, то мы определяем $v(x)$ в T как

$$v = \hat{v} \circ \mathcal{F}^{-1} = \tilde{f}(\hat{v}).$$

При этом имеют место классические формулы

$$\begin{aligned} \text{grad } v &= (D\mathcal{F}^{-1})^T \text{grad } \hat{v} \circ \mathcal{F}^{-1} = \tilde{f}(D\mathcal{F}^{-1}) \text{grad } v, \\ \int_{T_h} \tilde{f}(\hat{v}) d\vec{x} &= \int_{\hat{T}} \hat{v} J d\hat{x}, \quad \int_{\Gamma_h} \tilde{f}(\hat{v}) ds = \int_{\hat{\Gamma}} \tilde{f}(\hat{v}) J_{\hat{n}} d\hat{s}. \end{aligned}$$

В данном пункте мы будем рассматривать только *конформные аппроксимации*, которые являются наиболее естественными для МКЭ, в том смысле, что они обеспечивают *внутренние*

приближения (в себе), т.е. аппроксимации сходного пространства строятся с помощью принадлежащих ему конечно-мерных подпространств.

Более конкретно, будут рассматриваться только полиномиальные аппроксимации, в связи с чем дадим следующие определения и обозначения:

- $P_K(T)$ есть пространство полиномов степени не больше K ; размерности данных пространств равны $(K+1)(K+2)/2$ для $d=2$ и $(K+1)(K+2)(K+3)/6$ при $d=3$;
- $P_{K_1, K_2}(T) = \{p(x_1, x_2) | p(x_1, x_2) = \sum_{\substack{i \leq K_1 \\ j \leq K_2}} a_{i,j} x_1^i x_2^j\}$ пространства полиномов степени $\leq K_1$ по x_1 и $\leq K_2$ по x_2 (аналогично определяются пространства полиномов $P_{K_1, K_2, K_3}(T)$ для $d=3$; размерности этих пространств равны $(K_1+1)(K_2+1)$ и $(K_1+1)(K_2+1)(K_3+1)$ соответственно;

•

$$Q_k(T) = \begin{cases} P_{K,K}(T) & \text{для } d=2, \\ P_{K,K,K}(T) & \text{для } d=3; \end{cases}$$

- пространства полиномов, определенные на гранях E_i элемента:

$$R_K(\Gamma_h) = \{\varphi | \varphi \in L^2(\Gamma_h), \varphi|_{E_i} \in P_K(E_i), \forall E_i\},$$

$$R_K^0(\Gamma_h) = \{\varphi | \varphi \in R_K(\Gamma_h) \cap C^0(\Gamma_h)\}.$$

Перейдем теперь непосредственно к аппроксимациям функций из $H(\operatorname{div}; \Omega)$ с помощью базисных функций Равьяра-Тома. Для случая d -симплициальных элементов (треугольники или тетраэдры) определим пространства

$$RT_K(T_h) = (P_K(T_h))^d + \tilde{x}P_K(T_h),$$

размерности которых равны $(K+1)(K+3)$ для $d=2$ и $(K+1)(K+2)(K+4)/2$ для $d=3$. Предварительно установим несколько важных свойств для пространств Равьяра-Тома.

- Для любого вектора $\vec{q} \in RT_K(T)$ справедливы следующие соотношения:

$$\operatorname{div} \vec{q} \in P_K(T_h), \quad \vec{q} \cdot \vec{n}|_{\Gamma_h} \in R_K(\Gamma_h).$$

- Если определить пространство бездивергентных функций

$$RT_K^0(T_h) = \{\vec{q} \in RT_K(T_h) \mid \operatorname{div} \vec{q} = 0\},$$

то для него справедливо $RT_K^0(T_h) \subset (P_K(T_h))^d$. При $d = 2$ любой вектор $\vec{q} \in RT_K^0(T_h)$ представим в виде $\operatorname{rot} \psi$, где $\psi \in P_{K+1}(T_h)$ есть некоторая функция тока. Данный результат был обобщен Ж.Неделеком на случай $d = 3$ в известной работе J.C.Nedelec "A new family of mixed finite elements in $\mathcal{R}^3$ ". Numer. Math., v. 50, 1986, 57-81, где были предложены конечно-элементные аппроксимации для другого важного класса пространств –

$$H(\operatorname{rot} 6; \Omega) = \{\vec{\psi} \mid \vec{\psi} \in (L^2(\Omega))^d, \operatorname{rot} \psi \in (L^2(\Omega))^d\}.$$

- При $K \geq 0$ для любого $\vec{q} \in RT_K(T_h)$ из соотношений

$$\begin{aligned} \int_{\Gamma_h} \vec{q} \cdot \vec{n} p_k ds &= 0 \quad \forall p_k \in R_k(\Gamma_h), \\ \int_{\Gamma_h} \vec{q} p_{k-1} d\vec{x} &= 0 \quad \forall p_{k-1} \in (p_{k-1}(T_h))^d, \end{aligned}$$

следует бездивергентность данного вектора ($\operatorname{div} \vec{q} = 0$).

Из приведенных утверждений следует, что в качестве степеней свободы для $RT_k(T_h)$ можно взять моменты вплоть до степени k на гранях Γ_h , а также моменты до порядка до $K - 1$ от вектора $\vec{q}$ внутри T_h . Важно отметить, что данные множества степеней свободы, фактически представляемых набором линейных форм, являются унисольвентными, т.е. однозначно определяют все полиномы в рассматриваемых пространствах.

Например, в двумерном случае пространство RT_0 имеет размерность 3 и состоит из векторных функций, имеющих компоненты

$$q_1(x_1, x_2) = a + cx_1, \quad q_2(x_1, x_2) = b + cx_2,$$

коэффициенты a, b, c которых определяются по следам нормалей на сторонах треугольника. Аналогично, пространство RT_1 определяется по двум значениям $\vec{q} \cdot \vec{n}$ на каждой из трех сторон, а также одному значению вектора $\vec{q}$ в центре треугольника T_h . Внутри тетраэдра пространства RT_0 и RT_1 имеют размерности 4 и 13 соответственно.

Если конечные элементы являются не симплицальными, а прямоугольными, то пространства Равьяра-Тома строятся сначала на стандартных элементах $\hat{T} = (-1, +1)^d$:

$$RT_{[k]} = (Q_k)^d + \vec{x}Q_k.$$

При этом справедливы также следующие представления:

$$RT_{[k]} = \begin{cases} P_{k+1,k} \times P_{k,k+1} & \text{при } d = 2, \\ P_{k+1,k,k} \times P_{k,k+1,k} \times P_{k,k,k+1} & \text{при } d = 3. \end{cases}$$

Размерности этих пространств равны $2(k+1)(k+2)$ $3(k+1)^2(k+2)$ соответственно. Как и в симплицальных элементах, построение унисольвентных базисных функций строится по совокупности моментов, представляемых линейными формами $l_i(\hat{p})$, $i = 1, \dots, \dim \hat{p}$, коэффициенты которых находятся из формируемых условий интерполяции. Например, в конечных элементах лагранжевого типа соответствующие уравнения формируются по точечным значениям функций в узлах интерполяции $a_i \in \hat{T}$:

$$l_i(\hat{p}) = \hat{p}(\vec{a}_i), \quad i = 1, \dots, \dim \hat{p}.$$

При использовании эрмитовых элементов условия интерполяции включают уравнения, соответствующие производным различных порядков в точках граней и внутри элемента.

В общем случае множество функционалов определяется также следами нормалей на гранях стандартного элемента. Например, для квадратного элемента при $d = 2$ пространство $RT_{[0]}(\hat{T})$ размерности $\dim = 4$ определяется по значениям нормалей $\vec{q} \cdot \vec{n}$ на четырех сторонах. Для пространств $RT_{[1]}$ и $RT_{[2]}$ на каждой стороне задаются по 2 или 3 нормали соответственно и добавляются по 4 или 12 внутренних значений векторной функции, так что размерности этих пространств равны $4 \cdot 2 + 4 = 12$ и $4 \cdot 3 + 12 = 24$. Переход же от стандартного элемента к конкретным осуществляется путем замены переменных $T_h = \hat{f}(\hat{T})$.

Следует заметить, что реализация аппроксимационных алгоритмов в многомерных случаях требует громоздкой технической работы, поскольку итоговые формулы для конечно-элементных матриц могут занимать до нескольких страниц. Здесь существенно помогают средства автоматизации построения алгоритмов на основе компьютерного выполнения аналитических преобразований, которые успешно применяются в ряде пакетов прикладных программ типа FENICS. Поэтому, опуская конструктивные детали построения базисных функций Равьяра-Тома, мы закончим данный пункт сводкой результатов по их аппроксимационным свойствам.

Предварительно заметим, что фактически будут строиться оценки приближений для функций не из $H(\operatorname{div}; \Omega)$, а для несколько более узких пространств

$$W = H(\operatorname{div}; \Omega) \cap (L^s(\Omega))^d, \quad s > 2.$$

Аналогичные пространства $W(T_h)$ определяются на конечных элементах T_h , для которых введем также аппроксимирующие пространства $M_k(T_h)$, удовлетворяющие условиям

$$(P_k(T_h))^d \subseteq M_k(T_h), \quad (P_{k+1}(T_h))^d \not\subseteq M_k(T_h).$$

Утверждение 5.1. Пусть T_h будет аффинный элемент, а через ρ_{T_h} обозначим оператор интерполяции $W_{T_h} \rightarrow M_k(T_h)$. Пусть также π_{T_h} будет L^2 -проектор на $\operatorname{div}(M_k(T_h))$. Тогда для всех $\vec{q} \in W(T_h)$ справедливо соотношение

$$\operatorname{div}(\rho_{T_h} \vec{q}) = \pi_{T_h} \operatorname{div} \vec{q}.$$

Из предыдущего утверждения следует более сильный результат.

Утверждение 5.2. Пусть выполняются условия утверждения 5.1. Тогда существует постоянная c , зависящая только от k и от формы элемента T_h такая, что для $\vec{q} \in RT_k(T_h)$ и $1 \leq m \leq k+1$ справедливо неравенство

$$\|\operatorname{div}(\vec{q} - \rho_{T_h} \vec{q})\|_{0,T_h} \leq ch_{T_h}^m \|\operatorname{div} \vec{q}\|_{m,T_h}.$$

Как видно, данная оценка имеет локальный характер, но выбор степеней свободы на каждом элементе может быть сделан таким образом, чтобы обеспечить непрерывность следов $\vec{q} \cdot \vec{n}$ на интерфейсах конечных элементов во всем пространстве $H(\operatorname{div}; \omega)$. Для этого при сделанном определении $M_k(T_h)$ введем пространство

$$M_k(\Omega, \Omega^h) = \{\vec{q} \in H(\operatorname{div}; \omega), \quad \vec{q}|_{T_h} \in M_k(T_h)\}.$$

Теперь мы можем определить оператор $\prod_h$ глобальной интерполяции из W в $M_k(\Omega, \Omega^h)$, просто полагая

$$(\prod_h \vec{q})_{T_h} = \rho_{T_h}(\vec{q}|_{T_h}).$$

И наконец, сформулируем главный результат данного пункта.

Утверждение 5.3. Пусть Ω^h есть регулярная декомпозиция, или триангуляция области Ω . Тогда существует не зависящая от h константа c такая, что для ошибки глобальной интерполяции для всех $\vec{q} \in W$ справедливы неравенства

$$\begin{aligned} \|\vec{q} - \prod_h \vec{q}\|_{0,\Omega} &\leq ch^{k+1} |q|_{k+1, \Omega}, \\ \|\operatorname{div} (\vec{q} - \prod_h \vec{q})\|_{0,\Omega} &\leq ch^s |\operatorname{div} \vec{q}|_{s,\Omega}, \quad s \leq k+1. \end{aligned}$$

5.4. Разрывные методы Галеркина

В последние десятилетия активное развитие получил оригинальный вариационный метод решения краевых задач, обладающий достаточной общностью и прокладывающий в определенном смысле мостик между методами конечных элементов и конечных объемов. Его достоинствами являются, с одной стороны, применимость к многомерным задачам со сложными конфигурациями расчетных областей и контрастными свойствами материальных сред, а с другой — возможности достижения высоких порядков аппроксимации и точности. Данное направление получило название *разрывные методы Галеркина* с уже широко распространенной англоязычной аббревиатурой DG, от Discontinuity Galerkin. Достаточно содержательный сравнительный анализ работ с различными подходами за 30 лет приведен в обзорной работе Д. Арнольда и Ф. Бреци с соавторами (D. Arnold, F. Brezzi, B. Cockburn, L. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal., v. 39, N 5, 2002, 1749–1779).

5.4.1. Одномерные иммерсионные алгоритмы

Мы в данном параграфе сначала остановимся на изложении эффективного аппроксимационного принципа для одномерного случая: иммерсионных (от английского immersed — дословно “погруженный”) разрывных методов Галеркина РМГ(k) с базисными функциями высоких порядков k , которые строятся с уче-

том свойств решений дифференциальных уравнений, имеющих разрывные коэффициенты.

Такие алгоритмы, во-первых, могут строиться не на адаптивных сетках, привязанных жестко к геометрическим и материальным свойствам исходной задачи (точки разрыва коэффициентов не обязаны быть узлами сетки). Во-вторых, в данном случае строятся легко аппроксимации с переменными не только шагами, но и с меняющимися порядками на различных сеточных интервалах. Мы будем следовать в основном методике работы С. Аджериды и Т. Лина (S. Adjrid, T. Lin. High-order Immersed Discontinuous Galerkin Methods. Internat. J. of Information and System Siences, N 3, 2007, 555–568).

Итак, пусть $u(x)$ есть решение краевой задачи Дирихле для уравнения диффузии

$$\begin{aligned}(p u')' &= f(x), \quad a < x < b, \\ u(a) &= u_a, \quad u(b) = u_b.\end{aligned}\tag{5.83}$$

Без ограничения общности будем считать, что расчетная область $\Omega = (a, b)$ разбита на две подобласти — (a, α) и (α, b) такие, что в точке $\alpha \in (a, b)$ диффузионный коэффициент $p(x)$ претерпевает скачок:

$$p(x) = \begin{cases} p^-, & x \in (a, \alpha), \\ p^+, & x \in (\alpha, b), \end{cases}\tag{5.84}$$

причем в интерфейсной точке α выполняются условия сопряжения

$$[u](\alpha) \equiv u(\alpha^+) - u(\alpha^-) = 0, \quad [p u'](\alpha) = 0.\tag{5.85}$$

Результаты, излагаемые ниже, в дальнейшем без труда могут быть перенесены на случай, когда в Ω имеется несколько точек со скачками коэффициента $p(x)$.

Применение РМГ к решению дифференциальных уравнений второго порядка заключается, прежде всего, в переходе к системе уравнений первого порядка, т. е. к так называемой смешанной постановке:

$$\begin{aligned}q' &= f(x), \quad a < x < b, \\ p u' &= q(x).\end{aligned}\tag{5.86}$$

Введем теперь в расчетной области Ω сетку Ω^h , определяя на (a, b) множество узлов

$$a = x_0 < x_1 < \dots < x_N < x_{N+1} = b,$$

без какой-либо их привязки, или локализации, по отношению к интерфейсной точке α . Далее, умножая первое и второе уравнения (5.86) на некоторые пробные функции $v(x)$ и $w(x)$ соответственно и применяя интегрирование по частям на интервалах

$$I_i = [x_i, x_{i+1}], \quad i = 0, 1, \dots, N,$$

получаем так называемую локальную слабую постановку:

$$\begin{aligned} -q v \Big|_{x_i}^{x_{i+1}} - (q, v')_{I_i} &= (f, v)_{I_i} \equiv \int_{x_i}^{x_{i+1}} f v dx, \\ p u w \Big|_{x_i}^{x_{i+1}} - [p] u w(\alpha) - (p u, w') \Big|_{I_i} &= (q, w) \Big|_{I_i}, \quad \alpha \in I_i, \\ p u w \Big|_{x_i}^{x_{i+1}} - (p u, w') \Big|_{I_i} &= (q, w) \Big|_{I_i}, \quad \alpha \notin I_i. \end{aligned} \quad (5.87)$$

Отсюда получаем *иммерсионный локальный разрывный метод Галеркина k -го порядка* ИЛРМГ(k), если для аппроксимации $q(x)$ определим пространство разрывных кусочно-гладких полиномиальных функций

$$S^{N,k} = \{V | V|_{I_i} \in \mathcal{P}^{(k)}, \quad i = 0, 1, \dots, N\},$$

где $\mathcal{P}^{(k)}$ есть пространство многочленов степени не выше k , а для аппроксимации $u(x)$ введем пространство *иммерсионных конечных элементов* $S_I^{N,k}$, которые представляются полиномами из $\mathcal{P}^{(k)}$ на каждом из интервалов I_i , не содержащем интерфейсную точку α , и кусочно-полиномиальными функциями из $\mathcal{P}^{(k)}$ на интервале, включающем точку α . Построение таких функций мы рассмотрим несколько позже.

Отсюда ИЛРМГ(k) заключается в нахождении пары функций

$$(u^h, q^h) \in X^{N,k} = S_I^{N,k} \times S^{N,k}$$

такой, что для всех $(V, W) \in X^{N,k}$ выполняются равенства

$$\begin{aligned} \hat{q}^h v_h \Big|_{x_i}^{x_{i+1}} + (q^h v_h') \Big|_{I_i} &= (f, v_h)_{I_i}, \\ p \hat{u}^h w_h \Big|_{x_i}^{x_{i+1}} - (p u, w_h') \Big|_{I_i} &= (q^h, w_h) \Big|_{I_i}, \quad \alpha \notin I_i, \\ p \hat{u}^h w_h \Big|_{x_i}^{x_{i+1}} - [p](\alpha)(u^h w_h)(\alpha) - (p u^h, w_h') \Big|_{I_i} &= \\ (q^h, w_h) \Big|_{I_i}, \quad \alpha \in I_i. \end{aligned} \quad (5.88)$$

Здесь величины $\hat{u}^h$ и $\hat{q}^h$ означают следы функций $u^h(x)$ и $q^h(x)$ в узлах сетки. В точках неоднозначности этих разрывных функций следы могут быть определены, в принципе, различным образом, и в рассматриваемых нами методах используется их следующая конкретизация:

$$\hat{u}^h(x_i) = u^h(x_i^-), \quad \hat{q}^h(x_i) = q^h(x_i^+), \quad i = 1, \dots, N. \quad (5.89)$$

При этом на границах расчетной области для следов получаются уравнения

$$\begin{aligned} \hat{u}^h(a) &= u(a), \quad \hat{q}^h(a) = q^h(a^+), \\ \hat{u}^h(b) &= u(b), \quad \hat{q}^h(b) = q^h(b^-) - \frac{\gamma k}{h}(u(b) - u^h(b^-)) \end{aligned} \quad (5.90)$$

Таким образом, получаем систему сеточных алгебраических уравнений порядка $2(N+2)$ для неизвестных $\hat{u}_i^h = \hat{u}^h(x_i)$, $\hat{q}_i^h = \hat{q}^h(x_i)$, $i = 0, 1, \dots, N+1$. Отметим, что последний член в (5.90) с коэффициентом γ введен, с формальной точки зрения, несколько искусственно. В дальнейшем полагается $\gamma = 1$, а введение этого члена связано с идеологией множителей Лагранжа в методах штрафа, см. подробнее упомянутую выше обзорную работу Д. Арнольда с соавторами, а также приведенную в ней обширную специальную литературу. В данной статье дается сравнительный анализ и систематизация существующих подходов к построению разрывных методов Галеркина, основанных фактически на различных определениях следов функций $\hat{u}^h$ и $\hat{q}^h$, отличных от (5.89), (5.90).

Как видно из формулировки РМГ (5.89)–(5.90), данный алгоритм является обобщением методов конечных объемов с возможностями повышения порядков аппроксимации. С другой стороны, он является развитием смешанных методов конечных элементов.

Приведем теперь формулы, необходимые для вычисления интегралов в вариационных уравнениях (5.88), которые выводятся на основе представления базисных функций в “обычных” конечно-элементных пространствах $S^{N,k}$ и в ИКЭ – “интерфейсных” конечно-элементных пространствах $S_I^{N,k}$. В силу их разрывного характера, нам достаточно привести выражения на одном “стандартном” интервале $\hat{I} = (-1, 1)$, полученном после аффинных преобразований каждого из I_i . На интервале $\hat{I}$ иерархи-

ческие “функции формы”, как их иногда называют, имеют следующий вид:

$$\begin{aligned}\bar{\varphi}_0(t) &= (1-t)/2, \quad \bar{\varphi}_1(t) = (1+t)/2, \quad t \in \bar{I}, \\ \bar{\varphi}_j(t) &= c_j(P_j(t) - P_{j-2}(t)), \quad j = 2, \dots, k,\end{aligned}\tag{5.91}$$

где коэффициенты c_j определяются по условиям нормировки. Таким образом мы имеем пространства

$$\bar{\mathcal{P}}^{(k)} = \text{span}\{\bar{\varphi}_j(t), \quad j = 0, 1, \dots, k\}.$$

с помощью которых получаем локальные пространства $\bar{S}_I^{N,k}$ и $\bar{S}^{N,k}$ на интерфейсных и неинтерфейсных сеточных интервалах соответственно. Мы приведем базисные функции только для $\bar{S}_I^{N,k}$, которые точно удовлетворяют условиям сопряжения в точке $\bar{\alpha} \in (-1, 1)$, а гладкие функции из $\bar{S}^{N,k}$ различных порядков, называемые также иерархическим базисом Лобатто, получают из кусочно-гладких как частный случай.

Две линейные функции ИКЭ-базиса имеют следующий вид ($i = 0, 1$):

$$\bar{\varphi}_i(t) = \begin{cases} a_{i,1}t + b_{i,1}, & -1 \leq t \leq \bar{\alpha}, \\ a_{i,2}t + b_{i,2}, & \bar{\alpha} \leq t \leq 1, \end{cases}\tag{5.92}$$

где восемь неизвестных коэффициентов определяются однозначно из такого же количества уравнений

$$\begin{aligned}\bar{\varphi}_0(-1) &= 1, \quad \bar{\varphi}_0(1) = 0, \quad \bar{\varphi}_{-1}(1) = 0, \quad \bar{\varphi}_1(1) = 1, \\ [\bar{\varphi}_i](\bar{\alpha}) &= 0, \quad [p\bar{\varphi}'_i](\bar{\alpha}) = 0, \quad i = 0, 1.\end{aligned}$$

Квадратичная интерфейсная конечно-элементная базисная функция описывается формулой

$$\bar{\varphi}_2(t) = \begin{cases} (1+t)(a_{2,1}t + b_{2,1}), & -1 \leq t \leq \bar{\alpha}, \\ (t-1)(a_{2,2}t + b_{2,2}), & \bar{\alpha} \leq t \leq 1, \end{cases}$$

в которой четыре коэффициента находятся из трех условий сопряжения в интерфейсной точке $\bar{\alpha}$:

$$[\bar{\varphi}_2](\bar{\alpha}) = 0, \quad [p\bar{\varphi}'_2](\bar{\alpha}) = 0, \quad [p\bar{\varphi}''_2](\bar{\alpha}) = 0,$$

а также из условия нормировки $a_{2,2} = 1$. Заметим, что при этом выполняется условие ортогональности в смысле

$$\langle \bar{\varphi}_2, \bar{\varphi}_1 \rangle_{\bar{I}} \equiv \int_{-1}^{\bar{\alpha}} p^- \varphi_2' \varphi_1' dx + \int_{\bar{\alpha}}^1 p^+ \varphi_2' \varphi_1' dx = 0.$$

ИКЭ-функции формы k -х порядков для $k \geq 3$ имеют вид

$$\bar{\varphi}_k(t) = \begin{cases} (1+t) \sum_{j=0}^{k-1} a_{k,j}^{(1)} t^j, & -1 \leq t \leq \bar{\alpha}, \\ (t-1) \sum_{j=0}^{k-1} a_{k,j}^{(2)} t^j, & \bar{\alpha} \leq t \leq 1. \end{cases}$$

Здесь $2k$ неизвестных коэффициента определяются однозначно из $2k - 1$ уравнения

$$\begin{aligned} [\varphi_k](\bar{\alpha}) = 0, \quad [p\varphi_k^{(j)}](\bar{\alpha}) = 0, \quad j = 1, \dots, k+1, \\ \langle \bar{\varphi}_k, \bar{\varphi}_j \rangle_{\bar{I}} = 0, \quad j = 2, 3, \dots, k-1, \end{aligned}$$

а также из какого-либо условия нормировки, например, $a_{k,2}^{(2)} = 1$.

Локальные пространства $S_I^{N,k}$ на интерфейсном элементе формируются с помощью пространств

$$\bar{\mathcal{P}}^{(k)} = \text{span}\{\bar{\varphi}_i, \quad i = 0, 1, \dots, k\}.$$

Если $\bar{\alpha} = \pm 1$ или если диффузионный коэффициент p непрерывен в интерфейсной точке $\bar{\alpha}$, то данные функции переходят в определяемые формулами (5.91) иерархические многочлены Лобатто. Отметим также, что при переходе от t к исходной переменной x иммерсионные базисные функции удовлетворяют условиям $[p\varphi^{(j)}](\alpha) = 0$ для $j \geq 1$.

Примеры иммерсионных (для $\bar{\alpha} = -0.5$) и стандартных базисных функций первого-третьего порядков приведены на Рис. 5.15.

5.4.2. Многомерное квазилинейное эллиптическое уравнение

Разрывные методы Галеркина с аппроксимациями высокого порядка для многомерных краевых задач мы опишем на примере квазилинейной модели диффузии с конвекцией:

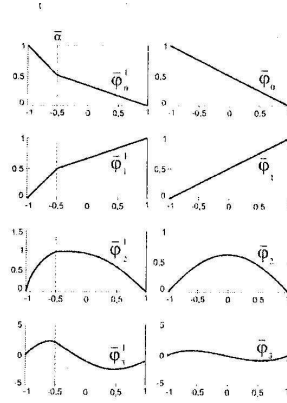


Рис. 5.15. Иммерсионные и стандартные базисные функции 1-го–3-го порядков

$$\begin{aligned} -\nabla \cdot \vec{k}(\vec{r}, u, \nabla u) + \kappa(\vec{r}, u, \nabla u) &= f(\vec{r}), \quad \vec{r} \in \Omega, \\ u &= g_D, \quad \vec{r} \in \Gamma_D, \quad \nabla u \cdot \vec{n} = g_N, \quad \vec{r} \in \Gamma_N, \end{aligned} \quad (5.93)$$

где $\bar{\Omega} = \Omega \cup \Gamma$, $\Gamma = \Gamma_D \cup \Gamma_N$ и $\vec{k}(\vec{r}, \xi, \eta) = \{k_i(\vec{r}, \xi, \eta); i = 1, 2, 3\}$. В уравнении (5.89) мы можем формально предполагать, что в нестационарном случае член с производной по времени перенесен в правую часть, и при этом после пространственной аппроксимации уравнения получаемая система обыкновенных дифференциальных уравнений будет решаться с помощью известных алгоритмов дискретизироваться по времени. Принципы построения данных вычислительных схем мы представим в соответствии с работой казанских математиков Р.З.Даутова и Е.М.Федотова “Разрывный смешанный метод Галеркина без штрафа для квазилинейных эллиптических уравнений второго порядка”, ЖВММФ, т. 53, N 11, 2013, 1791-1803.

Уравнение (5.89) перепишем предварительно в смешанной постановке как систему уравнений первого порядка:

$$-\nabla \cdot \vec{q} + q_0 = f, \quad \vec{q} = \vec{k}(\vec{r}, u, \vec{\sigma}), \quad q_0 = \kappa(\vec{r}, u, \sigma), \quad \vec{\sigma} = \nabla u. \quad (5.94)$$

Предполагая для простоты изложения область Ω полигональной, обозначим через Ω регулярное семейство ее разбиений на тетраэдры (конечные элементы, или объемы) K_ℓ^h с характерными диаметрами $h > 0$ и с граничными поверхностями F_ℓ^h , $\ell =$

$1, \dots, N_t$, где N_t – количество тетраэдров. Обозначим далее через $\vec{n}_\ell$ поле внешних единичных нормалей ℓ -го тетраэдра, через $P_m(K_\ell^h)$ – множество определенных на K_ℓ^h полиномов степени m по совокупности независимых переменных, а через F_h^i – множество “внутренних” граней тетраэдров, т.е. не лежащих на внешней сеточной границе $\Gamma_h^e = \Gamma_D^h \cup \Gamma_N^h$. Пространство приближенных решений u^h при заданном $m \geq 1$ определим как

$$V_h = \{v \in L_\infty(\Omega^h) : v|_{K_\ell^h} \in P_m(K_\ell^h), v|_{\Gamma_D^h} = 0, K_\ell^h \in \Omega^h\}.$$

Аналогично введем пространство для аппроксимации q_0 и векторных компонент $\vec{q}, \vec{\sigma}$:

$$W_h = \{w \in L_\infty(\Omega^h) : w|_{K_\ell^h} \in P_m(K_\ell^h), K_\ell^h \in \Omega^h\}.$$

Очевидно, что функции из этих пространств не обязаны быть непрерывными (на множестве F_h^i они являются двузначными) и что $V_h \subset W_h$, $\nabla_h V_h \subset W_h^3$, где оператор ∇_h определяется через ∇ поэлементно. Для однозначного определения предельных (односторонних) значений функций, терпящих разрыв на F_h^i , введем векторное кусочно-постоянное поле $\vec{p}$ единичных нормалей на $F_h^i \cup \Gamma$, которое назовем потоком. Поле $\vec{p}$ определяется произвольным образом, но с единственным условием: для любого элемента K_ℓ^h должна существовать выходящая из него нормаль $\vec{n}_f$, т.е. $(\vec{p} \cdot \vec{n}_f)^-|_{K_\ell^h} = 0$. Здесь и далее через $v^\pm = \max\{\pm v, 0\}$ обозначается положительная и отрицательная часть функций v , так что $v = v^+ - v^-$. Определим также такие понятия, как одностороннее значение и скачок функции w в точке $\vec{r} \in F_h^i$ вдоль нормали $\vec{p}(\vec{r})$:

$$w_{\pm\vec{p}}(\vec{r}) = \lim_{\eta \rightarrow 0} [w]_{\vec{p}(\vec{r} = w_{+\vec{p}})(\vec{r}) - w_{-\vec{p}}(\vec{r})},$$

причем на внешней границе Ω^h скачок полагаем нулевым, т.е. для $\vec{r} \in \Gamma_h^e$ $[w]_{\vec{p}(\vec{r})}|_{\Gamma_h^e} = 0$.

Для единственного определения значения (следа) функции $u_h \in V_h$ на общей грани F двух контактирующих тетраэдров K_h и K'_h принимается следующий противопотоковый способ:

$$u_h|_F = u_{h,-\vec{p}} = (\vec{p} \cdot \vec{n}_K)^+ u_h|_K + (\vec{p} \cdot \vec{n}_K)^- u_h|_{K'}.$$

Для векторных функций $\vec{q}_h \in W_h^3$, наоборот, принимается значение по потоку:

$$\vec{q}_h|_F = \vec{q}_{h,+ \vec{p}} = (\vec{p} \cdot \vec{n}_K)^+ \vec{q}_h|_{K'} + (\vec{p} \cdot \vec{n}_K)^- \vec{q}_h|_K.$$

Под разрывным смешанным методом Галеркина для приближенного решения задачи (5.93) на сетке $\bar{\Omega}^h = \Omega^h \cup \Gamma_h$ понимается следующая дискретная задача: найти функцию $u_h \in V_h$, а также вектор-функции $\vec{q}_h, \vec{\sigma}_h \in W_h^3$ такие, что для выбранного поля $\vec{p}$ выполняется равенство

$$\alpha_h(u_h, v_h) = f_h(v_h) \quad \forall v_h \in V_h, \quad (5.95)$$

где билинейная и линейная формы определяются следующим образом:

$$\begin{aligned} a_h(u_h, v_h) &= \int_{\Omega^h} q_{0,h} v_h d\vec{r} + \sum_{K \in \Omega^h} \int_K u_h \nabla \cdot \vec{w}_h d\vec{r} + \int_{\Gamma_h} \vec{q}_{h,+ \vec{p}} \vec{p} [v_h]_{\vec{p}} ds, \\ f_h(v_h) &= \int_{\Omega^h} f v_h d\vec{r}, \end{aligned} \quad (5.96)$$

а функции $\vec{\sigma}_h = \vec{\sigma}_h(u_h)$, $\vec{q}_h = \vec{q}_h(u_h)$ и $q_{0,h} = q_{0,h}(u_h)$ определяются по u_h из соотношений

$$\begin{aligned} \int_{\Omega_h} \vec{\sigma}_h \cdot \vec{W}_h d\vec{r} + \sum_{K \in \Omega^h} \int_K u_h \nabla \cdot \vec{W}_h d\vec{r} + \int_{\Gamma_h} u_h, -\vec{p} [\vec{W}_h]_{\vec{p}} \vec{p} ds &= 0, \quad \forall \vec{w}_h \in W_h^3, \\ \int_{\Omega_h} [(\vec{q}_h - \vec{k}(\vec{r}, u_h, \vec{\sigma}_h)) \cdot \vec{w}_h + \kappa(\vec{r}, u_h, \vec{\sigma}_h) w_{0,h}] d\vec{r} &= 0, \\ \forall \vec{w}_h \in W_h^3, \quad w_{0,h} \in W_h. \end{aligned} \quad (5.97)$$

Уравнения (5.95)–(5.96) определяют систему алгебраических (в общем случае нелинейных) уравнений для определения искоемых приближенных решений $u_h, \vec{q}_h$ и $\vec{\sigma}_h$. Мы не будем вдаваться в детали алгоритмов их решения, а также нетривиального обоснования данного варианта РМГ, предложенного Р.З.Даутовым и Е.М.Федотовым. Отметим только, что если выполнены достаточные условия монотонности операторов, а также существования и единственности обобщенного решения задачи (5.89) в пространстве Соболева $H_0^1(\Omega)$, и если к тому же $u \in H^{m+1}(\Omega)$, $q \in [H^m(\Omega)]^3$, то погрешности построенных приближенных решений в соответствующих нормах суть величины порядка $O(h^m)$.

5.4.3. Нестационарная диффузионно-конвективная задача

В данном пункте мы изложим, следуя работе R.Z.Dautov, E.M.Fedotov. Discontinuous Galerkin schemes for nonstationary convection-diffusion equation. Proc. 11-th Internat. Conf. on “Mesh methods for boundary value problems and applications”. IOP Publ. Conf. Series: Materials Science and Engin., v. 158, 2016, 012027, построение гибридизированного разрывного метода Галеркина для аппроксимации многомерной параболической начально-краевой задачи, описанной выше в п. 1.3.4:

$$\begin{aligned} c\partial u/\partial t - \nabla \cdot (\varepsilon A \nabla u - \vec{b}u) + au &= f(\vec{x}, t), \\ (x, t) \in Q &= \Omega \times (0, t_e), \quad t_e > 0, \\ u(\vec{x}, 0) &= u^0(\vec{x}), \quad \vec{x} \in \Omega. \end{aligned} \quad (5.98)$$

Относительно функциональных свойств исходных данных приведенной задачи мы предполагаем, что они удовлетворяют условиям (1.51)–(1.57) из п. 1.3.4, которые мы ради краткости здесь не повторяем. Для построения дискретных схем уравнение (5.94) перепишем в форме системы уравнений первого порядка

$$\vec{\sigma} = \nabla u, \quad \vec{q} = \varepsilon A \sigma, \quad \partial u/\partial t - \nabla \cdot (\vec{q} - \vec{b}u) + au = f, \quad (5.99)$$

где векторные величины $\vec{q}, \vec{\sigma}$ имеют смысл потока и градиента решения.

Приведенная постановка обладает тремя отличительными особенностями. Во-первых, введение матрицы A позволяет рассматривать анизотропные диффузионные процессы, что является актуальным для многих приложений. Во-вторых, явное выделение числового параметра ε дает возможность при $\varepsilon \rightarrow 0$ конструировать эффективные методы решения сингулярно возмущенных уравнений с преобладающей конвекцией и пограничными или внутренними слоями. В третьих, уравнение (5.98) представлено в дивергентной форме, с целью удобства построения консервативных вычислительных схем.

Пусть $\Omega_h = \cup T_h$ обозначает сеточную расчетную область, или триангуляцию исходной полигональной расчетной d -мерной ($d = 2, 3$) области Ω с границей Γ в совокупность конечных элементов T_h , у которых параметр h означает максимальный диаметр. Триангуляцию Ω_h будем считать регулярной в том смысле

ле, что каждый элемент T_h может быть получен аффинным преобразованием из некоторого базового, или стандартного элемента $\hat{T} \subset \mathcal{R}^d$ (единичный треугольник или квадрат для $d = 2$, тетраэдр, куб или призма для $d = 3$). Кроме того, считаем *триангуляцию конформной* (грань $\mathcal{F}_h$ каждого элемента или является гранью соседнего элемента, или принадлежит границе, т.е. $\mathcal{F}_h = \mathcal{F}_h^\Gamma \in \Gamma$) и адаптивной, или согласованной с границей (все вершины и ребра Γ принадлежат Γ_h – границе $\bar{\Omega}_h = \Omega^h \cup \Gamma^h$). Пусть также $\bar{\omega}_\tau = \{t_n, n = 0, \dots, N_e\}$ означает множество узлов дискретизации по времени $t \in [0, t]$ $\omega_\tau = \bar{\omega}_\tau \setminus \{t_{N_e}\}$. Предполагаем, что разбиение $\Omega_\tau = \{\tau_n = t_{n+1} - t_n\}^{N_e-1}$ является регулярным или квазиравномерным, т.е. при $\tau \rightarrow 0$ для всех n имеем $\tau_n = 0(\tau)$.

Скалярные произведения в пространствах $L_2(\Omega)$ и $L_2(\Omega)^d$, а также нормы и полунормы скалярных или векторных функций $\vec{v} = (v_1, \dots, v_d)$ из соответствующих соболевских пространств порядка s ($H^s(\Omega)$ или $H^s(\Omega)^d$), обозначаются, как и в п. 1.3.4, через $(\cdot, \cdot)_\Omega$, $\|\cdot\|_{s,\Omega}$ и $|\cdot|_{s,\Omega}$. Согласно (1.60) определяются также пространства $L_p(a, b; [H^s(\Omega)]^d)$ и $H^p(a, b; [H^s(\Omega)]^d)$, $1 \leq p \leq \infty$, измеримых функций $u(t) \in \{(a, b) \rightarrow [H^s(\Omega)]^d\}$.

Для дискретизированных, или конечно-элементных, функций скалярные произведения, нормы и полунормы вводятся аддитивным образом, т.е. как суммы соответствующих величин по всем конечным элементам пространственной сетки Ω^h или по “глобальной” дискретизации $\Omega_{h,\tau} = \Omega_h \times \Omega_\tau$.

Пусть $P_m(\Omega)$ (или $P_m(T_h)$) обозначает пространство полиномов степени не выше $m \geq 0$, определенных на Ω (или на T_h). На стандартном элементе $\hat{T}$ с границей $\hat{\Gamma}$ введем конечно-мерные пространства векторных и скалярных функций

$$\begin{aligned} P_0(\hat{T}) &\subseteq V(\hat{T}) \subset H^1(\hat{T}), \quad [P_0(\hat{T})]^d \subseteq W(\hat{T}) \subset [H^1(\hat{T})]^d, \\ p_0(\hat{\mathcal{F}}) &\subseteq \Lambda(\hat{\mathcal{F}}) \subset L_2(\hat{\mathcal{F}}), \quad \hat{\mathcal{F}} \subset \hat{\Gamma}. \end{aligned} \tag{5.100}$$

Отметим, что скалярное и векторное пространства $V(\hat{T}), W(\hat{T})$ будут использоваться для определенных в элементе функций u и $\vec{\sigma}, \vec{q}$, а пространство $\Lambda(\hat{\Gamma})$ – для следов скалярной функции u на границе элемента $\hat{\Gamma}$.

С помощью этих пространств вводятся аналогичные пространства функций, определенных на произвольном конечном элементе T_h с границей Γ_h , состоящей из граней $\mathcal{F}_h$. Предпола-

гая, что для $\vec{x} \in T_h$ и $\hat{\vec{x}} \in \hat{T}$ преобразование

$$\vec{x} = \mathcal{F}_T(\hat{\vec{x}}) = B_T \hat{\vec{x}} + c_T$$

является аффинным и обратимым, мы полагаем:

$$\begin{aligned} V(T_h) &= \{v = \hat{v} \circ \mathcal{F}_T^{-1}, \hat{v} \in (\hat{T})\}, \quad \Lambda_{T_h}(\mathcal{F}_h) = \{\lambda = \hat{\lambda} \circ \mathcal{F}_T^{-1}, \hat{\lambda} \in \Lambda(\hat{\mathcal{F}})\}, \\ W(T_h) &= \{\mu = |B_T|^{-1} B_T(\hat{\mu} \circ \mathcal{F}_T^{-1}), \hat{\mu} \in W(\hat{T})\}, \\ \Lambda(\mathcal{F}_h) &= \Lambda_{T_h}(\mathcal{F}_h), \quad \mathcal{F}_h \in \Gamma_h \subset \hat{T}_h. \end{aligned} \quad (5.101)$$

Обозначая через $I_\Omega \in P_0(\Omega)$ функцию, тождественно равную единице в Ω , определим еще пространства

$$\Lambda(\Gamma_h) = \{\lambda \in L_2(\Gamma_h) : \lambda|_{\mathcal{F}_h} \in \Lambda(\mathcal{F}_h), \mathcal{F}_h \in \Gamma_h\}, \quad (5.102)$$

$$\begin{aligned} \bar{V}(T_h) &= V(T_h) \times \Lambda(\Gamma_h), \\ \bar{V}_0(T_h) &= \{(v, \lambda) \in \bar{V}(T_h) : (v, \lambda) = c(I_{T_h}, I_{\Gamma_h}), c \in \mathcal{R}\}, \end{aligned}$$

где $|B_T| = \det|B_T|$. Относительно введенных в (5.100)–(5.102) пространств предполагается выполнение следующих ограничений:

H_1 : для всех $(v, \lambda) \in \bar{V}(T_h)$ из соотношения

$$(\lambda, \vec{\tau} \cdot \vec{n}_{T_h})_{\Gamma_h} - (v, \nabla \cdot \vec{\tau})_{T_h} = 0, \quad \vec{\tau} \in \bar{W}(T_h),$$

следует, что $(v, \lambda) \in V_0(T_h)$;

$$H_2 : \nabla \cdot \bar{W}(T_h) = \{\nabla \cdot \vec{w} : \vec{w} \in \bar{W}(T_h)\} \subseteq \bar{V}(T_h);$$

$$H_3 : \Lambda(\mathcal{F}_h) = \{\vec{w} \cdot \vec{n}_{T_h} : \vec{w} \in \bar{W}(T_h)\} \text{ для каждой грани } \mathcal{F}_h \subset \Gamma_h;$$

$$H_4 : \nabla V(T_h) = \{\nabla v : v \in V(T_h)\} \subset [V(T_h)]^d.$$

Далее мы определяем конечно-элементные пространства

$$\begin{aligned} V_h &= \{v \in L_2(\Omega) : v|_{T_h} \in V(T_h), T_h \in \Omega_h\}, \\ W_h &= \{\vec{w} \in [L_2(\Omega)]^d : \vec{w}|_{T_h} \in W(T_h), T_h \in \Omega_h\}, \\ \Lambda_h &= \{\lambda \in L_2(\Gamma_h) : \lambda|_{\mathcal{F}_h} \in \Lambda(\mathcal{F}_h), \mathcal{F}_h \in \Gamma_h; \lambda|_{\mathcal{F}_h^\Gamma} = 0\}, \end{aligned} \quad (5.103)$$

на основе которых, в свою очередь, вводим конечно-мерные пространства $V_{h,\tau}, W_{h,\tau}, \Lambda_{h,\tau}$ полиномов степени $r \geq 0$ от переменной $t \in \Delta \in \Omega_\tau$, имеющих свои значения в V_h, W_h, Λ_h соответственно. При этом пространства $V_{h,\tau}$ используются для аппроксимации скалярного решения u , $W_{h,\tau}$ – для векторных функций

$\vec{\sigma}$ и $\vec{q}$, а $\Lambda_{h,\tau}$ – для аппроксимации следов решения на гранях конечных элементов.

Семейство гибридизированных разрывных методов Галеркина определяется следующим образом: найти решение $(\vec{\sigma}_h, \vec{q}_h, u_h, \lambda_h) \in W_{h,\tau} \times W_{h,\tau} \times V_{h,\tau} \times \Lambda_{h,\tau}$, удовлетворяющее следующим вариационным равенствам:

$$\begin{aligned} (\vec{\sigma}_h, \vec{w})_{\Omega_h \times \Omega_\tau} + (u_h, \nabla \cdot \vec{w})_{\Omega_h \times \Omega_\tau} - (\lambda_h, \vec{w} \cdot \vec{n}_{T_h})_{\Gamma_h \times \Omega_\tau} &= 0 \quad \forall \vec{w} \in W_{h,\tau}, \\ (\vec{q}_h - \varepsilon A \vec{\sigma}_h, \vec{w})|_{\Gamma_h \times \Omega_\tau} &= 0 \quad \forall \vec{w} \in W_{h,\tau}, \end{aligned} \quad (5.104)$$

$$\begin{aligned} &-(u_h, \partial v / \partial t)_{\Omega_h \times \Omega_t} + \sum_{t \in \omega_\tau \setminus \{0\}} (\bar{u}_h, \bar{v} - v)_{\Omega_h \times \Omega_t} + \\ &+ (\bar{u}_h, \bar{v})_{\Omega_h} \Big|_{t=t_e} - (u_0, v|_{t=0})_{T_h} + U_{T_h}((u_h, \vec{q}_h, \lambda_h), v), I)_{\Omega_\tau} = \\ &= (f, v)_{\Omega_h \times \Omega_\tau} \quad \forall v \in V_{h,\tau}, \quad T_h \in \Omega_h, \end{aligned} \quad (5.105)$$

$$(\vec{q}_h \cdot \vec{n}, \mu)_{\Gamma_{h,\tau}} - (\vec{b}_d(\vec{u}_h, \mu) \cdot \vec{n}, \mu)_{\Gamma_{h,\tau}} = 0 \quad \forall \mu \in \Lambda_{h,\tau}. \quad (5.106)$$

Здесь $v = v(t+0)$ и $\bar{v} = v(t-0)$ означают предельные односторонние значения функции v в точках $t \in \omega_\tau$, а также введены обозначения $\vec{u}_h = (u_h, \lambda_h)$ и

$$\begin{aligned} U_{T_h}((u, \vec{q}, \lambda), v) &= (-\nabla \cdot \vec{q} + \alpha u, v)_{T_h} - \vec{b}u, \nabla v)_{T_h} + \\ &+ (\vec{b}_\alpha(\vec{u}_h) \cdot \vec{n}_{T_h}, \mu)_{\Gamma_h} \quad \forall v \in V_h, \\ \vec{b}_\alpha(\vec{u}_h)|_{\Gamma_h} &= 0.5(u_h + \lambda_h)_{\Gamma_h} + \alpha(u_h - \lambda_h)_{\Gamma_h} n_{T_h}, \end{aligned}$$

где $\alpha = \alpha(\vec{x})$ есть некоторый стабилизирующий параметр, который является скалярной функцией на Γ_h . Относительно него предполагается выполнение неравенств

$$\alpha_0 |\vec{b} \cdot \vec{n}_{T_h}| \leq \alpha(\vec{x}) \leq C \quad \forall \vec{x} \in T_h, \quad T_h \in \Omega_h,$$

где α и C суть некоторые не зависящие от h и τ постоянные.

Нетрудно видеть, что если мы положим $\alpha|_{\Gamma_h} = 0.5|\vec{b} \cdot \vec{n}_{T_h}|$, то будет выполняться соотношение

$$\vec{b}_\alpha(\vec{u}_h) \cdot \vec{n}_{T_h}|_{\Gamma_h} = (\vec{b} \cdot \vec{n}_{T_h})^+ u_h - (\vec{b} \cdot \vec{n}_{T_h})^- \lambda_h|_{\Gamma_h},$$

где $v^\pm = \max\{0, \pm v\}$ суть положительная и отрицательная части v . В этом случае равенство (*) означает аппроксимацию “по потоку” диффузионно-конвективного уравнения.

Если $r = 0$, то приведенная аппроксимация по времени является неявной схемой Эйлера, а при $r > 0$ – это $(r + 2)$ -слойные неявные вычислительные схемы, широко употребляемые при численном интегрировании систем обыкновенных дифференциальных уравнений, см. [192]– [194].

Рассмотренный в данном пункте подход основан на общем принципе единообразной дискретизации по временной и пространственным переменным с помощью разрывного метода Галеркина. Как видно из соотношений (5.104)–(5.106), получаемый алгоритм является достаточно громоздким, поскольку он содержит много неизвестных функций (однако это является неизбежной платой за высокие аппроксимационные качества метода). Уменьшить число степеней свободы дискретизированной задачи можно, исключив из полученной вычислительной схемы величины $\vec{\sigma}_h, \vec{q}_h$, в результате чего получим алгебраическую систему относительно скалярных величин u_h, λ_h .

В данном пункте мы ограничиваемся изложением основных идей построения разрывных методов Галеркина высокого порядка точности, отсылая заинтересованного читателя за алгоритмическими деталями и вопросами теоретического обоснования к упомянутым работам Р.З.Даутова и Е.М.Федотова, которыми показано, что при выполнении приведенных выше условий $H_1 - H_4$ и выборе соответствующих полиномиальных базисных функций описанные вычислительные схемы являются устойчивыми и обладают скоростью сходимости (в специальных нормах, на конкретизации которых мы не останавливаемся) с порядком $O(\tau^{r+1}(1 + \varepsilon^{-1/2}) + h^m)$.

В заключение данного параграфа отметим, что разрывные методы Галеркина в настоящее время представляют собой активно развиваемое направление, перспективное с точки зрения параллельных вычислительных технологий. Для различных типов задач российскими и зарубежными показано, как теоретически, так и экспериментально, высокая точность для численных решений и потоков.

Применение РМГ динамически расширяется для все новых классов задач, в том числе сингулярных, многомерных, нелинейных, нестационарных и т. д. Активное теоретическое и прикладное развитие разрывных методов Галеркина отражено в многочисленных журнальных статьях, но пока еще ждет своего обобщения и систематизации в монографической литературе.

В этом плане исключение составляет изданная в 2008 г. книга Б.Ривiere [149], хотя в некотором смысле она успела устареть, так как после ее выхода уже вышел ряд журнальных статей с новыми оригинальными идеями.

5.5. Общие замечания о построении матриц МКЭ

Резюмируя проведенные выше рассмотрения, можно сказать, что конкретный метод конечных элементов для решения поставленной краевой задачи полностью определяется способом разбиения расчетной области на конечные элементы Ω_k^h , видом искомой аппроксимации искомого решения и совокупностью функционалов, характеризующих параметры выбранного приближения. Совокупность этих факторов и порождает конечномерное пространство функций V^h .

Очевидно, что все МКЭ можно разделить на однородные и неоднородные по степени «одинаковости» алгоритмов в области. К однородным отнесем методы с одинаковой формой конечных элементов (например, треугольники), с одинаковой на каждом элементе полиномиальной формой интерполянта и с однотипным набором функционалов (значений функций и/или их производных в заданном количестве точек). В принципе же представления приближенных решений на каждом элементе могут быть никак не связаны между собой, за исключением, возможно, условий непрерывности на смежных ребрах.

Последнее означает естественное требование, что определяемые на ребрах искомые величины должны быть общими для смежных элементарных ячеек, то же самое относится и к узловым неизвестным.

Остановимся теперь на некоторых конструктивных вопросах формирования локальных матриц МКЭ. Пусть $\mathcal{P}_k$ означает конечномерное подпространство (размерности N_k) функций, определенных на элементе Ω_k^h . Обозначим через $Q_k = \{q_1, \dots, q_{N_k}\}$ базис в пространстве функционалов $\mathcal{P}'_k$, где N_k — число степеней свободы конечного элемента Ω_k^h . Пусть каждая величина $q_l^{(k)}$ соответствует значению какой-то производной $D_j u(P_i)$ порядка $|D_j|$ в точке $P_i \in \bar{\Omega}_k^h$, где $u \in \mathcal{P}_k$, $i = 1, \dots, M_k$, $j = 0, \dots, \alpha_i^{(k)} - 1$, $\alpha_i^{(k)}$ — кратность каждого узла P_i , а M_k — их количество в $\bar{\Omega}_k^h$ (при этом каждый номер l соответствует паре индексов i, j , и мы можем записать $q_l^{(k)} = q_{i,j}^{(k)}$). Тогда совокупность линейно

независимых функций $\Phi_k = \{\varphi_l^{(k)}, i = 1, \dots, N_k\}$, линейная оболочка которых составляет $\mathcal{P}_k$, является двойственным базисом к Q_k , если каждая базисная функция $\varphi_l^{(k)} = \varphi_{i,j}^{(k)}$ соответствует величине $q_l^{(k)} = q_{i,j}^{(k)}$ в смысле удовлетворения условиям $D_j \varphi_{i,j}^{(k)}(P_i) = \delta_{i,j}$. Поскольку элементы локальных матриц МКЭ определяются через скалярные произведения базисных функций, возникает не только теоретическая проблема их существования, но и алгоритмическая задача нахождения $\varphi_l^{(k)}$ по заданным $\mathcal{P}_k$ и Q_k .

В одномерном случае на эти вопросы находятся исчерпывающие ответы с помощью теории интерполирования. Пусть конечный элемент Ω_k^h есть отрезок $[x_k, x_{k+1}]$, на котором задано m_k различных точек $x_i^{(k)}$ и $N_k \geq m_k$ линейно независимых функций $\psi_j^{(k)}(x)$. Определим пространство $\mathcal{P}_k$ как совокупность всевозможных линейных комбинаций

$$\psi_k(x) = c_1^{(k)} \psi_1^{(k)} + \dots + c_{N_k}^{(k)} \psi_{N_k}^{(k)} \quad (5.107)$$

с некоторыми коэффициентами $c_j^{(k)}$.

Теперь будем искать из всевозможных комбинаций (5.107) такую, которая удовлетворяет условиям обобщенной задачи интерполирования, т. е. в каждой из точек $x_i^{(k)}, i = 1, \dots, m_k$, функция $\Psi_k(x)$ и ее производные до $(\alpha_i^{(k)} - 1)$ -го порядка включительно принимают заданные значения (число $\alpha_i^{(k)}$ называется *кратностью интерполяционного узла* $x_i^{(k)}$). Отсюда для N_k коэффициентов $c_j^{(k)}$ получаем систему линейных алгебраических уравнений

$$\begin{aligned} \psi_k(x_i^{(k)}) = q_j, \quad \psi'_k(x_i^{(k)}) = q_j^{(1)}, \dots, \psi_k^{(\alpha_i^{(k)}-1)}(x_i^{(k)}) = q_j^{(\alpha_i^{(k)}-1)}, \\ i = 1, \dots, m_k, \end{aligned} \quad (5.108)$$

где $q_i^{(l)}$ — заданные числа, а порядок системы равен

$$M_k = \alpha_1^{(k)} + \dots + \alpha_m^{(k)}.$$

Для существования единственного решения системы (5.108) требуется, чтобы выполнялось равенство $N_k = M_k$ и квадратная матрица

$$B = \begin{bmatrix} \psi_1(x_1) & \psi_2(x_1) & \dots & \psi_N(x_1) \\ \psi'_1(x_1) & \psi'_2(x_1) & \dots & \psi'_N(x_1) \\ \dots & \dots & \dots & \dots \\ \psi_1^{(\alpha_1-1)}(x_1) & \psi_2^{(\alpha_1-1)}(x_1) & \dots & \psi_N^{(\alpha_1-1)}(x_1) \\ \psi_1(x_2) & \psi_2(x_2) & \dots & \psi_N(x_2) \\ \dots & \dots & \dots & \dots \\ \psi_1^{(\alpha_m-1)}(x_m) & \psi_2^{(\alpha_m-1)}(x_m) & \dots & \psi_N^{(\alpha_m-1)}(x_m) \end{bmatrix} \quad (5.109)$$

была невырождена (здесь и ниже мы для краткости опускаем индекс k).

Обозначая через $c = \{c_j\}$, $q = \{q_i^{(l)}\}$ векторы порядка N , систему (5.108) запишем в матричном виде

$$Bc = q. \quad (5.110)$$

Установим некоторые факты о существовании и единственности решения системы (5.110), см. подробнее [82].

Определение 5.3. Совокупность функций $\psi_j(x)$, $j=1, 2, \dots, m$, образует на отрезке $[a, b]$ систему Чебышева, если никакая их линейная комбинация

$$c_1\psi_1(x) + c_2\psi_2(x) + \dots + c_m\psi_m(x)$$

не обращается в ноль в точках $x_j \in [a, b]$, $j = 1, 2, \dots, m$, т.е. не имеет на $[a, b]$ m различных корней.

Теорема 5.1. Если функции $\psi_1(x), \psi_2(x), \dots, \psi_m(x)$ m раз дифференцируемы на отрезке $[a, b]$ и вронскиан

$$W[\psi_1, \dots, \psi_k] = \begin{vmatrix} \psi_1(x) & \psi_2(x) & \dots & \psi_k(x) \\ \psi'_1(x) & \psi'_2(x) & \dots & \psi'_k(x) \\ \dots & \dots & \dots & \dots \\ \psi_1^{(k)}(x) & \psi_2^{(k)}(x) & \dots & \psi_k^{(k)}(x) \end{vmatrix}$$

не равен нулю на $[a, b]$ при всех $k = 1, \dots, m$, то эти функции образуют систему Чебышева.

Подчеркнем, что для того, чтобы функции $\psi_j(x)$ образовывали систему Чебышева, их линейная независимость является необходимой, но не достаточной. Простейшую систему Чебышева составляют мономы $\psi_j(x) = x^j$, $j = 0, 1, \dots$. Из теоремы 5.1

следует, в частности, существование и единственность интерполяционного многочлена Лагранжа, соответствующего случаю $\alpha_1 = \dots = \alpha_m = 1$, когда все узлы x_j являются некротными, т. е. простыми.

В общем случае с кратными узлами и степенными функциями $\psi_j(x) = x^j$ система уравнений (5.108) определяет единственным образом *интерполяционный многочлен Эрмита* порядка $N_k - 1$. Для его построения необходимо найти вектор c из решения системы (5.110):

$$c = B^{-1}q.$$

Пусть теперь требуется найти коэффициенты $b_l^{(i)}$, $l = 1, \dots, N$, многочленного представления базисной функции МКЭ:

$$\varphi_i(x) = b_1^{(i)} + b_2^{(i)}x + \dots + b_N^{(i)}x^N, \quad (5.111)$$

соответствующей решению системы (5.110) с правой частью $q = e_i$ (e_i — вектор-орд с единственной ненулевой i -й компонентой, равной единице).

Очевидно, что вектор $b^{(i)} = \{b_l^{(i)}\}$ есть просто i -й столбец обратной матрицы $B^{-1} = \{b^{(i)}, i = 1, \dots, N\}$. Таким образом, однократное обращение матрицы B определяет все базисные функции подпространства $\mathcal{P}_k$, а локальные матрицы МКЭ находят путем вычисления соответствующих билинейных форм:

$$a(\varphi_i, \varphi_j) = a(b_1^{(i)} + \dots + b_N^{(i)}x^N, b_1^{(j)} + b_2^{(j)}x + \dots + b_N^{(j)}x^N). \quad (5.112)$$

Аналогичные рассуждения легко провести и для многомерных задач, которые можно наглядно проиллюстрировать на случае двух независимых переменных. Конечные элементы Ω_k^h при этом являются многоугольниками, а конечномерное подпространство функций $\mathcal{P}_k$ можно определить как линейную оболочку многочленов суммарного p -го порядка от двух переменных:

$$\begin{aligned} \psi_k(x, y) &= c_1^{(k)} + c_2^{(k)}x + c_3^{(k)}y + \dots + \sum_{j=0}^p c_{pj}^{(k)}x^{p-j}y^j, \\ c_{pj}^{(k)} &= \frac{p(p+1)}{2} + j + 1. \end{aligned} \quad (5.113)$$

Размерность такого подпространства равна $N = \frac{(p+1)(p+2)}{2}$. Базис $Q_k = \{q_1, \dots, q_N\}$ в пространстве функционалов $\mathcal{P}'_k$ в данном

случае имеет компонентами смешанные производные различных порядков или в вершинах, или на ребрах, или во внутренних точках элементарного многоугольника. Отметим, что в некоторых алгоритмах размерность N «волевым образом» занижается путем выбрасывания некоторых членов полиномиального представления (5.113), с целью упрощения структур матриц МКЭ.

Обобщенная задача двумерной интерполяции также алгебраически описывается системой уравнений (5.110), и ее порядок (число степеней свободы МКЭ) N равен размерности подпространства $\mathcal{P}_k$. Базисные функции $\varphi_i^{(k)}(x, y)$ этого подпространства описываются формулой, аналогичной (5.113):

$$\varphi_i^{(k)}(x, y) = b_{1,i}^{(k)} + b_{2,i}^{(k)}x + b_{3,i}^{(k)}y + \dots + \sum_{j=0}^p b_{p,j,i}^{(k)}x^{p-j}y^j, \quad (5.114)$$

а вектор $b_i^{(k)}$, составленный из коэффициентов $b_{l,i}^{(k)}$, $l = 1, \dots, N_k$, есть i -й столбец обратной матрицы B_k^{-1} , т. е.

$$b_i^{(k)} = B_k^{-1}e_i.$$

Элементы матриц МКЭ находятся путем вычисления билинейных форм от базисных функций аналогично (5.112).

Напомним, что на четырехугольных элементах мы рассматривали биполиномиальные аппроксимации, в которых порядок по каждой из переменных x, y равен одинаковому числу q . В этом случае вместо (5.113) можно записать

$$\psi_k(x, y) = \sum_{j=0}^q c_{j,x}^{(k)}x^j \sum_{j=0}^q c_{j,y}^{(k)}y^j, \quad (5.115)$$

и порядок соответствующего подпространства $\mathcal{P}_k$ равен $N = 2(q+1)$, а двумерные базисные функции в соответствии с (5.115) представляются произведениями одномерных, что означает соответствующую блочную структуру матрицы B_k .

5.6. Многомасштабные методы и технологии

В этом параграфе мы рассмотрим специальные подходы к решению задач, характеризующихся резко меняющимися пространственными и/или временными масштабами изучаемых процессов и явлений. Примерами таких трудных для традиционных

численных методов проблем могут быть моделирование физических полей или в композитных материалах с мелкими неоднородными включениями, или в активной зоне ядерного реактора с большим количеством топливных элементов, или в нелинейных задачах с локальными “обострениями” решений. При этом характерным для решаемых уравнений является наличие контрастных скачков в значениях коэффициентов (диффузии, теплопроводности и др.), что может привести к пространственным неоднородностям решений, масштаб которых мал не только по отношению к размеру расчетной области, но даже в сравнении с приемлемой величиной шага сетки h .

Возможным подходом к преодолению этих трудностей является *гомогенизация* среды, основанная на аналитическом *осреднении коэффициентов*, позволяющем получить приемлемое решение исходной задачи, см., например, книги Г. И. Марчука и А. А. Самарского [121], [163].

В последние десятилетия получили активное развитие численные многомасштабные методы (multi-scale), использующие сетки двух или большего числа уровней. По-видимому, одни из первых независимых опытов в этом направлении получены авторами статей Brandt A. Multiscale adaptive solutions to boundary value problems. Math. Comp., 1977, Vol.31, P. 333–390 и Козлов С. М. Осреднение дифференциальных операторов с почти периодическими быстроосциллирующими решениями. Матем. сборник, 1978, Т. 107, № 2, С. 199–217. В книге Р. П. Федоренко [183] аналогичная идея описана под названием *метод суперэлементов*, развитие и обоснование которого дано в работе М. П. Галанина и С. А. Лазарева “Метод конечных суперэлементов и его применение для решения задач науки и техники”. Матем. моделирование, 2013, Т. 26, № 6, с. 32–40. Обзор работ по данному направлению имеется в [203], [204]. Отметим еще, что многомасштабный подход у некоторых авторов носит наименование *метод подсеток*, см., например, статью К. Ю. Богачева и Н. С. Мельниченко “О пространственной аппроксимации методом подсеток для задачи фильтрации вязкой сжимаемой жидкости в пористой среде”, Вычислит. методы и программирование, 2008, Т. 9, С. 191–199.

Идею метода суперэлементов, как варианта многомасштабного подхода, мы изложим на модельной задаче о решении двумерной задачи Дирихле для уравнения Лапласа в квадратной

расчетной области $\bar{\Omega} = \Omega \cup \Gamma = (0, 1)^2$, в которой расположена одна круглая “дыра” $\bar{\Omega}_0 = \Omega_0 \cup \Gamma_0$ радиуса $R \ll 1$, которую можно рассматривать как скважину в нефтеносном полигоне, см. Рис. 5.16:

$$\begin{aligned} \Delta u &= 0, & (x, y) &\in \Omega/\Omega_0, \\ u|_{\Gamma} &= g(x, y), & u|_{\Gamma_0} &= g_0, \end{aligned} \quad (5.116)$$

где g_0 есть заданная константа.

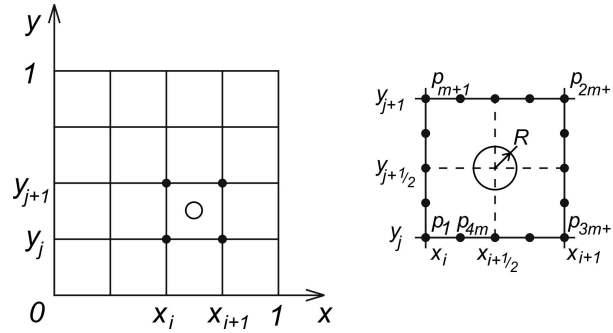


Рис. 5.16. Расчетная область модельной задачи с макросеткой и дискретизация ребер макроэлемента

В области Ω построим “макросетку” Ω^H , состоящую из квадратных суперэлементов (макроячеек) с “большой” стороной H и будем считать, что скважина расположена в центре макроячейки $\Omega_{i,j} = [x_i, x_{i+1}] \times [y_i, y_{i+1}]$. Далее каждое ребро макросетки Ω^H разобьем для простоты равномерно с шагом $h = H/m$ так, что на периметре макроячейки будет расположено $4m$ точек P_k , которым сопоставляем “макро-реберные” кусочно-линейные базисные функции $\varphi_k(x, y)$ со свойствами

$$\varphi_k(P_{k'}) = \delta_{k,k'}, \quad k, k' = 1, \dots, N,$$

где $\delta_{k,k'}$ – символ Кронекера, а N – общее количество реберных базисных функций в Ω^H . Отметим, что $\varphi_k(x, y)$ в рассматриваемом случае квадратной макросетки является фактически на каждом макроребре функцией только одной переменной, x или y .

Для макроэлемента $\Omega_{i,j}$ “со скважиной” мы дополнительно вводим базисную функцию $\varphi_0(x, y)$, как решение уравнения Лапласа, равное g_0 на Γ_0 и нулю – в точках $P_1, \dots, P_{4m}$ на его

сторонах. В силу известного аналитического представления для однородного неограниченного пространства, имеем

$$\varphi_0(x, y) = g_0 \left(1 - \ln \frac{R}{[(x - x_{i+1/2})^2 + (y - y_{i+1/2})^2]^{1/2}} \right). \quad (5.117)$$

Если в точках P_k , $k = 1, \dots, 4m$, на сторонах квадрата $\Omega_{i,j}$ известны значения u_k , то в области $\Omega_{i,j}/\Omega_0$ решение задачи Дирихле для уравнения Лапласа (5.116) может быть представлено как

$$u(x, y) = \varphi_0(x, y) + \sum_{k=1}^{4m} u_k \Phi_k(x, y), \quad (x, y) \in \Omega_{i,j}/\Omega_0. \quad (5.118)$$

Здесь функция $\Phi_k(x, y)$ означает “парциальное” решение задачи Дирихле в $\Omega_{i,j}/\Omega_0$ (индексы i, j в Φ_k и $u(x, y)$ ради краткости опущены):

$$\begin{aligned} \Delta \Phi_k &= 0, & \Phi_k|_{\Gamma^H} &= \varphi_k(x, y), \\ \Phi_k|_{\Gamma_0} &= 0, & k &= 1, \dots, 4m, \end{aligned} \quad (5.119)$$

которое может быть найдено с помощью “стандартного” метода конечных элементов в $\Omega_{i,j}/\Omega_0$. Эта подзадача является достаточно автономной и в алгоритме ее решения имеется большой произвол: связь с остальными суперэлементами заключается только в том, что на их границах Γ^H должны быть общие узлы P_k и базисные функции $\varphi_k(x, y)$ в качестве краевых условий. Внутри же макроячеек могут использоваться различные “макросетки” и аппроксимации.

Отметим, что для суперэлементов “без скважин” в выражении вида (5.118) для локального решения член $\varphi_0(x, y)$ отсутствует, а при определении подзадачи (5.119) условие для Φ_k на Γ_0 также опускается. Заметим также, что если в (5.118) какие-то точки $P_k(x, y)$ лежат на внешней границе Γ расчетной области Ω , то очевидным образом полагается $u_k = g(x_k, y_k)$.

Для вычисления используемых в (5.118) неизвестных значений u_k , $k = 1, \dots, N$, можно воспользоваться оператором Пуанкаре–Стеклова для следов решения задачи (5.116) на совокупности Γ^H внутренних (по отношению к Ω) суперэлементов. В каждой подзадаче типа (5.119) (неважно, с наличием скважины Γ_0 или без) оператор Пуанкаре–Стеклова S_k определяется как

$$S_k \varphi_k = \delta_n \Phi_k, \quad (5.120)$$

где δ_n – внешняя нормальная производная на границе Γ^H рассматриваемой макроячейки. Таким образом, следу φ_k функции Φ_k на границе S_k сопоставляет значение граничной нормальной производной. Обратная связь между этими величинами устанавливается с помощью обратного оператора Пуанкаре–Стеклова:

$$\varphi_k = S_k^{-1} \delta_n \Phi_k. \quad (5.121)$$

Очевидно, что оператор S_k обратим, поскольку краевая задача (5.119) имеет единственное решение. О свойствах операторов Пуанкаре–Стеклова и возможностях их использования см. подробнее [106].

Для замыкания системы уравнений решаемой задачи, т.е. для нахождения значений $u_k = u(P_k)$, $P_k \in \Gamma^H$, определяющих следы искомого решения на внутренних ребрах макросетки, фактически достаточно использовать интерфейсные условия непрерывности решения и его потоков. Очевидно, что этому эквивалентна непрерывность их линейной комбинации $U|_{\Gamma^H} = (\alpha u + \beta \partial u / \partial n)|_{\Gamma^H}$ при некоторых значениях коэффициентов α и β .

Если обозначить через $U|_{\Gamma_+^H}$, $U|_{\Gamma_-^H}$ двусторонние значения функции U , определяемые на двух разных примыкающих к Γ^H суперэлементах, то в терминах оператора Пуанкаре–Стеклова интерфейсное условие можно записать в виде

$$(\alpha u + \beta S u)_{\Gamma_+^H} = (\alpha u + \beta S u)_{\Gamma_-^H}. \quad (5.122)$$

В целом рассматриваемый вариант многомасштабного подхода можно представить следующим образом:

- а) задается макросетка Ω^H и согласованные микросетки Ω^h , имеющие на макроребрах Γ^H общие узлы P_k ;
- б) с помощью МКЭ решаются автономно краевые подзадачи (5.119) в макроячейках, или суперэлементах;
- в) из полученных решений Φ_k вычисляются их нормальные производные на Γ^H и составляется алгебраическая система интерфейсных уравнений относительно значений u_k во всех точках P_k , $k = 1, \dots, N$;

- г) после решения полученной СЛАУ решение в каждом из суперэлементов определяется по формуле (5.118).

Данная методика может быть модифицирована и обобщена в различных направлениях. В первую очередь это касается расширения круга решаемых задач. Например, вместо задачи Дирихле для уравнения Лапласа в прямоугольнике могут рассматриваться уравнения диффузии или фильтрации в многомерных областях со сложной конфигурацией кусочно-гладких многосвязных границ и контрастными материальными свойствами сред, когда коэффициенты резко меняются на порядки (именно такие проблемы и инициировали многомасштабные методы).

Подсеточные подходы также эффективно применяются для методов конечных объемов и разрывных алгоритмов Галеркина, в том числе с аппроксимациями повышенного порядка точности. Далее, как уже отмечалось, интерфейсные условия на внутренних границах контактирующих суперэлементов, или макроячеек, также могут отличаться большим разнообразием. При этом данное направление перекликается с другой активно развиваемой областью вычислительной математики – методами декомпозиции областей, являющимися главным современным оружием распараллеливания алгоритмов. Данные проблемы выходят за рамки этой книги, хотя некоторые аспекты декомпозиции сеточных областей описываются в шестой главе.

Глава 6

Алгоритмы и технологии построения сеток

Построение сетки является важным звеном в математическом моделировании природных явлений и научно-технических процессов. От ее выбора во многом зависит эффективность применяемых численных алгоритмов, включая точность и вычислительную сложность решения поставленной дифференциальной начально-краевой задачи (или соответствующей вариационной формулировки). Большой вклад в исследование этих проблем внесли российские математики: академики С.К. Годунов и А.Ф. Сидоров с учениками, а также В.А. Гаранжа, С.А. Иваненко, А.Д. Лисейкин с коллегами, А.М. Мацокин, Г.И. Шипкин, и многие зарубежные специалисты. Вопросам построения сеток посвящена обширная литература (см., например, монографии [45], [78], [110]–[113], [130], [201], [202] и цитируемые там работы). Данным проблемам посвящены конференции, специальные выпуски журналов, научно-технические отчеты и обширные материалы в интернете. Кроме того, имеется большое количество программных реализаций генераторов сеток, как общедоступных, так и коммерческих, адреса многочисленных сайтов которых легко находятся с помощью поисковиков. Примерами могут служить широко распространенные программные комплексы NETGEN и GMESH.

Сеточная тематика включает значительное количество разнообразных вопросов. Важные аспекты касаются способов дискретизации исходной расчетной области (т.е. построения сеточной расчетной области), которые охватывают конформные и квазиконформные отображения, вариационные принципы, методы дифференциальной геометрии, преобразования графов и различные технологические приемы.

Здесь следует отметить направление, активно развиваемое после пионерской работы А. Уинслоу 1967 г., где предлагалось строить регулярные сетки на основе гармонического отображения заданной двумерной области на единичный квадрат. Успешное продвижение этого подхода, связанное с его обобщением и расширением области применимости, получено в работах московской математической школы, обзор которых приведен в ста-

тье: Б.Н. Азаренок, А.А. Черяхчян. Об одной трудности построения двумерных регулярных сеток с помощью отображений. – Математическое моделирование. – 2014. – Т. 26, № 12. – С. 48–64. Ряд оригинальных подходов и новых алгоритмов получен также представителями екатеринбургской научной школы “сеточников”, из которых мы приведем две характерные публикации: О.В. Ушакова. О невырожденности трехмерных сеток. – Труды ИММ УрО РАН. – 2004. – Т. 12, № 1. – С. 78–100 и М.Ф. Прохорова. Проблемы гомеоморфизма, возникающие в теории построения сеток. – Труды ИММ УрО РАН. – 2008. – Т. 14, № 1. – С. 112–129. Исторически данные направления активно развиваются и в Новосибирске. Например, успешное применение квазиконформных отображений представлено в работе С.К. Годунова, Е.И. Роменского, Г.А. Чумакова. “Построение разностных сеток в сложных областях с помощью квазиконформных отображений” в книге “Вычислительные проблемы в задачах математической физики”, Новосибирск: Изд-во “Наука”, 1990 (ред. М.М. Лаврентьев).

Актуальными проблемами являются определение критериев качества сеток, для которых имеется много разнообразных количественных характеристик, а также подходы к улучшению или оптимизации сеток, основанные на последовательных приближениях. Достаточно содержательный обзор данных направлений имеется в работе V.A. Garanzha, L.N. Kudryavtseva, S.V. Utyuzhnikov. “Variational method for untangling and optimization of spatial meshes”. – J. of Comput. and Appl. Math., v. 264, 2014, 24–41.

Вопросы оптимальной дискретизации многомерной краевой задачи со сложными многосвязными кусочно-гладкими граничными поверхностями и разномасштабными деталями составляют фундаментальную математическую проблему, решение которой видится только в отдаленном будущем. Можно сказать, что зачастую построить оптимальную сетку “дороже”, чем воспользоваться достаточно “хорошей” сеткой и решить на ней поставленную задачу.

Не менее существенны, особенно с практической точки зрения, и технологии реализации: организация сеточных структур данных (ССД), допускающих множественные представления с возможностью их конвертирования и переиспользования существующих программ построения сеток, согласованное информа-

ционное взаимодействие с инструментариями геометрического моделирования, постобработки и визуализации результатов расчета, представляемых в виде многомерных сеточных функций. Данные актуальные вопросы вычислительной геометрии, решаемые эффективно с помощью конечно-элементных аппроксимаций, сложились в самостоятельное научное направление, называемое *изогеометрический анализ* и уже получившее свое монографическое изложение в [94]. Здесь важно отметить следующую происходящую в мире тенденцию: конвергенцию бурно развивающихся программных комплексов математического моделирования с системами автоматизации проектирования (САПР), уже давно составляющих огромный рынок программных продуктов: CAD (Computer Aided Design), CAE (Engineering), CAM (Manufacturing), PLM (Product Lifecycle Management), см. [108], [109], [147], [181], [187], [200]. Данные технологии цифрового проектирования становятся основой очередной технологической революцией (четвертая индустриализация), ассоциируемой с терминами “интернет вещей”, “аддитивные технологии” и “3D-принтеры”.

Построение сетки в сложных случаях (например, при формировании дискретной модели проектируемого самолета или корабля) может составлять по времени значительную часть крупного вычислительного эксперимента. Поэтому чрезвычайно актуальной является проблема автоматической генерации сеток с заданными критериями качества, которая в значительной степени относится к прерогативе искусственного интеллекта. Популярная задача в данном случае носит название *оптимизация геометрических форм* “Shape Optimization” и представляет интересное направление вычислительной топологии, которое уже выходит за рамки нашей книги, см. монографию М.Дельфора, Дж.Золесио [58] и приведенную там многочисленную литературу. В этой связи нельзя не упомянуть о результатах 1970–80 гг. харьковской математической школы В.Л. Рвачева (см. [95], [148], [170], и большую имеющуюся там библиографию), представителями которой разработан аналитический аппарат *R*-функций, чрезвычайно перспективный для вычислительной геометрии и, можно так сказать, геометрической информатики.

Задача данной главы заключается в том, чтобы только подвести читателя к пониманию широкого круга постановочных проблем, а также вычислительных методов и технологий дис-

кретизации многомерных задач с реальными данными, составляющих одну из ключевых стадий математического моделирования.

6.1. Формальные аспекты математических постановок многомерных краевых задач

Пусть в замкнутой ограниченной области

$$\bar{\Omega} = \Omega \cup \Gamma = \bigcup_k \bar{\Omega}_k \quad (6.1)$$

требуется найти вектор-функцию $\vec{u} = \{u_\mu, \mu = 1, \dots, m\}$, удовлетворяющую дифференциальному уравнению

$$L\vec{u}(\vec{x}, t) = \vec{f}(\vec{x}, t), \quad \vec{x} \in \Omega, \quad (6.2)$$

а также граничным и начальным условиям

$$l\vec{u} = \vec{g}(\vec{x}, t), \quad \vec{x} \in \Gamma, \quad \vec{u}(\vec{x}, 0) = \vec{u}^0(\vec{x}). \quad (6.3)$$

Здесь Ω – открытая односвязная область в трехмерном или двумерном евклидовом пространстве, Γ – ее граница, L и l – операторы исходного уравнения и граничных условий, по разному определенные в подобластях Ω_k и граничных сегментах $\Gamma_k = \Gamma \cap \bar{\Omega}_k$ соответственно, где $k = 1, \dots, N_\Omega$, а N_Ω – количество подобластей, на которое по каким-либо соображениям разбита расчетная область Ω .

Например, оператор L может иметь вид

$$L = A \frac{\partial}{\partial t} + \nabla B \nabla + C \nabla + D, \quad (6.4)$$

где A, B, C, D – матрицы, зависящие кусочно-гладким образом от независимой переменной $\vec{x}$, представляющей радиус-вектор в декартовой, цилиндрической или какой-то другой системе координат. Граница представляется в виде двух частей $\Gamma = \Gamma_D \cup \Gamma_N$, на первой из которых задаются условия первого рода (Дирихле)

$$\vec{u} = \vec{g}_D, \quad \vec{x} \in \Gamma_D, \quad (6.5)$$

а на другой – условия второго или третьего рода

$$D_N \vec{u} + A_N \nabla_n \vec{u} = \vec{g}_N, \quad \vec{x} \in \Gamma_N, \quad (6.6)$$

где ∇_n означает внешнюю нормальную производную к Γ_N , а D_N и A_N – в общем случае матричные коэффициенты. Отметим также, что на общих границах смежных подобластей $\Gamma_{k,k'} = \bar{\Omega}_k \cap \bar{\Omega}_{k'}$ могут ставиться условия сопряжения, конкретный вид которых зависит от свойств решаемых уравнений и здесь нами не рассматривается.

Уравнение (6.4) может принимать различный вид в разных подобластях $\bar{\Omega}_k$ расчетной области. Зачастую это означает, что в подобластях заданы среды с различными материальными свойствами или даже решаются разные уравнения. С точки зрения приложений это означает “многофизичность”, или междисциплинарность, задач, когда, например, необходимо моделировать взаимосвязанные процессы тепло-массопереноса, упруго-пластичности, гидро-газодинамики, электромагнетизма и химической кинетики.

С предметной, или прикладной, точки зрения естественно разбивать расчетную область $\bar{\Omega}$ на непересекающиеся подобласти $\bar{\Omega}_k$, в соответствии с материальными свойствами заполняющих их сред. Однако такой подход не является единственно возможным, поскольку при использовании параллельных методов декомпозиции для решения больших задач может сказаться целесообразным формирование подобластей с пересечениями из алгоритмических принципов.

Если в уравнениях (6.4)–(6.6) какие-то из членов зависят от искомого решения, то после применения методов линеаризации алгоритмы приводят к итерационному процессу, на каждом шаге которого решается линейное уравнение. С точки зрения дискретизации задачи, это может привести только к необходимости перестройки сетки на некоторых “нелинейных” итерациях.

Характерной особенностью ряда актуальных задач является наличие сингулярностей, т.е. наличие подобластей с резко меняющимися свойствами решения. К такому классу относятся уравнения с малыми коэффициентами при старших производных, для которых типичным является появление пограничных и/или внутренних слоев с огромными градиентами, а также присутствие границ с входящими углами, имеющими соответствующие асимптотические формулы об особенностях поведения решения в их окрестностях. В таких случаях адаптация сетки произво-

дится путем ее сгущения по определенному закону или априори, или на основе апостериорного анализа полученных предварительных результатов. Последняя ситуация связана с перестройкой сетки в процессе счета. Динамические дискретизации неизбежны также в задачах с движущимися границами, перемещения которых или заданы, или определяются в процессе моделирования.

Наряду с прямыми задачами вида (6.1)–(6.6), когда все коэффициенты уравнений заданы, мы рассматриваем *обратные задачи*, удовлетворяющие следующей математической постановке. Требуется найти минимум *целевого функционала*

$$\Phi_0(\vec{u}(\vec{x}, t, \vec{p}_{opt})) = \min_{\vec{p}} \Phi_0(\vec{u}(\vec{x}, t, \vec{p})), \quad (6.7)$$

определенного на решении $\vec{u}$ и зависящего от вектора оптимизируемых параметров $\vec{p} = (p_1, \dots, p_m)$, при соблюдении линейных или нелинейных (функциональных) ограничений

$$p_k^{min} \leq p_k \leq p_k^{max}, \quad k = 1, \dots, m_1 \quad (6.8)$$

$$\Phi_l(\vec{u}(\vec{x}, t, \vec{p})) \leq \delta_l, \quad l = 1, \dots, m_2, \quad m_1 + m_2 = m. \quad (6.9)$$

Оптимизационная формулировка (6.7)–(6.9) предполагает, как правило, построение последовательных приближений к искомому решению $\vec{u}(\vec{x}, t, \vec{p}_{opt})$, причем на каждой итерации необходимо решать прямую задачу вида (6.2) (называемую в данном случае уравнением состояния) с текущим значением параметров $p_1, \dots, p_m$. Такой итерационный процесс сходится, если начальные значения параметров достаточно близки к точке их локального (единственного) минимума. Однако такая ситуация составляет только частный случай из всех возможных. Более общая проблема – глобальная оптимизация многоэкстремального функционала Φ_0 , см. [73].

Классические исследования по вопросам существования, единственности, корректности и методов решения обратных задач составляют достаточно самостоятельную и развитую область прикладной математики, где значительный вклад внесли школы А. Н. Тихонова, В. В. Иванова и М. М. Лаврентьева, обзор достижений которых ярко представлен в монографии С. И. Кабанихина [86].

Мы предусматриваем решение обратных задач с помощью оптимизационных подходов (см. [84] и цитируемую там литературу), что приводит к необходимости многократной реализации прямых задач, а это требует (как и в “нелинейных” итерациях) неоднократного перестроения сеток. Один из востребованных типов обратных задач связан с оптимизацией геометрии расчетной области, что обуславливает варьирование отдельных частей ее границы и, как следствие, динамическую адаптацию сетки.

В заключение данного параграфа отметим, что в целом вся совокупность исходных данных по начально-краевой задаче составляет формально геометрическую и функциональную структуру данных (ГСД и ФСД), являющихся отправной точкой для генерации сетки. Функциональные объекты — это виды решаемых уравнений, представления их коэффициентов, граничных условий и т.д., описания которых должны быть привязаны к соответствующим геометрическим объектам.

6.2. Геометрические и функциональные объекты

Естественно, что построение сетки является вычислительным этапом, предваряющим аппроксимацию краевой задачи. Поэтому дискретизация расчетной области должна осуществляться с учетом геометрической конфигурации ее границы, а также вида коэффициентов в различных подобластях и характера краевых условий на различных граничных сегментах.

В последние десятилетия математика, не только вычислительная, но и теоретическая, и прикладная, в значительной степени “геометризуется”, в силу чего, например, активно развиваются такие направления, как дифференциальная геометрия и вычислительная топология. Это объясняется появлением новых классов практических задач, связанных с геометрическими объектами сложной конфигурации, требующих развитых средств по исследованию их свойств и преобразований.

Характерным примером современных исследований в данной области является книга М. Дельфора и Дж. Золесио [58], посвященная анализу геометрических форм, метрик, дифференциального исчисления и оптимизации. Здесь также следует отметить результаты харьковской школы В. Л. Рвачева, активно развиваемые еще в 80-е годы прошлого века на основе так на-

зываемых R -функций (см. [148], [170], а также более позднюю книгу В. Ф. Кравченко [95]), позволяющих описывать в аналитической форме сложные геометрические объекты и теоретико-множественные операции над ними.

Геометрические аспекты другого плана сформировали в последние годы направление вычислительной математики, получившее название *изогеометрический анализ*, см. [94]. В данном случае акцент делается на автоматизации построения адаптивных сеток, в том числе криволинейных, при решении многомерных задач со сложными конфигурациями граничных кусочно-гладких поверхностей. При этом широко используются такие базисные функции, как B -сплайны, “нурбсы” (NURBS, Non-Uniform Rational B -Splines, русского эквивалентного термина здесь пока не сложилось), кривые и поверхности Безье, см. [82]. Фактически здесь складывается новая вычислительная методология, называемая *конечно-элементный анализ*, которая непосредственно интегрирует технологии математического моделирования с промышленными САПР-ами.

В данном параграфе мы кратко рассмотрим на содержательном уровне структуры геометрических и функциональных данных, а также основные операции с ними, более подробно представленных в работе Л.А.Голубевой, В.П.Ильина и А.Н.Козырева “О программных технологиях в геометрических аспектах математического моделирования”. – Вестник НГУ, серия Информационные технологии, т. 10, N 2, 2012, 25-33. Будем для простоты сначала предполагать, что подобласти Ω_k в Ω не пересекаются, то есть

$$\Omega_k \cap \Omega_{k'} = 0, \quad k \neq k', \quad k, k' = 1, \dots, N_\Omega, \quad (6.10)$$

где N_Ω – заданное число подобластей. Каждая подобласть вместе со своей границей Γ_k образует замыкание, т.е. $\bar{\Omega}_k = \Omega_k \cup \Gamma_k$.

Общую границу двух соседних подобластей $\Omega_k, \Omega_{k'}$ обозначаем через $\Gamma_{k,k'} \subset \Gamma_k$:

$$\bar{\Omega}_k \cap \bar{\Omega}_{k'} = \Gamma_{k,k'} = \Gamma_{k',k} \neq 0. \quad (6.11)$$

Внешняя граница расчетной области Γ представляется в виде совокупности поверхностных сегментов $\Gamma_{k,0}$, $\Gamma = \bigcup_k \Gamma_{k,0}$, $(\Gamma_{k,0} = \bar{\Omega}_k \cap \bar{\Omega}_0$, где через $\bar{\Omega}_0$ обозначается формально вся “внешность”

по отношению к Ω), каждый из которых отличается своей геометрией или/и заданным на нем краевым условием. Границы подобластей представляются в виде

$$\Gamma_k = \Gamma_k^{(e)} \cup \Gamma_k^{(i)}, \quad (6.12)$$

где $\Gamma_k^{(e)}$ – сегменты, принадлежащие внешней границе, а $\Gamma_k^{(i)} = \bigcup_{k' \neq 0} \Gamma_{k,k'}$ – фрагменты внутренней границы. Отметим, что на $\Gamma_k^{(i)}$ не ставится никаких граничных условий, если соответствующая кривая или поверхность вводятся только из информационного удобства и не являются границами раздела различных сред.

Расчетная область в целом задается при помощи следующих геометрических объектов, помимо подобластей (объемов Ω_k): вершины (точки) $V_p = (x_p, y_p, z_p)$, $p = 1, \dots, N_V$, поверхности $S_s(x, y, z) = 0$, $s = 1, \dots, N_S$, линии C_l $l = 1, \dots, N_C$, поверхностные сегменты (границы) F_f , $f = 1, \dots, N_F$, отрезки (или ребра – односвязные куски линий, в том числе криволинейных, снабженные своей начальной и конечной точками) E_e , $e = 1, \dots, N_E$. Кроме того, для экономии объема задаваемой информации целесообразно ввести понятие базовых, или опорных, координат (например, $X_c = \{X_1, \dots, X_{NX}\}$, $Y_c = \{Y_1, \dots, Y_{NY}\}$ в двумерной декартовой системе), с помощью номеров которых удобно задавать координаты вершин.

Помимо использованной выше декартовой системы координат, могут использоваться и другие: цилиндрическая, сферическая и какие-либо специальные координаты. Более того, для удобства могут применяться различные локальные системы координат, с указанием их положения и ориентации относительно (единственной!) глобальной системы. Например, локальные координаты используются для эффективного задания типовых кривых или поверхностей в их каноническом виде (сфера, цилиндр, конус и т.д.).

В более сложных случаях целесообразно задание отрезков линий и сегментов поверхностей (в том числе криволинейных) в параметрическом виде (например, в форме кривых или *поверхностей Безье* [82]):

$$\begin{aligned} x &= x(\tau), \quad y = y(\tau), \quad z = z(\tau), \\ x &= x(\tau^{(1)}, \tau^{(2)}), \quad y = y(\tau^{(1)}, \tau^{(2)}), \quad z = z(\tau^{(1)}, \tau^{(2)}), \end{aligned} \quad (6.13)$$

где $\tau, \tau^{(1)}, \tau^{(2)}$ – параметры с задаваемыми границами их изменения.

Существует большое разнообразие способов задания геометрических объектов. Например, линия в трехмерном пространстве может задаваться или уравнениями двух пересекающихся поверхностей, или каким-то специальным образом (например, окружность задается тремя точками). Канонические поверхности экономично также описываются: сфера – своим центром и радиусом, конус – основанием и вершиной, и т.д. При использовании множественных представлений геометрических объектов необходимо, естественно, иметь средства перевода информации из одних форматов в другие.

Между геометрическими объектами устанавливаются топологические связи при помощи матриц инцидентности (смежности):

$$M_{VE} = \{m_{i,j}^{VE} : i = 1, \dots, N_V; j = 1, \dots, N_E\} \text{ – (вершины – ребра),}$$

$$M_{EF} = \{m_{i,j}^{EF} : i = 1, \dots, N_E; j = 1, \dots, N_F\} \text{ – (ребра – грани),}$$

$$M_{F\Omega} = \{m_{i,j}^{F\Omega} : i = 1, \dots, N_F; j = 1, \dots, N_\Omega\} \text{ – (грани – подобласти).}$$

Здесь матричными элементами являются номера соответствующих объектов, а сами матрицы M_{VE}, M_{EF} и $M_{F\Omega}$ устанавливают взаимосвязи между вершинами и ребрами, между ребрами и поверхностными сегментами, между поверхностными сегментами и подобластями соответственно.

Отметим, что иногда из алгоритмических соображений, например, при использовании метода декомпозиции областей, целесообразно задавать пересекающиеся подобласти. В этом случае общая область перекрытия может принадлежать нескольким подобластям. Принципиально это не меняет геометрического формализма, надо только контролировать корректность определения материальных, или функциональных, свойств в пересечениях.

Другой пример декомпозиции по методическим соображениям – выделение подобластей в форме прямоугольников или параллелепипедов (в зависимости от размерности задачи), на которых задаются равномерные сетки, с целью применения эконо-

мичных аппроксимаций высоких порядков и специальных быстрых методов решения получаемых алгебраических подсистем.

Из комбинаций простейших, или примитивных, объектов могут составляться типовые фигуры, которые целесообразно использовать при конструировании сложных расчетных областей. Более того, для ускорения сборки можно предусмотреть также иерархию составных тел.

Функциональные объекты, отображающие материальные свойства сред и другие условия задачи, устанавливаются с приведением в соответствие с подобластями и граничными сегментами. К ним относятся типы уравнений и их коэффициенты (например, в подобласти Ω_k может быть определено скалярное уравнение Гельмгольца $-\nabla \lambda_k \nabla u + \varkappa_k u = f_k$), виды и коэффициенты краевых условий (пример – задание на граничном сегменте Γ_l условия Ньютона $\alpha_l u_l + \beta_l \frac{\partial u}{\partial n} = \gamma_l$), а также – в нестационарных задачах – начальные данные для искомого решения. При этом числовые или функциональные представления величин задаются с помощью своих значений или определяемых пользователем процедур.

Материальные свойства в простейших случаях задаются просто числами, например, постоянный коэффициент диффузии или теплопроводности в какой-то подобласти. Более общие модели включают зависимости или от независимых переменных, или от искомых решений. Самые сложные ситуации возникают в многофазных и многокомпонентных средах. Например, простая вода может находиться в жидком, твердом и парообразном состояниях, а ее термодинамические характеристики сильно зависят от давления и температуры. Аналогичная картина зачастую наблюдается и для разных сплавов, описываемых с помощью сложных диаграм состояния, построение которых — это сложная исследовательская работа как в экспериментальном, так и в теоретическом плане. И здесь также далеко не последняя роль принадлежит математическому моделированию.

Отметим также, что некоторые из числовых исходных данных могут быть определены как варьируемые параметры в целях организации многовариантных расчетов или оптимизационных алгоритмов решения обратных задач. В этих случаях их изменения осуществляются пользовательскими функциями и средствами управления вычислительными процессами.

Для обеспечения вариаций геометрических примитивов или более сложных тел необходимо предусмотреть набор операций над объектами: сдвиг, поворот, масштабирование и др., – которые должны сопровождаться согласованным изменением взаимосвязанной информацией и не нарушать топологию расчетной области в целом. Более сложные геометрические модификации должны предусматривать и изменение топологии, например, добавление или удаление каких-то фигур, т.е. тел с какими-то своими материальными свойствами, что может привносить более кардинальные изменения математической постановки, которое должно обеспечиваться диалоговыми средствами для поддержки “человеческого фактора”.

6.3. Сеточные объекты и операции

Основные задачи данного раздела – формализация основных понятий сеточных технологий, а также краткое описание принципов построения и модификации сеток, которые составляют алгоритмическое (функциональное) наполнение сеточных технологий.

6.3.1. Основные понятия и спецификации сеток

Дискретизация расчетной области заключается в построении сетки $\bar{\Omega}_h$, которая представляет собой сеточную расчетную область и включает следующие сеточные объекты, аналогичные “макрообъектам” расчетной области:

$$\text{сеточные подобласти } \bar{\Omega}_k^h: \bar{\Omega}^h = \bigcup_k \bar{\Omega}_k^h, \bar{\Omega}_k^h = \Omega_k^h \cup \Gamma_k,$$

$$k = 1, \dots, N_\Omega^h, \Omega_k^h \approx \Omega_k,$$

сеточные границы подобластей

$$\Gamma_{k,k'}^h = \bar{\Omega}_{k'}^h \cap \bar{\Omega}_k^h, \Gamma_l^h = \bigcup_k \Gamma_{k,l}^h,$$

$$\Gamma_k^h \approx \Gamma_k, k, k', l = 1, \dots, N_\Gamma^h,$$

$$\text{узлы сетки } V_p^h = (x_p, y_p, z_p), p = 1, \dots, N_V^h,$$

$$\text{сеточные ребра } E_s^h, s = 1, \dots, N_E^h,$$

$$\text{сеточные грани } F_l^h, l = 1, \dots, N_F^h,$$

$$\text{сеточные объемы, или элементы, } T_m^h, m = 1, \dots, N_T^h,$$

где $N_\Omega^h, N_\Gamma^h, N_V^h, N_E^h, N_F^h, N_T^h$ означают соответственно общее количество сеточных подобластей, граней, узлов, ребер, граней и элементарных объемов.

Мы будем рассматривать *адаптивные сетки* в следующем смысле: вершины расчетной области являются узлами сетки, а ребра (макроребра), грани (макрограни) и подобласти аппроксимируются соответствующими сеточными (элементарными) объектами, а в идеальном случае макрообъекты состоят из совокупности элементарных объектов. Последнее означает, что сетки в общем случае могут быть криволинейными или, по крайней мере, в конечных элементах, или объемах, некоторые ребра и грани допускаются непрямолинейными.

Между сеточными объектами существуют следующие топологические связи. Сетку можно представить как совокупность конечных объемов $\bar{\Omega}^h = \bigcup_m \bar{T}_m^h$, грани – как поверхности пересечения соседних конечных объемов с номерами m и m' , $F_l^h = F_{m,m'}^h \equiv \bar{T}_m^h \cap \bar{T}_{m'}^h$, ребра – как линии пересечения смежных граней ($E_s^h = E_{l,l'}^h \equiv \bar{F}_l^h \cap \bar{F}_{l'}^h$) и, наконец, узлы – как точки пересечения ребер e и e' , $V_p^h = V_{e,e'}^h \equiv \bar{E}_e^h \cap \bar{E}_{e'}^h$, или как точки пересечения граней l, l', l'' , $V_{l,l',l''}^h \equiv \bar{F}_l^h \cup \bar{F}_{l'}^h \cup \bar{F}_{l''}^h$. Аналогично тому, как это делается с геометрическими объектами, можно ввести матрицы инцидентности, устанавливающие следующие связи между сеточными объектами: матрицы инцидентности $M_{V,E}^h$ – (узлы–ребра), $M_{V,T}^h$ – (узлы–объемы), $M_{E,F}^h$ – (ребра–грани), $M_{F,T}^h$ – (грани–объемы).

Кроме того, указывается связь сеточных объектов с функциональными данными, а именно принадлежность конечного объема T^h сеточной подобласти ($T_m^h \in \Omega_k^h$) и сеточной грани F_l^h – границе Γ_k^h ($F_l^h \in \Gamma_k^h$). Имеется в виду, что по номерам подобластей и граничных сегментов из ГСД и ФСД могут быть определены коэффициенты решаемых уравнений и граничных условий исходной краевой задачи, необходимые для выполнения сеточной аппроксимации (с помощью МКО или МКЭ) математической постановки. Совокупность всей описанной информации о сеточных объектах и их топологических взаимосвязях составляет сеточную структуру данных.

По структурным, или топологическим, свойствам мы будем различать следующие типы сеток: структурированные, неструк-

турированные, квазиструктурированные, согласованные и несогласованные, регулярные и нерегулярные, конформные и неконформные.

Сетку Ω_h будем называть *структурированной*, если она образована при помощи семейства однотипных координатных плоскостей (линий – в двумерном случае), имеет топологически подобные конечные объемы и простые правила нумерации ее элементов, граней, ребер и узлов, позволяющие легко находить соседний объект. В противном случае будем называть сетку *неструктурированной*. Под этим можно также понимать такую сетку, в которой для каждого узла номера ближайших, в геометрическом смысле, соседних узлов можно задать только их перечислением. Частным случаем неструктурированной сетки является *квазиструктурированная* сетка, состоящая из сеточных подобластей Ω_k^h , в каждой из которых задана своя структурированная подсетка. Если $\Gamma_{k,k'}^h \neq \Gamma_{k',k}^h$, то есть сеточные границы двух смежных сеточных подобластей $\Omega_k^h, \bar{\Omega}_{k'}^h$ совпадают, то такие подсетки назовем *согласованными*, а в противном случае – *несогласованными*.

Качественная дискретизация расчетной области должна обладать двумя следующими свойствами, определяющими *адаптивность* сетки, а именно: обеспечивать хорошую аппроксимацию границы и иметь возможность изменения шага вблизи больших градиентов решения. Первое свойство означает выполнение следующих требований: вершины подобластей являются узлами сетки $V \subset V^h$, а границы и ребра подобластей достаточно точно аппроксимируются соответствующими сеточными объектами ($E \tilde{\subset} E^h, F \tilde{\subset} F^h$). Второе свойство обеспечивается за счет регулировки плотности узлов сетки в сеточных подобластях.

При решении стационарных прямых задач мы имеем дело со статическими сетками, а при решении задач геометрической оптимизации и/или при решении нестационарных задач с движущимися границами или перемещающимися сингулярностями — с динамическими сетками.

Сетки могут содержать следующие виды конечных, т.е. микроуровня, объемов, причем в различных сеточных подобластях могут быть различные типы сеточных элементов: тетраэдры, шестигранники (в том числе – параллелепипеды), призмы или симплексы разных конфигураций и т.д.

Среди структурированных сеток можно выделить подкласс сеток, называемых *регулярными*, которые используют только одинаковую топологию конечных элементов, или объемов. Хотя этот вопрос носит чисто терминологический характер, поскольку нерегулярные сетки можно отнести к неструктурированным.

Для формализации процесса дискретизации полезно выделить две структуры, связанные с построением сетки. Первая – это топология, определяемая комбинаторной связностью вершин, а вторая – геометрическая форма, зависящая не только от связности, но и от локализации узлов. Большой класс сеток основывается на понятии геометрического симплекса.

Пусть $\vec{x}_i = (x_{1,i}, \dots, x_{n,i})^T, i = 1, \dots, n+1$ есть набор из $(n+1)$ -й точки в n -мерном пространстве $\mathbb{R}^n$. Будем говорить, что точки $x_1, \dots, x_{n+1}$ не лежат в одной гиперплоскости, если n векторов $\vec{x}_1\vec{x}_2, \dots, \vec{x}_1\vec{x}_{n+1}$ линейно независимы. Это эквивалентно невырожденности матрицы

$$A = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n+1} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n+1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,n+1} \\ 1 & 1 & \cdots & 1 \end{bmatrix}.$$

Для любой точки с радиус-вектором $\vec{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ из решения невырожденной системы линейных алгебраических уравнений

$$A \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \\ \lambda_{n+1} \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{bmatrix} \quad (6.14)$$

мы получаем единственный набор чисел $\lambda_i(\vec{x}), 1 \leq i \leq n+1$, удовлетворяющий условию $\sum_{i=1}^{n+1} \lambda_i = 1$. При этом произвольный набор из $d+1$ точек $\vec{x}_1, \dots, \vec{x}_{d+1}$, не лежащих в одной гиперплоскости, определяет выпуклую оболочку

$$S := \{\vec{x} = \sum_{i=1}^{d+1} \lambda_i x_i \mid 0 \leq \lambda_i \leq 1, \sum_{i=1}^{d+1} \lambda_i = 1\},$$

которая называется d -симплексом, образованным вершинами $\vec{x}_1, \dots, \vec{x}_{d+1}$.

Например, треугольник есть 2-симплекс, а тетраэдр – 3-симплекс. Для любого целого $0 \leq m \leq d-1$ m -мерная грань S есть m -симплекс, образованный m вершинами. Нуль-мерная грань есть вершина, а одномерная – это ребро S . Величины $\lambda_1(\vec{x}), \dots, \lambda_{d+1}(\vec{x})$ называются *барицентрическими координатами* точки $\vec{x}$ в d -симплексе с вершинами $\vec{x}_1, \dots, \vec{x}_{d+1}$. Эти координаты имеют простой геометрический смысл. А именно, если для любого $\vec{x} \in S$ обозначить через S_i симплекс, получаемый заменой вершины $\vec{x}_i$ на точку $\vec{x}$, то из вышеприведенной системы (6.14) для λ_i по правилу Крамера получаем

$$\lambda_i(\vec{x}) = |S_i(\vec{x})| / |S(x)|,$$

где $|\cdot|$ обозначает меру Лебега в $\mathbb{R}^d$, т.е. длину, площадь и объем для $d = 1, 2, 3$ соответственно.

В сеточных технологиях оказывается удобным вводить *стандартные симплексы* $S^{(n)}$, образованные вершинами $0, e_1, \dots, e_n$, где $e_i = (0, \dots, 1, \dots, 0)$ суть орты. С их помощью произвольный n -симплекс $S \subset \mathbb{R}_n$ является образом $S^{(n)}$, полученным из аффинного преобразования $B : S^{(n)} \rightarrow S$ при $B(\vec{e}_i) = \vec{x}_i, i = 1, \dots, n+1$. Примеры стандартных симплексов для $n = 1, 2, 3$ даны на рис.6.1.

Если Ω представляет собой полигональную область в $\mathbb{R}^n$, то ее *симплектическая дискретизация*, или *триангуляция*, представляет собой разбиение на треугольники при $n = 2$ и на тетраэдры – при $n = 3$. Получаемая симплектическая сетка, представляющая собой набор симплексов, или элементов,

$$S^h = S_k^h, \quad k = 1, \dots, M, \quad M = O(h^{-n}),$$

называется *конформной*, если выполняются условия

$$\bar{\Omega} = \bigcup_{k=1}^m \bar{S}_k^h, \quad S_k^h \cap S_{k'}^h \quad \text{при} \quad k \neq k',$$

причем пересечением замыканий разных симплексов $\bar{S}_k^h$ и $\bar{S}_{k'}^h$ может быть только их общая грань. Если последнее условие не выполняется, то сетка называется *неконформной*. В этом случае сетка содержит *висячие*, или *терминальные*, узлы P_l , не

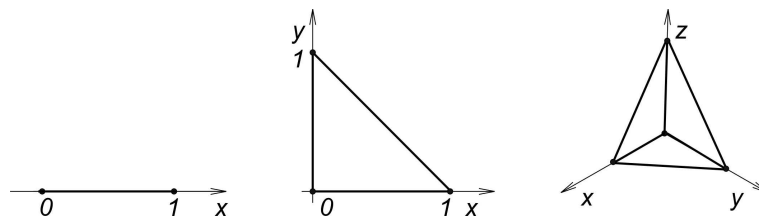


Рис. 6.1. Примеры стандартных симплексов

являющиеся вершиной какого-то из инцидентного ($P_l \in \bar{S}_k^h$) замыкания симплекса $\bar{S}_k^h$. Примеры конформной и неконформной триангуляций для двумерного случая изображены на рис. 6.2.

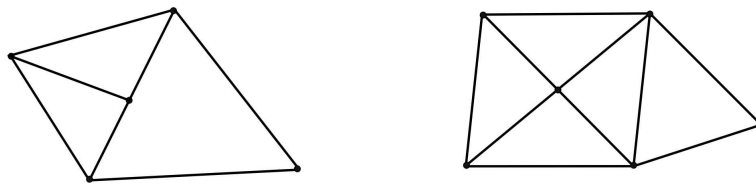


Рис. 6.2. Примеры конформной (слева) и неконформной (справа) триангуляций двумерных сеток

Важно понимать, что построение сетки в задачах математического моделирования — это не самоцель, а лишь средство эффективного решения исходной задачи. В этом плане надо отметить два основных последующих вычислительных этапа. Первый заключается в аппроксимации функциональных уравнений с помощью МКО или МКЭ, а второй — в решении систем полу-

чаемых линейных или нелинейных алгебраических уравнений. Как правило, и аппроксимационные, и алгебраические свойства дискретных систем являются наилучшими в случае равномерных сеток, которые состоят из одинаковых элементов (например, прямоугольников или треугольников). Более того, идеальные сетки — это правильные, или равносторонние (например, когда элемент — это квадрат или куб). Заметим также, что такие сетки хороши не только в математическом отношении, но и в технологическом: для их конструирования и хранения нужен минимум операций и объема оперативной памяти.

Хотя конфигурация расчетных областей в практических задачах очень редко позволяет такие сетки применять, зачастую можно выбрать относительно большую подобласть и построить в ней равномерную или даже правильную сетку, что в целом даст значительную экономию вычислительных ресурсов. Минимальное обобщение таких сеток — переход на кусочно-равномерные сетки, в которых по координатным линиям задаются “зоны равномерности”, например,

$$X^{(k)}, k = 0, 1, \dots, N_x; Y^{(l)}, l = 0, 1, \dots, N_y,$$

в каждой из которых шаги выбираются одинаковыми:

$$h_k^x = (X^{(k)} - X^{(k-1)})/N_k^x, h_l^y = (Y^{(l)} - Y^{(l-1)})/N_l^y,$$

где N_k^x и N_l^y — количество шагов в каждой из зон. Дальнейшее обобщение таких сеток состоит в том, что в каждой зоне шаги изменяются по заданному правилу арифметической или геометрической прогрессии. Укажем здесь и на полезность таких сеток, как квазиравномерные, в которых соседние шаги меняются на относительно малую величину, что не меняет практически аппроксимационных и алгебраических свойств алгоритмов.

В качестве простого обобщения кусочно-равномерных сеток приведем еще примеры двумерных прямоугольных кусочно-аналитических сеток, в которых шаги изменяются по арифметической или геометрической прогрессии соответственно. При этом координаты узлов определяются по формулам

$$x_{i+1} = x_i + h_i^x, y_{j+1} = y_j + h_j^y, i = 0, 1, \dots, L-1; j = 0, 1, \dots, M-1.$$

Разности арифметических прогрессий шагов сеток в зонах для первого случая вычисляются следующим образом:

$$d_k^x = \left(\frac{X^{(k)} - X^{k-1}}{N_k^x} - h_{(k)}^x \right) \frac{2}{N_k^x - 1}, \quad k = 1, \dots, N_x,$$

$$d_l^y = \left(\frac{Y^{(l)} - Y^{l-1}}{N_l^y} - h_{(l)}^y \right) \frac{2}{N_l^y - 1}, \quad l = 1, \dots, N_y,$$

где начальные шаги в каждой зоне равны

$$h_{(k)}^x = \frac{X^{(k)} - X^{k-1}}{N_k^x} - d_k^x \frac{N_k^x - 1}{2},$$

$$h_{(l)}^y = \frac{Y^{(l)} - Y^{l-1}}{N_l^y} - d_l^y \frac{N_l^y - 1}{2},$$

а сами шаги определяются как

$$h_i^x = h_{(k)}^y + (i - \bar{N}_{k-1}^x) d_k^x, \quad i = \bar{N}_{k-1}^x + 1, \dots, \bar{N}_k^x,$$

$$h_j^y = h_{(l)}^y + (j - \bar{N}_{l-1}^y) d_l^y, \quad j = \bar{N}_{l-1}^y + 1, \dots, \bar{N}_l^y.$$

Здесь $\bar{N}_k^x = \sum_{s=1}^k N_s^x$ и $\bar{N}_l^y = \sum_{s=1}^l N_s^y$ суть суммарные количества шагов в первых k и l зонах по x и y соответственно.

Если знаменатели геометрических прогрессий сеточных шагов по x, y равны α_k, β_l , то для второго случая первые шаги в зонах вычисляются по формулам

$$h_{(k)}^x = (X^k - X^{k-1})(\alpha_k - 1)/(\alpha_k^{N_k^x} - 1), \quad k = 1, \dots, N_x,$$

$$h_{(l)}^y = (Y^l - Y^{l-1})(\beta_l - 1)/(\beta_l^{N_l^y} - 1), \quad l = 1, \dots, N_y,$$

а сами шаги в соответствующих зонах равны

$$h_i^x = h_{(k)}^x \alpha_k^{i - \bar{N}_{k-1}^x}, \quad i = \bar{N}_{k-1}^x + 1, \dots, \bar{N}_k^x,$$

$$h_j^y = h_{(l)}^y \beta_l^{j - \bar{N}_{l-1}^y}, \quad j = \bar{N}_{l-1}^y + 1, \dots, \bar{N}_l^y.$$

Для неструктурированных многомерных сеток под шагами подразумеваются расстояния между ближайшими узлами. Одной из количественных характеристик сетки является отношение максимального шага к минимальному: $\varkappa = h_{\max}/h_{\min}$. Надо сказать, что набросанные по какому-то закону в области точки еще не определяют сетку, так как при этом ребра, грани и конечные объемы могут строиться разными способами. Наиболее

естественный подход заключается в построении ячеек Вороного, или Дирихле–Вороного: для данного узла такая ячейка — это множество (всегда однозначное!) точек, расположенных к нему ближе, чем к остальным узлам. Формально для этого достаточно соединить точку со всеми остальными узлами, провести через середины полученных отрезков перпендикулярные к ним плоскости и выбрать определяемый ими минимальный объем, см. пример на рис. 6.3.

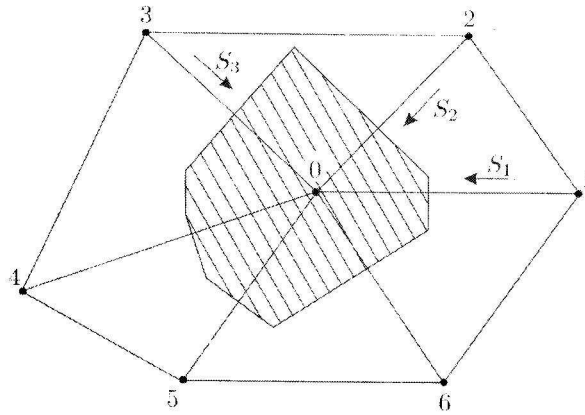


Рис. 6.3. Ячейка Дирихле в двумерном случае

Отметим, что такая дискретизация является двойственной к так называемой триангуляции Делоне, т.е. получаемые в ячейке Дирихле треугольники (или тетраэдры в трехмерном случае) обладают тем свойством, что построенные по их вершинам круги (шары для 3D) не содержат внутри себя других узлов сетки.

Для характеристики сеток имеется большое количество критериев качества, например, — величина минимального угла в ячейке. Узлы построенной сетки, которые соединены ребрами, называются соседними, а совокупность узлов, соседних к узлу $V_p, p = 1, \dots, N_V^h$, называется соответствующим сеточным шаблоном и обозначается через ω_p . Фактически ω_p можно определить как совокупность номеров ($p' \in \omega_p$) узлов, являющихся соседними к V_p . Введенную выше матрицу инцидентности $M_{V,E}^h$ (узлы — ребра) можно интерпретировать как определение графа, изоморфного сетке.

Для большого класса сеточных МКЭ- или МКО-аппроксимаций (которые будем называть узловыми) получаемая система алгебраических уравнений (СЛАУ) имеет вид

$$(Au)_p = a_{p,p}u_p + \sum_{p' \in \omega_p} a_{p,p'}u_{p'} = f_p, \quad p = 1, \dots, N_V^h, \quad (6.15)$$

где ненулевые внедиагональные элементы матрицы $A = \{a_{p,p'}\}$ соответствуют ребрам сетки. Таким образом, портрет матрицы, или матричный граф, характеризующий расположение ненулевых элементов, является изоморфным сеточному графу.

В уравнении (6.15) через ω_p обозначено множество узлов (точнее говоря, их номеров), называемое сеточным шаблоном p -го узла. Более строго, сам “центральный” узел с номером “ p ” также включается в состав шаблона ω_p . На алгебраическом языке “ p ” означает номер строки матрицы A , а совокупность “ p ” – номера столбцов ненулевых элементов $a_{p,p'}$ в p -й строке. По количеству узлов в ω_p сеточные алгебраические уравнения классифицируются на трехточечные, пятиточечные, семиточечные и т.д. Конечно, тип шаблона еще определяется геометрией, или топологией, расположения его узлов. Если для различных узлов шаблоны одинаковы, то такие сеточные вычислительные схемы называются однородными, а в противном случае – неоднородными. “Одинаковость” шаблона понимается в смысле совпадения относительных номеров $p'' = p' - p$ внедиагональных элементов матрицы. Например, для простейшего трехточечного одномерного сеточного уравнения $a_{i,i}u_i + a_{i,i-1}u_{i-1} + a_{i,i+1}u_{i+1} = f_i$ такой “относительный” шаблон можно определить как $\omega_i = \{-1, 0, +1\}$.

Одна из важнейших сеточных технологий заключается в использовании для решения задачи не одной, а нескольких сеток $\Omega_0^h, \Omega_1^h, \dots$, как правило, сгущающихся, т.е. имеющих характерные шаги $h_0 > h_1 > \dots$, теоретически стремящиеся к нулю. Такие *многосеточные методы* позволяют строить оптимальные по порядку методы, в которых объем вычислений пропорционален общему количеству узлов. Наиболее распространенный класс таких алгоритмов – алгебраические (в отличие от геометрических) многосеточные методы (AMG). Здесь важное место занимают последовательности вложенных сеток, когда узлы редких сеток являются также узлами более густых сеток. Простейшим

примером построения вложенной сетки является деление каждого ребра исходной сетки пополам, а также добавление новых узлов в центрах тяжести граней и конечных объемов. Помимо “глобального” сгущения, используются и локальные сгущения сетки, использующие мельчение шагов только в некоторых подобластях. На последовательности неравномерных сеток важный класс составляют *квазиравномерные сетки*, у которых при $h \rightarrow 0$ отношение максимального шага к минимальному является конечным (при теоретическом исследовании вопросов сходимости сеточных решений данное свойство фактически является обязательным).

Отметим еще одну важную операцию над сетками – *декомпозиция сеточной расчетной области Ω^h на подобласти Ω_k^h* , которая связывается с проблемой масштабируемого распараллеливания алгоритмов. Здесь возникает много своих вопросов: сбалансирование разбиений, перенумерация узлов, переформирование сеточной структуры данных и др.

6.3.2. Некоторые подходы к генерации сеток

Один из естественных подходов — это конструирование сетки “сверху вниз”. При этом расчетная область Ω разбивается из каких-то соображений на подобласти Ω_k , которые в совокупности со своими гранями F_f , ребрами E_e и вершинами V_p могут рассматриваться как макросетка, у которой топологические связи между макрообъектами описываются с помощью матриц инцидентности M_{VE} , M_{EF} , $M_{F\Omega}$, см. выше п.6.2.

Иерархическая генерация сетки заключается в последовательной дискретизации макроребер, затем – макрограней (с использованием уже построенных узлов на их границах) и в заключение – макрообъемов, т.е. подобластей. При этом мы получаем согласованные сетки, которые обладают очевидными хорошими свойствами, но не могут быть просто использованы в случае, например, взаимного перемещения контактирующих моделируемых тел.

Описанный способ обладает рядом положительных качеств:

- легко обеспечивается адаптивность сетки к особенностям граничных поверхностей области, ребер и вершин;
- имеются хорошие возможности построения качественных сеток в подобластях за счет выбора их конфигураций, а

также подбора соответствующих алгоритмов типа конформных или квазиконформных преобразований;

- естественным образом осуществляется декомпозиционное представление сеточной расчетной области, необходимое для распараллеливания алгоритмов;
- решается связанный с предыдущим вопрос о создании иерархической структуры сеточных данных: последовательно нулевой размерности, первой, второй и третьей.

Самый сложный вопрос в данном случае — как автоматически реализовать “хорошую” декомпозицию области? Естественно, что опытный пользователь, оснащенный наглядным графическим интерфейсом, сделает это быстро, интуитивно и почти не задумываясь. Однако переложить эту задачу на искусственный интеллект — не совсем тривиальная проблема.

Альтернативный выше рассмотренному способ состоит в построении локально модифицированных сеток. На расчетную область (или подобласть), вместе с ее ближайшим окружением, наносится какая-то регулярная сетка, а затем точки пересечения ее ребер с границей (а скорее всего, и некоторые геометрически соседние узлы) каким-нибудь образом передвигаются, чтобы сделать ее адаптивной и удовлетворяющей некоторым критериям качества (например, условию Делоне). Очевидно, что если при этом окажется относительно много регулярных узлов, то это позволит сформировать экономичную ССД. А в случае присутствия разномасштабных объектов в области на последующих этапах сетку можно где-то сгущать, а где-то разреживать.

Наиболее универсальными генераторами сеток используются алгоритмы фронтального типа, в которых узлы сетки определяются последовательно от границы внутрь области. При этом степени сгущения или разрежения узлов могут регулироваться априорно заданными функциями плотности. Локальные сгущения сетки могут также производиться и апостериори, на основе анализа промежуточного характера численного решения с помощью функций, называемых “эстиматорами” (estimators).

6.3.3. Вопросы формирования поузловой информации

Различные модификации, декомпозиции или объединения сеток — это в первую очередь операции над структурами данных, а

другими словами — преобразования графов, отображающих топологические связи между сеточными объектами. Эти сеточные графы тесно связаны с соответствующими матричными графами, что особенно наглядно в простейшем случае, когда одному узлу соответствует после аппроксимации одно алгебраическое уравнение и в узле определена только одна неизвестная функция.

В этом случае топология неструктурированной сетки может быть описана с помощью целых массивов, составляющих *сжатый разреженный формат* (CSR – Compressed Sparse Row), который является стандартным для алгебраических универсальных решателей. В применении к сеткам CSR содержит для каждого узла минимально необходимую информацию: количество соседних (соединенных ребрами) узлов и номера самих соседей. ССД должна содержать также поузловые вещественные данные, т.е. координаты узлов и определенные в них (заданные значениями или вычисляемые) функции.

Содержание CSR зависит от нумерации узлов и естественной выглядит иерархическая упорядоченность, когда подряд нумеруются (неважно как) сначала внутренние узлы первой сеточной подобласти, затем – второй, и т.д., после чего следуют все граневые узлы, все реберные узлы (со своими внутренними разбивками) и макровершины. Для необходимой идентификации типа узла, т.е. какому макроребру, какой макрограницы или подобласти он принадлежит, требуется еще сформировать макроструктурный формат (MCSR), в котором указывается количество узлов различного типа, множества номеров этих типов и связи между ними. Таким образом, например, для каждого внутреннего узла подобласти можно определить ее номер, является ли он околোগраничным и, если да, к какой макрогранице примыкает.

Отсюда становится целесообразной иерархическая упорядоченность узлов, в том числе локальная и глобальная нумерации, а также неизбежность, в силу различных вычислительных потребностей, перенумерация узлов (в том числе из локальной в глобальную и наоборот).

6.3.4. Поэлементные технологии

Замечательные достижения методов конечных элементов обязаны не только развитию уникальной теории численного решения широкого класса дифференциальных и интегральных уравнений с применением финитных базисных функций высокого порядка на неструктурированных сетках, но и созданию эффективных вычислительных поэлементных технологий, основанных на определении локальных матриц и сборки глобальной матрицы задачи. Этот подход значительно упрощает программную реализацию алгоритмов, а также обеспечивает масштабируемое распараллеливание аппроксимационного этапа, поскольку обработка каждого конечного элемента выполняется независимо от других. Аналогичные поэлементные технологии с вычислениями локальных матриц баланса разработаны и в методах конечных объемов, см. [81], [83].

Пусть, например, треугольный конечный элемент T_m^h имеет вершинами точки $V_p^m, p = 1, 2, 3$, координаты которых x_p^m, y_p^m могут быть найдены в ССД по номерам p, m . Предполагается также, что по значению m могут быть определены “материальные свойства среды” внутри T_m^h , т.е. коэффициенты исходного уравнения (2) и краевых условий (3) на гранях T_m^h , если они являются граничными сегментами расчетной области Ω . Тогда по этим данным с помощью некоторых формул, характеризующих коэффициентами решаемого исходного уравнения и видом алгоритма МКО или МКЭ, определяются элементы соответствующей локальной матрицы $A_m^{loc} = \{a_{i,j}^{(m)}; i, j = 1, 2, 3\}$, из которых, в свою очередь, складываются элементы глобальной матрицы задачи, определяемой с помощью скалярного произведения

$$(Au, v) = \sum_{m=1}^{N_T^h} \sum_{i,j=1}^3 a_{i,j}^{(m)} \bar{u}_i^{(m)} \bar{v}_j^{(m)}. \quad (6.16)$$

Здесь обозначены через $\bar{u}^{(m)} = \{u_1^{(m)}, u_2^{(m)}, u_3^{(m)}\}$ “локальные” подвекторы третьего порядка, а через u, v – “глобальные” подвекторы размерности, равной количеству узлов N_V^h . Компоненты локальных и глобальных векторов совпадают для одинаковых узлов, надо только установить взаимнооднозначное соответствие их локальных и глобальных номеров.

Таким образом, возникает задача формирования поэлементной сеточной структуры данных, состоящей из совокупности данных $EIN(m)$ о каждом T_m^h : номера вершин-узлов, номера соответствующей подобласти и граничных сегментов, которым принадлежат грани (ребра – в 2D) элемента T_m^h (соответствующий номер полагается равным нулю, если сеточная грань не лежит на границе расчетной области). В отличие от поузловой информации CSR, каждая m -я порция EIN содержит геометрическую и функциональную информацию только о своем элементе, и здесь формально никак не присутствуют межэлементные связи.

6.4. Примеры построения адаптивных квазиструктурированных сеток

Для иллюстрации возможных подходов к оптимизации сеток мы рассмотрим достаточно простую, но актуальную для нефтяной геофизики задачу. Пусть расчетная двумерная область представляет собой прямоугольник $\Omega = [a, b] \times [c, d]$, в котором размещены несколько круглых подобластей (для определенности берем два круга радиуса R с центрами в точках (x_1, y_1) и (x_2, y_2) , см. Рис. 6.4. Данная постановка представляет нефтеносный полигон со скважинами очень малого относительно радиуса, в окрестности которых распределения скоростей и давления имеют большие градиенты. Исходя из этих соображений, представляется целесообразным сгенерировать специальную квазиструктурированную дискретизацию, выделив околоскважинные подобласти, чтобы построить в них достаточно качественную треугольную сетку.

Рассмотрим для примера одну такую подобласть Ω_1 , ограничив ее квадратом с центром в точке (x_1, y_1) и длиной стороны 2Δ . Под область Ω_1 , в свою очередь, разобьем на две части: Ω'_1 – круг радиуса R и $\Omega''_1 = \Omega_1/\Omega'_1$, т.е. дополнение круга до квадрата. Внутри круга Ω'_1 строим сетку по радиальным слоям r_i , $i = 1, \dots, L_1$: $r_0 = 0, r_{L_1} = R$, $r_{i+1} = r_i + h_i^r$, где величины шагов h_i^r можно взять или постоянными, или как-то их варьировать из соображений желательного локального сгущения сетки, при сохранении качества треугольников. На первом слое строим 6 правильных треугольников, а на каждом последующем

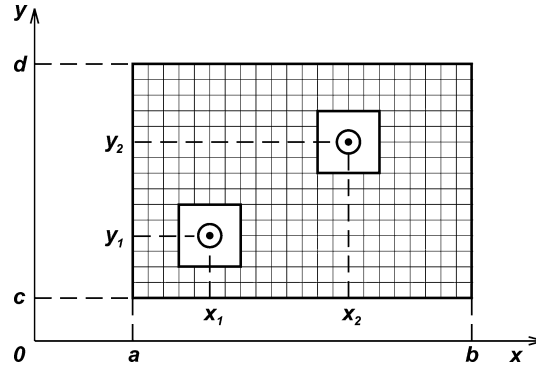


Рис. 6.4. Пример двумерной области с квазиструктурированной сеткой (нефтяной полигон со скважинами)

радиальном слое r_i располагаем равномерно по длине дуги узлы r_{i,j_i} , $j_i = j_{i-1} + 1, \dots, j_{i-1} + 6i$, $j_0 = 1$, как это показано на Рис. 6.5.

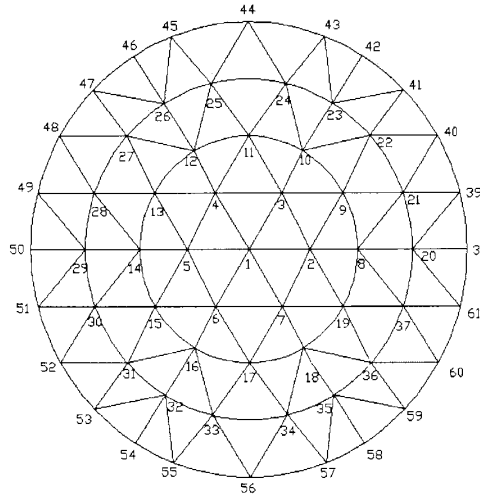


Рис. 6.5. Построение квазиравномерной треугольной сетки в круге

В итоге на окружности радиуса $r_{L_1} = R$ будем иметь $6L_1$ узлов, равномерно расположенных с шагом $\pi R/3L_1$. Эти узлы берем в качестве заданных для построения треугольной сетки в подобласти Ω_1'' , на внешней границе которой также ставим целью конструирование равномерной адаптивной сетки (верши-

ны квадрата должны являться узлами). В Ω_1'' нетрудно построить квазиравномерную треугольную сетку с узлами z_{i,j_i} , $i = L_1, \dots, L_2$ расположенными по слоям, постепенно переходящими из окружности в квадрат, которым соответствуют номера $i = L_1$ и $i = L_2$. На Рис. 6.6 представлена сетка в Ω_1'' , полученная с помощью простого алгоритма, в котором число узлов на каждом i -м слое равно $M_i = M_{i-1} + 4$, $M_{L_1} = 6L_1$ – количество узлов, построенных предварительно на окружности.

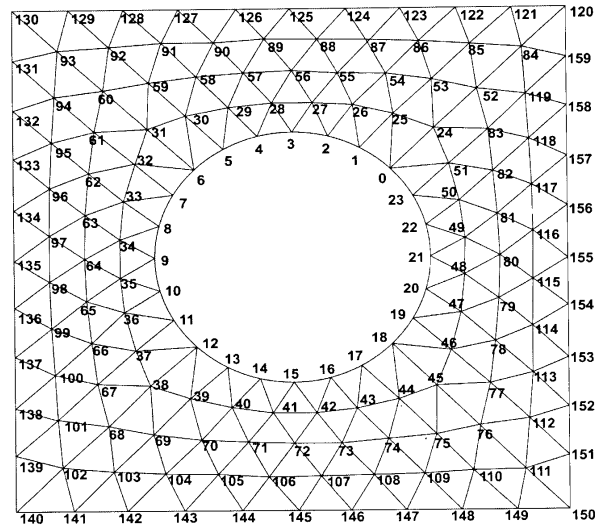


Рис. 6.6. Квазиравномерная треугольная сетка в подобласти “квадрат/круг”

Если в расчетной области Ω имеется P произвольно расположенных скважин, и для каждой из соответствующей околоскважинной подобласти Ω_p построена описанная выше сетка, то во внескважинном пространстве $\Omega / \bigcup_{q=1}^P \Omega_p$ уже нетрудно сгенерировать, например, кусочно-равномерную прямоугольную сетку.

Рассмотренный алгоритм может быть несложно перенесен и на трехмерные задачи, особенно если скважины вертикальные. В этом случае легко строится сетка из треугольных призм, и соответствующий пример представлен на Рис. 6.7.

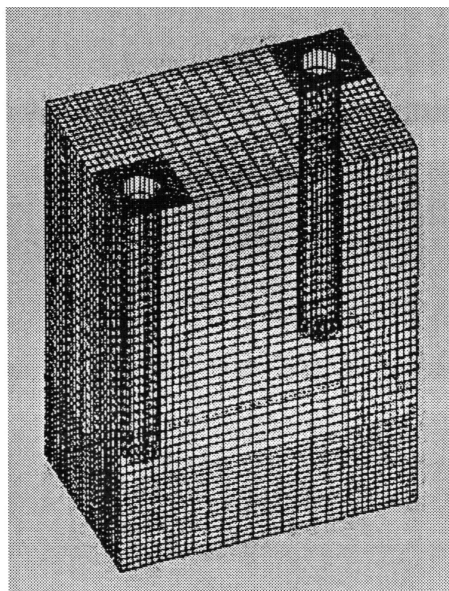


Рис. 6.7. Пример квазиструктурированной сетки в трехмерной области «нефтяной полигон со скважинами»

В заключение данного пункта сделаем замечание. Для построения сетки в круге очень легко построить равномерную сетку в локальной полярной системе координат. Однако такой способ является отнюдь не оптимальным, так как в этом случае при большом числе узлов вблизи полюса получаются слишком «вытянутые» по радиусу треугольники.

6.5. Задача декомпозиции сеточных областей

Мы рассмотрим *метод декомпозиции областей* (МДО), который имеет свою замечательную историю. Возникший на заре XIX-го века под названием *альтернирующий метод Шварца*, он долгое время служил чисто теоретическим средством исследования сложных краевых задач математической физики. Во второй половине XX-го века появилась его модификация – аддитивный метод декомпозиции, отличие которого от альтернирующего, являющегося по своему характеру мультипликативным, легко пояснить с помощью рис. 6.8.

Здесь расчетная L -образная область $\Omega = \Omega_1 \cup \Omega_2$, в кото-

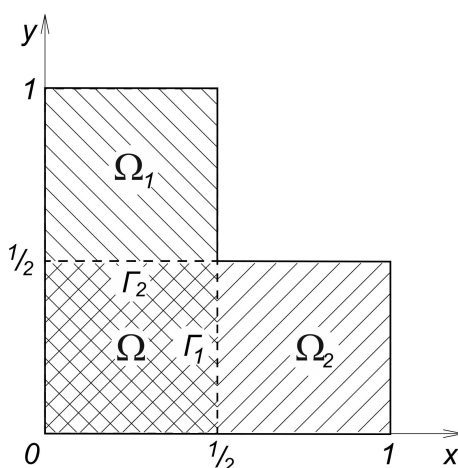


Рис. 6.8. Пример декомпозиции области на две подобласти

рой надо решить, например, задачу Дирихле для уравнения Лапласа, состоит из двух пересекающихся подобластей Ω_1 и Ω_2 , у которых на “внутренних” границах Γ_1 и Γ_2 граничные условия итерируются при решении вспомогательных задач в Ω_1 и Ω_2 . В мультипликативном, или альтернирующем, варианте эти подзадачи решаются последовательно одна за другой, а в аддитивном – одновременно.

После появления многопроцессорных ЭВМ аддитивный подход стал главным орудием распараллеливания, когда область Ω разбивается на большое количество подобластей Ω_k , $k = 1, \dots, P$, $P \gg 1$, в которых вспомогательные подзадачи решаются синхронно на “своих” процессорах. При этом технологически расчетная область Ω предварительно дискретизируется, так что фактически происходит декомпозиция сеточной области Ω на сеточные подобласти Ω_k .

В результате МДО стал бурно развиваемым направлением вычислительной математики, по которому имеется огромное количество специальной литературы и проходят регулярные международные конференции (см. в сети Интернет сайт DDM.org, от Domain Decomposition Method), поскольку при решении больших задач математического моделирования распараллеливание на МВС стало обязательным атрибутом.

Чтобы представить многообразие возникающих в данной об-

ласти идей и алгоритмов, мы сошлемся только на три монографии специалистов, внесших существенный вклад в развитие данной актуальной и по-прежнему перспективной тематики: В.И. Лебедева [106], О. Видлунда [179] и Ф. Натафа [64] с коллегами.

В данном пункте мы кратко рассмотрим нетрадиционную задачу дискретной математики, актуальную с точки зрения распараллеливания алгоритмов, – сбалансированное разбиение сеточной расчетной области на сеточные подобласти (с приблизительно равным количеством узлов в каждой). Цель такой декомпозиции заключается в том, что исходная задача разбивается на относительно независимые подзадачи – каждая для своей подобласти, – которые решаются одновременно на соответствующих процессорах, что в идеале дает ускорение вычислений, близкое к линейному, т.е. пропорциональному числу подобластей. Для полного решения рассматриваемой задачи между подобластями необходимы информационные обмены, что приводит к неизбежным коммуникационным потерям.

Возникающие в данном случае вопросы составляют актуальное направление вычислительных методов и технологий для решений больших задач (с числом узлов 10^8 – 10^{11}) на многопроцессорных вычислительных системах. Этим проблемам посвящено огромное число публикаций, из которых мы назовем только монографии последних лет [64], [179].

Проблема сбалансированного разбиения сеточной области не представляет труда для пользователя (или, выражаясь фигурально, – для естественного интеллекта), если он визуально наблюдает изображение сетки и может с помощью графических средств приближенно сделать это практически не задумываясь. Однако задача автоматической декомпозиции области для неструктурированной сетки с узлами, геометрические связи между которыми заданы фактически с помощью алгебраических CSR-форматов, содержащих информацию о ненулевых элементах больших разреженных матриц, требует анализа топологии сеточных графов, а реализация их расщепления совсем не тривиальна. В связи с этим даже терминологически различаются геометрические и алгебраические методы декомпозиции областей (МДО), которые сейчас являются основным орудием масштабируемого распараллеливания алгоритмов на многопроцессорных вычислительных системах. Можно сказать, что рассмат-

риваемая постановка представляет актуальную проблему искусственного интеллекта. Формализация данных подходов может быть представлена следующим образом.

Рассмотрим сеточную расчетную область Ω с границей Γ , разделенную на P непересекающихся (nonoverlapping) сеточных подобластей $\Omega_q \subset \Omega$ с границами Γ_q (здесь и далее индексы “ h ” у сеточных объектов опускаем):

$$\bar{\Omega} = \bigcup_{q=1}^P \bar{\Omega}_q, \quad \bar{\Omega} = \Omega \cup \Gamma, \quad \bar{\Omega}_q = \Omega_q \cup \Gamma_q, \quad (6.17)$$

причем $\Omega_{q'} \cap \Omega_{q''} = 0$ при $q' \neq q''$. Для q -й подобласти обозначим через $\hat{\omega}_q$ совокупность номеров соседних, или смежных, подобластей, имеющих общую часть границы с $\bar{\Omega}_q$. Границу каждой из подобластей можно представить суммой из M_q ее частей (M_q – общее количество подобластей, смежных с Ω_q), по их принадлежности к соседним подобластям:

$$\Gamma_q = \bigcup_{r \in \hat{\omega}_q} \Gamma_{q,r} \equiv \bigcup_{q'=1}^{M_q} \Gamma_{q,q'}, \quad \Gamma_{q,r} = \bar{\Omega}_q \cap \bar{\Omega}_{q'}. \quad (6.18)$$

Расчётные узлы могут лежать как в некоторой подобласти Ω_q , так и на её границе Γ_q .

При упоминании подобластей вида (6.17) под Ω_0 (если это необходимо) мы формально всегда будем понимать “внешнюю” подобласть, т.е. дополнение расчетной области до всего пространства: $\Omega_0 = \mathcal{R}^d / \Omega$. В каждой подобласти будем выделять фрагменты границы $\Gamma_{q,r'}$ с индексами $r' = 0$ и $r' \neq 0$, составляющие соответственно внешнюю и внутреннюю границы:

$$\Gamma_q^0 = \bigcup \Gamma_{q,0}, \quad \Gamma_q^i = \Gamma_q / \Gamma_q^0.$$

Подобласти, у которых часть границы является внешней, называются околোগраничными, а остальные – внутренними.

Отметим, что для узлов из $\Gamma_{q,0}$ с граничными условиями 2-го или 3-го рода (совокупность таких узлов обозначим через $\Gamma'_{q,0}$) на этапе аппроксимации исходной задачи формируется “обычное” сеточное уравнение

$$a_{l,l}u_l + \sum_{m \in \omega_l} a_{l,m}u_m = f_l, \quad (6.19)$$

где $\omega_l \in \Omega_q$ есть сеточный шаблон для узла с номером l , т.е. фактически такие узлы из $\Gamma_{q,0}$ в СЛАУ никак не выделяются. С другой стороны, если граничный узел является “дирихлёвым” (множество таких узлов обозначаем $\Gamma''_{q,0}$, причем $\Gamma_{q,0} = \Gamma'_{q,0} \cup \Gamma''_{q,0}$), то соответствующее сеточное уравнение имеет тривиальный вид $u_l = \text{const}$. В таких случаях практически всегда делается (и мы всегда эту процедуру предполагаем, хотя она, вообще говоря, не является обязательной) исключение условий Дирихле (constraining): значение u_l исключается из всех соседних узлов, у которых корректируется правая часть и затем зануляется соответствующий коэффициент, т.е. при $m \in \omega_l$ и $l \in \Gamma''_{q,0}$ имеем:

$$f_m := f_m - a_{m,l}u_l, \quad a_{m,l} := 0.$$

Таким образом, узлы из $\Gamma_{q,0}$ при этом в СЛАУ не входят.

Важно подчеркнуть, что формулы (6.17)–(6.18) определяют декомпозицию с явным определением внутренних сеточных границ $\Gamma_{q,r}$ между смежными подобластями. Совокупность таких границ-разделителей, не входящих ни в одну открытую подобласть, образуют каркас, или скелет, разделенных подобластей. Этот сеточный каркас, в свою очередь, можно сам рассматривать как подобласть. Для определенности дадим ей последний $(P + 1)$ -й номер.

Пример такого сеточного разделительного каркаса для двумерного случая приведен на рис. 6.9.

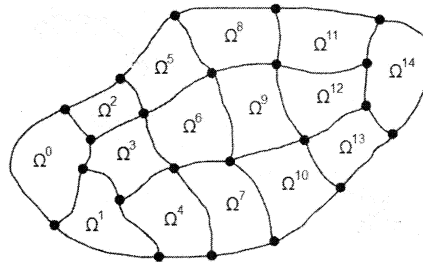


Рис. 6.9. Пример формирования разделительного каркаса для декомпозиции двумерной области

Рассмотрим сначала сеточную макроструктуру данных для простейшего типа СЛАУ, когда все уравнения являются узловыми и каждому l -му узлу сетки соответствует одна векторная компонента u_l , $l = 1, \dots, N$. В данном случае l является глобальным номером узла и соответствующих векторных компонент u_l, f_l . Отметим, что при переупорядочивании узлов сетки и соответствующих строк матрицы A ее структура меняется, но межузловые связи в сетке при этом сохраняются, т.е. они в некотором смысле являются инвариантом преобразования перестановок. Можно также сказать, что матричный портрет определяет некоторый граф, который является изоморфизмом по отношению к сеточной структуре. Для наглядной иллюстрации взаимосвязи матричных и сеточных структур мы рассмотрим модельную СЛАУ, представляющую собой простейшую семиточечную аппроксимацию задачи Дирихле для уравнения Пуассона в декартовой системе координат в кубической расчетной области на кубической же сетке с шагом $h = 1/(M + 1)$, для которой коэффициенты уравнений вида (6.19) при “естественной” глобальной нумерации узлов определяются следующим образом:

$$a_{l,l} = 6, \quad a_{l,m} = -1 \text{ или } 0, \\ l, m \in \omega_l = (l, l + 1, l - 1, l + M, l - M, l + M^2, l - M^2),$$

причем внедиагональные коэффициенты $a_{l,m}$ равны нулю, если l -й узел сетки в соответствующем координатном направлении примыкает к границе Γ .

Пусть сетка имеет по M узлов вдоль каждой координатной линии и $N = M^3$. Предполагаем, что макросетка из подобластей также кубическая и $P = M_1^3$. С учетом наличия сеточного разделительного каркаса $((P + 1)$ -й подобласти) справедливы соотношения

$$M = M_1 L_1 + M_1 - 1, \quad N_s = L_1^3, \quad s = 1, \dots, P, \quad (6.20)$$

$$N_{P+1} = 3L_1^2(M_1 - 1) + 3L_1(M_1 - 1) + (M_1 - 1)^3, \quad (6.21)$$

где N_s есть число узлов в s -й подобласти. В примере на рис. 6.9 для двумерного случая координатные линии “обычной” сетки нарисованы тонкими линиями, а макросетки (каркаса) – толстыми, причем узлы подобласти-разделителя обозначены крестиками.

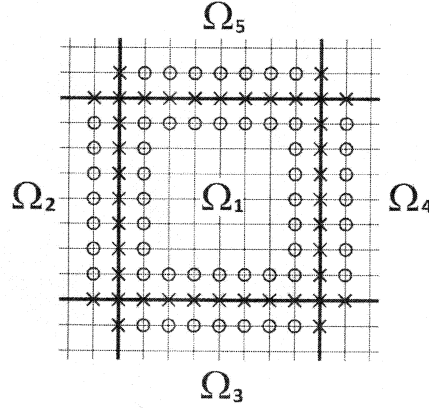


Рис. 6.10. Декомпозиция прямоугольной области с разделительной сеточной подобластью

Предположим, что глобальная нумерация узлов сеточной расчетной области, разбитой на P непересекающихся подобластей, выбрана следующим образом: сначала идут все узлы 1-й подобласти, затем – 2-й, и так до P -й подобласти, а в заключение нумеруются узлы $(P + 1)$ -й подобласти, т.е. сеточного каркаса декомпозиции с разделителями. При соответствующей упорядоченности элементов СЛАУ ее матрица принимает следующий блочный стрелообразный (стрела направлена из левого верхнего угла в правый нижний) вид:

$$A = \{A_{q,r}\} = \left[\begin{array}{ccc|c} A_{1,1} & & 0 & A_{1,P+1} \\ & \ddots & & \vdots \\ 0 & & A_{P,P} & A_{P,P+1} \\ \hline A_{P+1,1} & \cdots & A_{P+1,P} & A_{P+1,P+1} \end{array} \right]. \quad (6.22)$$

Другой возможный способ декомпозиции сеточной области – без формирования разделительных узлов. В этом случае можно представить, в отличие от определения $\Gamma_{q,q'} = \bar{\Omega}_q \cap \bar{\Omega}_{q'}$ из формулы (6.18), что геометрическая граница $\Gamma_{q,q'}$ проходит между узлами сетки: с одной стороны от нее находятся узлы Q_l из Ω_q , и с другой – из $\Omega_{q'}$. Формально это записывается следующим образом:

$$\Gamma_{q,q'} = \bigcup_{l'} \text{при } l' \in \hat{\omega}_l, \quad l \in \Omega_q, \quad l' \in \Omega_{q'},$$

$$\Gamma_{q',q} = \bigcup_l Q_l \text{ при } l \in \hat{\omega}_{l'}, \quad l \in \Omega_q, \quad l' \in \Omega_{q'}.$$

Соответствующий пример декомпозиции приведен на Рис. 6.11.

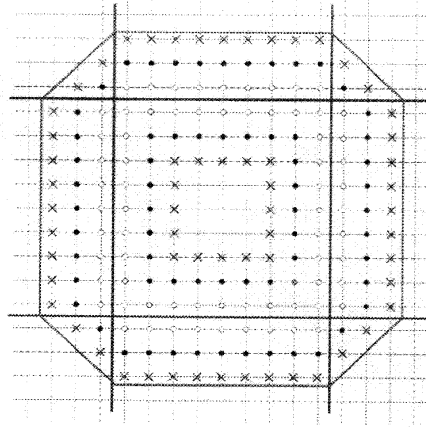


Рис. 6.11. Декомпозиция двумерной сеточной области без узлов-разделителей

Здесь напрашивается аналогия с интегральными граничными уравнениями для потенциалов простого и двойного слоя: в варианте с определением граничных разделителей на Рис. 6.11 узлы из $\Gamma_{q,q'}$ образуют простой сеточный слой, а в альтернативном варианте без разделителей совокупность узлов из $\Gamma_{q,q'}$ и $\Gamma_{q',q}$ образуют двойной сеточный слой.

В данном случае матрица A при “естественной” (лексикографической) нумерации подобластей будет иметь уже не экзотический вид (6.24), а ленточную блочную форму (например, блочно-трехдиагональную или блочно-пятидиагональную).

Необходимо также отметить, что при построении “расширения” подобласти (в случае применения МДО с перекрытием подобластей) принципиальная разница между каркасным и бескаркасным делением области заключается в том, что при каркасном делении расширяться будет только сам каркас, таким образом “наезжая” на соседние подобласти (например, для случая на рис. 6.10 каркас будет включать не только узлы, помеченные крестиками, но и узлы, помеченные кружочками), а при бес-

каркасном – расширяются все подобласти послойно (в смысле соседства узлов по аппроксимационному шаблону).

При всей алгоритмической сложности и многообразии современных алгоритмов декомпозиции областей, основой параллельного решения рассматриваемых СЛАУ (в алгебраических терминах) является блочный итерационный метод Якоби, или другими словами (в геометрическом смысле) – аддитивный метод Шварца. Если обозначить через $\hat{u}_q, \hat{f}_q$, $q = 1, \dots, P + 1$, подвекторы, соответствующие подобласти Ω_q , или соответствующей блочной строке матрицы A , то такой алгоритм записывается в виде

$$A_{q,q}\hat{u}_q^n = f_q - \sum_{r \in \hat{\omega}_q} A_{q,r}\hat{u}_r^{n-1} \equiv \hat{f}_q, \quad q = 1, \dots, P; \quad n = 1, 2, \dots, \quad (6.23)$$

где n есть номер итерации, а каждый диагональный блок $A_{q,q}$ – фактически матрица автономной вспомогательной СЛАУ для подобласти Ω_q . Внедиагональные же блоки $A_{q,r}$, $r \neq q$, отвечают за взаимосвязи между подобластью Ω_q и контактирующими с ней подобластями, имеющими номера r из множества $\hat{\omega}_q$. Итерационный процесс (6.23) легко формально обобщить, добавив к левой и правой частям формулы члены с некоторым одинаковым матричным множителем:

$$B_{q,q}\hat{u}_q^n \equiv (A_{q,q} + C_{q,q})\hat{u}_q^n = f_q + C_{q,q}\hat{u}_q^{n-1} - \sum_{r \in \hat{\omega}_q} A_{q,r}\hat{u}_r^{n-1}, \quad n = 1, 2, \dots \quad (6.24)$$

Для иллюстрации возможной реализации интерфейсных условий мы приводим на рис. 6.12 пример расположения узлов вблизи внутренних границ для контактирующих подобластей.

Очевидно, что если при $n \rightarrow \infty$ итерационные процессы (6.23), (6.24) сходятся, то предельный вектор $u = \{u_q\}$ у них один и тот же, а именно – решение исходной СЛАУ. Вводимые дополнительные блоки $C_{q,q}$ не меняют структуру и логику итерационного алгоритма, если они не изменяют портрет матрицы $B_{q,q}$, в сравнении с $A_{q,q}$. Обычно добавляются члены только к сеточным уравнениям для околограничных узлов из Ω_q , и эта процедура соответствует изменению интерфейсных, или краевых, итерируемых условий между контактирующими подобластями. В частности, если положить $C_{q,q} = \theta D_{q,q}$, где $D_{q,q}$ – диагональная матрица, каждый элемент которой равен сумме внедиагональных элементов матриц $A_{q,r}$, $r \in \hat{\omega}_q$, то при $\theta = 0$ интерфейс

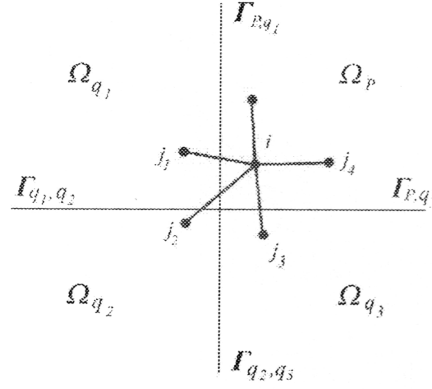


Рис. 6.12. Иллюстрация к реализации интерфейсных условий для узлов около внутренних границ смежных подобластей

между смежными подобластями соответствует условию Дирихле, а при $\theta = 1$ – условию Неймана.

Подчеркнем, что мы для простоты рассматриваем в (6.24) декомпозицию без пересечения подобластей, чтобы выделить главные алгебро-геометрические особенности алгоритмов. При наличии пересечений структура методов несколько усложняется. Например, блочное представление СЛАУ будет содержать перекрывающиеся блоки матрицы A и подвекторов. Само построение пересечений целесообразно делать последовательно: сначала осуществляется конструирование подобластей Ω_q без пересечения, а затем каждая из подобластей постепенно расширяется путем включения в нее по одному сеточному слою на очередном этапе. Величину пересечений естественно измерять параметром Δ – числом расширяющих сеточных слоев подобласти Ω_q , который для простоты будем считать одинаковым для всех q . Последовательность формируемых границ можно описать следующим образом:

$$\begin{aligned} \Gamma_q &\equiv \Gamma_q^0 = \{l' \in \hat{\omega}_l, \ l \in \Omega_q, \ l' \notin \Omega_q, \ \Omega_q^1 = \bar{\Omega}_q^0 = \Omega_q \cup \Gamma_q^0\}, \\ \Gamma_q^t &= \{l' \in \hat{\omega}_l, \ l \in \Omega_q^{t-1}, \ l' \in \Omega_q^{t-1}, \ \Omega_q^t = \bar{\Omega}_q^{t-1} = \Omega_q^{t-1} \cup \Gamma_q^{t-1}\}. \end{aligned}$$

В частности, на рис. 6.10 изображен пример декомпозиции области с пересечением контактирующих подобластей величиной в три сеточных слоя. А на Рис. 6.9 кружочками обозначены узлы

из Γ_{P+1}^0 , граничные к подобласти - разделителю. В дальнейшем расширенную подобласть с числом расширяющих слоев Δ будем обозначать через Ω_q^Δ .

Декомпозиция сетки осуществляется в два этапа. На первом разбиение осуществляется на непересекающейся подобласти, а на втором каждая из них расширяется на некоторое количество сеточных слоев, или фронтов. Иллюстрация такого расширения приводится на рис. 6.10 для простого случая квадратной сетки.

Необходимость такой декомпозиции с пересекающимися подобластями обуславливается скоростью сходимости итерационного метода Шварца, на котором основываются параллельные вычислительные процессы, см [64], [179].

Мы не будем останавливаться на алгоритмах решения данных логически сложных задач, а отметим только, что трудоемкость второго этапа (расширение уже построенных непересекающихся подобластей) значительно повышается в параллельном варианте, когда построенные на первом этапе сеточные подобласти располагаются в памяти различных процессоров. Данное обстоятельство является необходимым условием решения сверхбольших сеточных задач.

6.6. Принципы построения сеточных генераторов

Состав функциональных модулей библиотеки формально осуществляет преобразование геометрической и функциональной структур данных в сеточную структуру данных, т.е. реализует отображение

$$\Gamma CD + \Phi CD \rightarrow CCD.$$

Исходная информация полностью и однозначно описывает математическую постановку решаемой задачи, а итоговая — представляет все необходимые данные для реализации ее сеточной аппроксимации с помощью МКО или МКЭ и для формирования систем линейных или нелинейных алгебраических уравнений (СЛАУ или СНАУ).

Предполагается, что вычислительный процесс этапа генерации сетки выполняется в автоматическом или полуавтоматическом режиме. Последнее означает, что для частных подзадач пользователем могут указываться применяемые алгоритмы из состава библиотеки, а также их счетные параметры. В случае

отсутствия таких данных они выбираются по умолчанию. Кроме того, реализация скомпилированной исполняемой программы подразумевает визуальный контроль результатов, т.е. графическое и текстовое представление исходных данных, а также изображение конструируемой сетки и выдачу количественных характеристик ее качества. Важно также иметь в виду, что при решении задач или с динамическим изменением геометрии расчетной области, или с качественной перестройкой свойств моделируемого решения в процессе счета сетка может неоднократно модифицироваться.

Напомним, что исходными данными для построения сетки являются массивы данных по геометрическим и функциональным объектам, в том числе:

- описание расчетной области со списками макровершин, макроребер, макрограней и подобластей, а также сегментов граничных поверхностей;
- информация о коэффициентах решаемых уравнений и краевых условиях, с их привязкой к подобластям и граничным сегментам;
- дополнительная необходимая информация по методам построения или по модификациям сеток, характеризующая как общее число узлов, так и особенности их распределения в подобластях.

Создаваемая ССД должна содержать поэлементную информацию обо всех T_m^h с ее привязкой к узловым и функциональным данным, обеспечивающим выполнение последующего этапа аппроксимации.

Ориентация на решение больших задач математического моделирования на МВС определяет актуальную проблему формирования данных о сеточных сбалансированных подобластях, распределенных по различным процессорам. С одной стороны, это является главным средством масштабируемого распараллеливания алгоритмов (вспомогательные подзадачи в подобластях рассчитываются на соответствующих процессорах), а с другой – обеспечивает интеграцию вычислительных ресурсов для создания программных реализаций практических без ограничений на размер генерируемой сетки. Если для нумерации сеточных объектов использовать стандартное 64-битовое машинное

слово, то их допустимое количество в одной подобласти определяется объемом памяти процессора, а суммарное число узлов может быть увеличено фактически неограниченно за счет использования необходимого количества процессоров. Например, при оперативной памяти процессора в 4 Гбайта одна подобласть может насчитывать до $2 \cdot 10^5$ неизвестных, а на МВС с $5 \cdot 10^5$ процессорами их общее количество сможет достигать 10^{11} .

На мировом рынке программных продуктов имеется достаточно большое количество коммерческих или общедоступных прикладных пакетов по генерации многомерных сеток. В работе В.П.Ильина DELAUNAY: технологическая среда генерации сеток.–СибЖИМ, т. 16, N 2 (54), 2013, 83-97. Один из основных принципов организации библиотеки DELAUNAY заключается в возможности переиспользования внешних программных разработок, что означает, с одной стороны, овеществление огромного мирового интеллектуального потенциала, заключенного в существующих продуктах, а с другой – интегрирование результатов различных групп разработчиков. Такими иллюстрациями могут служить “чужие” сеточные генераторы для отдельных подобластей, использование для визуализации существующих прекрасных графических систем, а также согласованное взаимодействие с САПР-овскими системами (CAD, CAE, CAM и др.). Способы достижения этих целей – это наличие множественных представлений данных и переходников (конверторов), а также система гибких настраиваемых интерфейсов.

Главное назначение предлагаемой библиотеки – это функционирование в составе какого-либо приложения, осуществляющего решение задач моделирования в определенной области и реализуемого, например (но не только), конфигурационными инструментариями БСМ.

Эффективная реализация алгоритмических и системных компонент библиотеки представляется наиболее естественной с помощью средств объектно-ориентированного программирования, например, на языках C++ или FORTRAN 90.

Принципиальным моментом существования библиотеки DELAUNAY является ее открытость (Open Source) как инструментальной среды, из которой заинтересованный пользователь может брать необходимые ему компоненты, а внешний разработчик, наоборот, может вносить свои модули, расширяющие функциональность алгоритмического окружения. В этом смыс-

ле данная разработка рассматривается не как проект отдельной группы (group project), а как общественно развиваемый продукт (community project), разумеется, на основе некоторой системы регламентов. Данная концепция не противоречит созданию и коммерческих закрытых продуктов на основе общедоступных инструментариев, но при соблюдении определенной дисциплины и соглашений. Более того, такая парадигма хорошо согласуется с современным принципом открытых инноваций [197].

В заключение можно подчеркнуть двойственный характер предлагаемой концепции. С одной стороны, средства генерации сеток формируются как актуальная составляющая общей технологической цепочки математического моделирования в рамках БСМ, а с другой — библиотека DELAUNAY по своему замыслу является самостоятельным программным продуктом, который может автономно использоваться в других прикладных разработках. Однако в любом случае задачей сеточного генератора является создание комфортной информационной среды для функционирования следующего технологического этапа — реализации аппроксимационных подходов (на основе ССД, ГСД и ФСД) к построению дискретных систем алгебраических уравнений. Архитектурные принципы соответствующей технологической среды представлены в работе Д.С.Бутюгина, В.П.Ильина. CHEBYSHEV: принципы автоматизации построения алгоритмов в интегрированной среде для сеточных аппроксимаций начально-краевых задач.—Труды Международной конференции ПАВТ'2014. Челябинск, изд. ЮУрГУ, 2014, 42-50.

Литература

- [1] *Агошков В.И., Дубовский П.Б., Шутяев В.П.* Методы решения задач математической физики. — М.: Изд. ИВМ РАН, 2001.
- [2] *Аксельсон О., Баркер В.А.* (Axelsson O., Barker V.A.) Finite element solution of boundary value problems: Theory and computations. — New York: Acad. Press, 1984.
- [3] *Акулич А.Н., Смирнов Н.Н., Тюренкова В.В., Лапко В.А., Галкин В.А.* Математическое моделирование трещины гидроразрыва. — Сургут: ИЦСурГУ, 2016.
- [4] *Альтшуллер Г.С.* Найти идею. — Новосибирск: Наука, 1991.
- [5] *Арнольд В.И.* Математические методы классической механики. — М.: Наука, 1970.
- [6] *Арнольд В.И.* Что такое математика. — М.: МЦНМО, 2004.
- [7] *Арсеньев А.А., Самарский А.А.* Что такое математическая физика. — М.: Знание, 1983.
- [8] *Ахромеева Т.С., Курдюмов С.П., Малинецкий Г.Г., Самарский А.А.* Нестационарные структуры и диффузионный хаос. — М.: Наука, 1992.
- [9] *Ачиманов С.А.* Введение в математическую экономику. — М.: Наука, 1984.
- [10] *Бабушка И.М. и др.* (Babuska I.M., Nobile F., Oden J.T., Tempone R.) Reability, Uncertainty, Estimates Validation and Verification. — 2004. — (Tech. Rep. / ECES. Univ of Texas, Austin; N 04-05).
- [11] *Багаев Б.М., Кареева Е.Д., Шайдулов В.В.* Сеточные методы решения задач с пограничным слоем. Ч. 2. — Новосибирск: Наука, 2001.
- [12] *Багаев Б.М., Шайдулов В.В.* Сеточные методы решения задач с пограничным слоем. Ч. 1. — Новосибирск: Наука, 1998.
- [13] *Базаров М.Б., Сафаров И.И., Шокин Ю.И.* Численное моделирование колебаний диссипативно однородных и неоднородных механических систем. — Новосибирск: Изд-во СО РАН, 1998.
- [14] *Базыкин А.Д.* Математическая биофизика взаимодействующих популяций. — М.: Наука, 1985.

-
- [15] *Баренблатт Г.И., Ентов В.М., Рыжик В.М.* Движение жидкостей и газов в природных пластах. — М.: Недра, 1984.
- [16] *Басниев К.С., Власов А.М., Кочина П.Н., Максимов В.М.* Подземная гидравлика. — М.: Недра, 1986.
- [17] *Бахвалов Н.С., Корнеев А.А., Чижонков Е.В.* Численные методы решения задач и упражнения. — М.: Дрофа, 2009.
- [18] *Бежаев А.Ю., Василенко В.А.* (Bezhaev A.Yu., Vasilenko V.A.) Variational theory of splines. — New York: Kluwer Academic, 2001.
- [19] *Белолитецкий В.М., Генова С.Н., Туговиков В.Б., Шокин Ю.И.* Численное моделирование задач гидроледотермики водотоков. — Новосибирск: Наука, 1994.
- [20] *Белоцерковский О.М.* Численное моделирование в механике сплошных сред. — М.: Наука, 1994.
- [21] *Белоцерковский О.М., Лифанов О.М.* Численные методы в сингулярных интегральных уравнениях. — М.: Наука, 1985.
- [22] *Бенерджи П., Баттерфилд Р.* Методы граничных элементов в прикладных науках. — М.: Мир, 1984.
- [23] *Березин Ю.А., Вишнев М.А.* Метод частиц в динамике разреженной плазмы. — Новосибирск: Наука, 1980.
- [24] *Березин И.С., Жидков Н.П.* Методы вычислений. Т. 1. — М.: Физматлит, 1962.
- [25] *Бояринцев Ю.Е., Орлова И.В.* Пучки матриц и алгебро-дифференциальные системы. — Новосибирск: Наука, 2006.
- [26] *Бреннер С.К., Скотт Л.Р.* (Brenner S.C., Scott L.R.) The Mathematical Theory of Finite Element Methods. — New York: Springer-Verlag, 2008. — (Texts in Applied Mathematics; Vol. 15).
- [27] *Бреци Н., Фортин М.* (Brezzi N., Fortin M.) Mixed and Hybrid Element Method. — Series in Computational Mathematics. — Heidelberg: Springer, 1991.
- [28] *Бругнано Л., Иавернано Ф.* (Brugnano L., Iavernano F.) Line Integral Methods for Conservative Problems. — CRC Press, Taylor & Francis Group, N.Y., 2015.
- [29] *Булеев Н.И.* Пространственная модель турбулентного обмена. — М.: Наука, 1989.

-
- [30] *Вабищевич П.Н.* Численные методы решения задач со свободной границей. — М.: Изд-во МГУ, 1987.
- [31] *Вабищевич П.Н.* Метод фиктивных областей в задачах математической физики. — М.: Изд-во МГУ, 1991.
- [32] *Вабищевич П.Н.* Численное моделирование. — М.: Изд-во МГУ, 1993.
- [33] *Вабищевич П.Н.* Вычислительные методы математической физики. Стационарные задачи. — М.: «Вузовская книга», 2008.
- [34] *Вабищевич П.Н.* Аддитивные операторные разностные схемы. — М.: Красан, 2013.
- [35] *Вазов В., Форсайт Дж.* Разностные методы решения дифференциальных уравнений в частных производных. — М.: ИЛ, 1963.
- [36] *Василенко В.А., Зюзин М.В., Ковалков А.В.* Сплайн-функции и цифровые фильтры. — Новосибирск: Изд. ВЦ СО РАН, 1984.
- [37] *Васильев В.И., Максимов А.М., Петров Е.Е., Цыпкин Г.Г.* Тепломассоперенос в промерзающих и протаивающих грунтах. — М.: Наука, 1996.
- [38] *Васильев В.И., Попов В.В., Типофеева Т.С.* Вычислительные методы в разработке месторождений нефти и газа. — Новосибирск: Изд-во СО РАН, 2000.
- [39] *Власов В.И.* Краевые задачи в областях с криволинейной границей. — М.: Изд-во ВЦ АН СССР, 1987.
- [40] *Воеводин В.В., Кузнецов Ю.А.* Матрицы и вычисления. — М.: Наука, 1984.
- [41] *Галкин В.А.* Анализ математических моделей. Системы законов сохранения. Уравнения Больцмана и Смолуховского. — М.: БИНОМ. Лаборатория знаний, 2009.
- [42] Гатика Г.Н. (Gatica G.N.) A Simple Introduction to the Mixed Finite Element Methods. Theory and Applications. — Heidelberg: Springer Briefs in Mathematics, 2014.
- [43] *Гельфонд А.О.* Исчисление конечных разностей. — М.: Наука, 1967.

-
- [44] Гергель В.П., Стронгин Р.Г., Городецкий С.Ю., Гришагин В.А., Маркина М.В. Современные методы принятия оптимальных решений. – Нижний Новгород: Изд-во ННГУ, 2002.
- [45] Глассер А.Г., Лисейкин В.Д., Шокин Ю.И., Васева И.А., Лиханова Ю.В. Построение разностных сеток с помощью уравнений Бельтрами и диффузии. – Новосибирск: Наука, 2006.
- [46] Годунов С.К., Забродин А.В., Иванов М.Я., Крайко А.Н., Прокотов Г.П. Численное решение многомерных задач газовой динамики. – М.: Наука, 1976.
- [47] Годунов С.К., Киселев С.П., Куликов И.М., Мам В.И. Моделирование ударно-волновых процессов в упругопластических материалах на различных (атомный, мезо и термодинамический) структурных уровнях. – М.: Ижевск, Институт компьютерных исследований, 2014.
- [48] Годунов С.К., Роменский Е.И. Элементы механики сплошных сред и законы сохранения. – Новосибирск: Научная книга, 1998.
- [49] Годунов С.К., Рябенский В.С. Разностные схемы (введение в теорию). – М.: Наука, 1977.
- [50] Головизнин В.М., Зайцев М.А., Карабасов С.А., Короткин Н.А. Новые алгоритмы вычислительной гидродинамики для многопроцессорных вычислительных комплексов. – М.: Изд-во МГУ, 2013.
- [51] Гребенников А.И. Метод сплайнов и решение некорректных задач теории приближений. – М.: Изд. МГУ, 1983.
- [52] Григорьев А.Д. Методы вычислительной электродинамики. – М.: ФИЗМАТЛИТ, 2012.
- [53] Григорьев Ю.Н., Вишнев В.А., Федорук М.П. Численное моделирование методами частиц в ячейках. – Новосибирск: Изд-во СО РАН, 2004.
- [54] Гришин А.М., Зинченко В.И., Ефимов К.Н. и др. Интегроинтерполяционный метод и его приложения. – Томск: Изд. ТГУ, 2004.
- [55] Гришин А.М., Зинченко В.И., Кузин А.Я. и др. Решение некоторых обратных задач механики реагирующих сред. – Томск: Изд. ТГУ, 2006.

-
- [56] Данаев Н.Т., Корсакова Н.К., Пеньковский В.И. Массоперенос в прискважинной зоне и электромагнитный каротаж пластов. — Алматы: КазГУ, 2005.
- [57] Деккер К., Вервер Я. Устойчивость методов Рунге-Кутты для жестких нелинейных дифференциальных уравнений. — М.: Мир, 1988.
- [58] Дельфор М., Золесио Дж. (Delfour M., Zolesio J.) Shapes and Geometries. Metrics, Analysis, Differential Calculus, and Optimization. — Philadelphia: SIAM, 2011.
- [59] Дементьев А.Д., Назаров Л.А., Назарова Л.А. Прикладные задачи теории упругости. — Новосибирск: Изд. НГАУ, 2002.
- [60] Демиденко Н.Д., Потапов В.И., Шокин Ю.И. Моделирование и оптимизация систем с распределенными параметрами. — Новосибирск: Наука, 2006.
- [61] Демидович Б.П., Марон И.А., Шувалова Э.З. Численные методы анализа. — Санкт-Петербург: Изд. “Лань”, 2010.
- [62] Джанг Бо-нан (Bo-nan Jiang) The Least-Squares Finite Element Method. — Berlin: Springer, 1998.
- [63] Дерман Э. (Derman E.) The Future of Modelling. — [http : //www.ederman.com/new/docs/risk – futureofmodeling.html](http://www.ederman.com/new/docs/risk-futureofmodeling.html)
- [64] Долеан В., Джоливет П., Натаф Ф. (Dolean V., Jolivet P., Nataf F.) An Introduction to Domain Decomposition Methods: algorithms, theory and parallel implementataion. — Philadelphia: SIAM, 2015.
- [65] Дородницын А.А. Информатика: предмет и задачи. Кибернетика. Становление информатики. — М.: Наука, 1996.
- [66] Дулан Э., Миллер Дж, Шилдерс У. Равномерные численные методы решения задач с пограничным слоем. — М.: Мир, 1983.
- [67] Дымников В.П. Устойчивость и предсказуемость крупномасштабных атмосферных процессов. — М.: Изд. ИВМ РАН, 2007.
- [68] Дьяконов Е.Г. Минимизация вычислительной работы. — М.: Наука, 1989.
- [69] Дьяконов Е.Г. Разностные методы решения задач. Вып. 1 (стационарные задачи). — М.: Изд. МГУ, 1971.

-
- [70] *Ершов А.П.* Программирование – вторая грамотность. — Новосибирск, 1981. — (Препр. / АН СССР. Сиб.отд-ние ВЦ; N 293).
- [71] *Ершов А.П., Ильин В.П.* Пакеты программ — технология решения прикладных задач. — Новосибирск, 1978. — (Препр. / АН СССР. Сиб.отд-ние ВЦ; N 121).
- [72] *Жданов М.С.* Теория обратных задач и регуляризации в геофизике. — М.: Научный мир, 2007.
- [73] *Жиглявский А.А., Жилинскас А.Г.* Методы поиска глобального экстремума. — М.: Наука, 1991.
- [74] *Завьялов Ю.С., Леус В.А., Скороспелов В.А.* Сплайны в инженерной геометрии. — М.: Машиностроение, 1985.
- [75] *Задорин А.И.* Разностные схемы для задач с пограничным слоем. — Омск: Изд-во ОмГУ, 2002.
- [76] *Захаров В.Е., Манаков С.В., Новиков С.П., Питаевский И.П.* (Zakharov V.E., Manakov S.V., Novikov S.P., Pitaevsij I.P.) Theory of Solitons. — New York: Plenum Press, 1984.
- [77] *Зенкевич О., Морган К.* Конечные элементы и аппроксимации. — М.: Мир, 1986.
- [78] *Иваненко С.А.* Адаптивно-гармонические сетки. — М.: ВЦ РАН, 1997.
- [79] *Ильин В.П.* Численные методы решения задач электрофизики. — М.: Наука, 1985.
- [80] *Ильин В.П.* Вычислительная информатика: открытие науки. — Новосибирск: Наука, 1991.
- [81] *Ильин В.П.* Методы конечных разностей и конечных объемов для эллиптических уравнений. — Новосибирск: ИВМиМГ СО РАН, 2001.
- [82] *Ильин В.П.* Численный анализ. Ч.1. — Новосибирск: Изд. ИВМиМГ СО РАН, 2004.
- [83] *Ильин В.П.* Методы и технологии конечных элементов. — Новосибирск: Изд. ИВМиМГ СО РАН, 2007.
- [84] *Ильин В.П., Катешов В.А., Куликов Ю.В., Монастырский М.А.* Численные методы оптимизации эмиссионных электронно-оптических систем. — Новосибирск: Наука, 1987.

-
- [85] Ильин В.П., Кузнецов Ю.И. Трехдиагональные матрицы и их приложения. — М.: Наука, 1985.
- [86] Кабанихин С.И. Обратные и некорректные задачи. — Новосибирск: Сиб. науч. изд., 2009.
- [87] Калмыков С.А., Шокин Ю.И., Юлдашев З.Х. Методы интервального анализа. — Новосибирск: Наука, 1986.
- [88] Квасов Б.И. Изометрическая аппроксимация сплайнами. — Новосибирск: Изд. НГУ, 1998.
- [89] Kleppe A. (Kleppe A.) Software Language Engineering: Creating Domain-Specific Language Using Metamodels. — New York: Addison-Wesley, 2008. — 207 p.
- [90] Коновалов А.Н. Решение задач теории упругости в напряжениях. — Новосибирск: НГУ, 1979.
- [91] Коновалов А.Н. Задачи фильтрации многофазной несжимаемой жидкости. — Новосибирск: Наука, 1988.
- [92] Корнеев В.Г. Схемы метода конечных элементов высокой точности. — Л.: ЛГУ, 1972.
- [93] Коробейников В.П. Математическое моделирование катастрофических явлений природы. — М.: Знание, 1986.
- [94] Cottrell J., Hughes T., Bazilevs Y. (Cottrell J., Hughes T., Bazilevs Y.) Isogeometric Analysis. Towards Integration of CAD and FEA. — Singapore: John Wiley & Sons, 2009.
- [95] Кравченко В.Ф., Рвачев В.Л. Алгебра логики, атомарные функции и вейвлеты в физических приложениях. — М.: ФИЗМАТЛИТ, 2006.
- [96] Краснощечков П.С., Петров А.А. Принципы построения моделей. — М.: Изд. МГУ, 1983.
- [97] Крукиер Л.А., Чижин А.Л., Чижина Л.Г., Шабас И.Н. Моделирование гидрофизических процессов в водоемах с обширными районами мелководья. — Ростов-на-Дону: Изд. ЮФУ, 2009.
- [98] Kuksin S.B. (Kuksin S.B.) Analysis of Hamiltonian PDEs. — Oxford: Oxford University Press, 2000.
- [99] Куликовский А.Г., Погорелов Н.В., Семенов А.Ю. Математические вопросы численного решения гиперболических систем уравнений. — М.: Физматлит, 2001. — 607 с.

-
- [100] *Лаевский Ю.М.* Проекционно-сеточные методы решения двумерных параболических уравнений. — Новосибирск: Изд. ВЦ СО АН СССР, 1987.
- [101] *Лаевский Ю.М.* Метод конечных элементов решения многомерных параболических уравнений. — Новосибирск: Изд. НГУ, 1993.
- [102] *Лаевский Ю.М.* Метод конечных элементов (основы теории, задачи). — Новосибирск: Изд. НГУ, 1999.
- [103] *Лаймкюхлер Б., Райх С.* (Leimkuhler B., Reich S.) Simulating Hamiltonian Dynamics. — New York: Cambridge University Press, 2004.
- [104] *Лебедев В.И.* Функциональный анализ и вычислительная математика. — М.: ВИНТИ, 1994.
- [105] *Лебедев В.И.* Явные разностные схемы с переменными шагами по времени для решения жестких систем уравнений. — М., 1987. — (Препр. / ОВМ АН СССР; N 177).
- [106] *Лебедев В.И., Агошков В.И.* Операторы Пуанкаре–Стеклова и их приложения в анализе. — М.: Изд. ВИНТИ, 1983.
- [107] *Лебо Е.Г., Тиликин В.Ф.* Исследование гидродинамической неустойчивости в задачах ядерного термоядерного синтеза. — М.: Физматлит, 2006, 303 с.
- [108] *Левин Д.Я., Малюх Б.Н., Ушаков Д.М.* Энциклопедия PLM. — Новосибирск: Изд. Азия, 2008.
- [109] *Ли К.* Основы САПР (САД, САМ, САЕ). — С.-Петербург: Изд. Питер, 2004.
- [110] *Лисейкин В.Д.* Разностные сетки. Теория и приложения. — Новосибирск: Изд-во СО РАН, 2014.
- [111] *Лисейкин В.Д., Лиханова Ю.В., Шокин Ю.И.* Разностные сетки и координатные преобразования для численного решения сингулярно возмущенных задач. — Новосибирск: Наука, 2007.
- [112] *Лисейкин В.Д., Петренко В.Е.* Адаптивно-инвариантный метод численного решения задач с пограничными и внутренними слоями. — Новосибирск: ВЦ СО АН СССР, 1989.
- [113] *Лисейкин В.Д., Шокин Ю.И., Васева И.А., Лиханова Ю.В.* Технология построения разностных сеток. — Новосибирск: Наука, 2009.

-
- [114] Лоусон Ч., Хенсон Р. Численное решение задач метода наименьших квадратов. — М.: Наука, 1986.
- [115] Магомедов К.М., Холодов А.С. Сеточно-характеристические численные методы. — М.: Наука, 1988.
- [116] Мак-Кракен Д., Дорн У. Численные методы и программирование на Фортране. — М.: Мир, 1977.
- [117] Малоней Дж. (Maloney J.) Distributed COM Application Development Using Visual C++. — N.Y.: Prentice Hall, 1999.
- [118] Малышкин В.Э., Корнеев В.Д. Параллельное программирование мультимикомпьютеров. — Новосибирск: НГТУ, 2006.
- [119] Манин Ю.Н. Математика как метафора. — М.: Изд-во МЦНМО, 2008.
- [120] Мартыненко С.И. Многосеточная технология: Теория и приложения. — М.: ФИЗМАТЛИТ, 2015.
- [121] Марчук Г.И. Методы расчета ядерных реакторов. — М.: Атомиздат, 1958.
- [122] Марчук Г.И. Численное решение задач динамики атмосферы и океана. — Ленинград: Гидрометеиздат, 1974.
- [123] Марчук Г.И. Методы вычислительной математики. — М.: Наука, 1980.
- [124] Марчук Г.И. Методы расщепления и переменных направлений. — М.: ОВМ АН СССР, 1986.
- [125] Марчук Г.И. Сопряженные уравнения и анализ сложных систем. — М.: Наука, 1992.
- [126] Марчук Г.И. Наука управлять наукой. — Новосибирск: Изд-во СО РАН, 2014.
- [127] Марчук Г.И., Агошков В.И. Введение в проекционно-сеточные методы. — М.: Наука, 1981.
- [128] Марчук Г.И., Шайдунов В.В. Повышение точности решений разностных схем. — М.: Наука, 1979.
- [129] Массель Л.В., Болдырев Е.А., Горнов А.Ю. Интеграция информационных технологий в системных исследованиях энергетики. — Новосибирск: Наука, 2003.

-
- [130] *Мацочкин А.М.* Автоматизация триангуляции областей с гладкой границей при решении уравнений эллиптического типа. — Новосибирск, 1975. — (Препр. / ВЦ СО АН СССР; N 15).
- [131] *Михайлов Г.А., Войтишек А.В.* Численное статистическое моделирование. Методы Монте-Карло. — М.: Изд. Академия, 2006.
- [132] *Моисеев Н.Н.* Математические задачи системного анализа. — М.: Наука, 1981.
- [133] *Монк П.* (Monk P.) Finite Element Methods for Maxwell's Equations. — Oxford University Press, 2003.
- [134] *Новиков Е.А.* Явные методы для жестких систем. — Новосибирск: Наука, 1997.
- [135] *Новиков С.П., Тайманов И.А.* Современные геометрические структуры и поля. — М.: Изд. МЦНМО, 2005.
- [136] *Норри Д., де Фриз Ж.* Введение в метод конечных элементов. — М.: Мир, 1981.
- [137] *Обэн Ж.-П.* Приближенное решение эллиптических краевых задач. — М.: Мир, 1977.
- [138] *Оганесян Л.А., Ривкин В.Я., Руховец Л.А.* Вариационно-разностные методы решения эллиптических уравнений. П. — Дифференциальные уравнения и их применение. — Вильнюс: Изд. ИФМ АН ЛитССР, 1974. — Вып. 8.
- [139] *Оганесян Л.А., Руховец Л.А.* Вариационно-разностные методы решения эллиптических уравнений. — Ереван: Изд. АН Арм. ССР, 1979.
- [140] *Ольшанский М.А.* Лекции и упражнения по многосеточным методам. — М.: Физматлит, 2005.
- [141] *Павловский Ю.Н., Белотелов Н.В., Бродский Ю.И.* Имитационное моделирование. — М.: Изд-во "Академия", 2008.
- [142] *Патанкар С.* Численные методы решения задач теплообмена и динамики жидкости. — М.: Энергоатомиздат, 1984.
- [143] *Пененко В.В.* Методы численного моделирования атмосферных процессов. — Ленинград: Гидрометеиздат, 1981.
- [144] *Петров А.А.* Экономика. Модели. Вычислительный эксперимент. — М.: Наука, 1996.

-
- [145] *Петров А.А., Поспелов И.Г., Шананин А.А.* Опыт математического моделирования экономики. — М.: Энергоиздат, 1996.
- [146] *Попов Ю.П., Самарский А.А.* Вычислительный эксперимент. — Новое в жизни, науке, технике. Сер. Мат. кибернетика. — М.: Знание, 1983.
- [147] *Препарата Ф., Шеймос М.* Вычислительная геометрия: Введение. — М.: Мир, 1989. — 478 с.
- [148] *Рвачев В.Л.* Теория R-функций и некоторые ее приложения. — Киев: Наукова думка, 1982.
- [149] *Ривiere Б. (Riviere B.)* Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations. Theory and Implementation. — SIAM, Philadelphia, 2008.
- [150] *Рунтмайер Р.Д., Мортон К.* Разностные методы решения краевых задач. — М.: Мир, 1972.
- [151] *Рождественский Б.Л., Яненко Н.Н.* Системы квазилинейных уравнений и их приложения в газовой динамике. — М.: Наука, 1978.
- [152] *Роч П.* Вычислительная гидродинамика. — М.: Мир, 1980.
- [153] *Рябенский В.С.* Введение в вычислительную математику. — М.: ФИЗМАТЛИТ, 2008.
- [154] *Сабо Б.В., Бабушка И. (Szabo B.V., Babuska I.)*. Finite Element Analysis. — New York: John Wiley Sons, Inc., 1991.
- [155] *Садовская О.В., Садовский В.М.* Математическое моделирование в задачах механики сыпучих сред. — М.: Физматлит, 2008.
- [156] *Самарский А.А.* Теория разностных схем. — М.: Наука, 1983.
- [157] *Самарский А.А., Андреев В.Б.* Разностные методы для эллиптических уравнений. — М.: Наука, 1976.
- [158] *Самарский А.А., Вабищевич П.Н.* Численные методы решения задач конвекции-диффузии. — М.: Изд. УРСС, 2003.
- [159] *Самарский А.А., Вабищевич П.Н., Матус П.П.* Разностные схемы с операторными множителями. — Минск: ЗОА “ЦОТЖ”, 1998.
- [160] *Самарский А.А., Галактионов В.А., Курдюмов С.П., Михайлов А.П.* Режимы с обострением в задачах для квазилинейных параболических уравнений. — М.: Наука, 1987.

-
- [161] Самарский А.А., Гулин А.В. Численные методы. — М.: Наука, 1989.
- [162] Самарский А.А., Лазаров Р.Д., Макаров В.Л. Разностные схемы для дифференциальных уравнений с обобщенными решениями. — М.: Высш. школа, 1987.
- [163] Самарский А.А., Михайлов А.П. Математическое моделирование. — М.: Физматгиз, 2002.
- [164] Самарский А.А., Попов Ю.П. Разностные методы решения задач газовой динамики. — М.: Наука, 1981.
- [165] Самарский А.А. и др. Разностные схемы на нерегулярных сетках. — Минск: Изд. “Критерий”, 1996.
- [166] Сегерлинд Л. Применение метода конечных элементов. — М.: Мир, 1979.
- [167] Смелов В.В. Лекции по теории переноса нейтронов. — М.: Атомиздат, 1978.
- [168] Смирнов Ю.Г. Математические методы исследования задач электродинамики. — Пенза: Изд. ПГУ, 2009.
- [169] Соловейчик Ю.Г., Рояк М.Э., Перова М.Г. Метод конечных элементов для решения скалярных и векторных задач. — Новосибирск: Изд. НГТУ, 2007.
- [170] Стоян Ю.Г., Проценко В.С., Манько Г.П. и др. Теория R -функций и актуальные проблемы прикладной математики. — Киев: Наукова думка, 1986.
- [171] Стренг Г., Фикс Дж. Теория методов конечных элементов. — М.: Мир, 1977.
- [172] Сыровой В.А. Теория интенсивных пучков заряженных частиц. — М.: Энергоатомиздат, 2004.
- [173] Сьярле Ф. Метод конечных элементов для эллиптических задач. — М.: Мир, 1980.
- [174] Тахтаджян Л.А., Фаддеев Л.Д. Гамильтонов подход в теории солитонов. — М.: Наука, Физматлит, 1986.
- [175] Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. — М.: Наука, 1986.

-
- [176] *Тихонов А.Н., Костомаров Д.М.* Вводные лекции по прикладной математике. — М.: Наука, 1984.
- [177] *Тихонов А.Н., Самарский А.А.* Уравнения математической физики. — М.: Наука, 1966.
- [178] *Толстых А.И.* Компактные разностные схемы и их применение в задачах гидродинамики. — М.: Наука, 1990.
- [179] *Тоселли А., Видлунд О.* (Toselli A., Widlund O.) Domain Decomposition Methods. — Algorithms and Theory. — Springer Ser. in Comput. Math. — 2005. — Vol. 34.
- [180] *Тыртыхиников Е.Е.* Краткий курс численного анализа. — М.: Изд. ВИНТИ, 1994.
- [181] *Ушаков Д.М.* Введение в математические основы САПР. — Новосибирск: Изд. ЗАО “ЛЕДАС”, 2008.
- [182] *Федоренко Р.П.* Приближенное решение задач оптимального управления. — М.: Наука, 1978.
- [183] *Федоренко Р.П.* Введение в вычислительную физику. — М.: МФТИ, 1996.
- [184] *Фенг К., Кин М.* (Feng K., Qin M.) Symplectic Geometric Algorithms for Hamiltonian Systems. — New York: Springer, 2009.
- [185] *Флетчер К.* Вычислительные методы в динамике жидкостей. — М.: Мир, 1991.
- [186] *Флетчер К.* Численные методы на основе метода Галеркина. — М.: Мир, 1988.
- [187] *Фокс А., Пратт М.* Вычислительная геометрия. Применение в проектировании и на производстве. — М.: Мир, 1982.
- [188] *Форсайт Дж., Малькольм М., Молер К.* Машинные методы математических вычислений. — М.: Мир, 1980.
- [189] *Франк А.М.* Дискретные модели несжимаемой жидкости. — М.: Физматлит, 2001.
- [190] *Фурсиков А.В.* Оптимальное управление распределенными системами. Теория и приложения. — Новосибирск: Научная книга, 1999.
- [191] *тер Хаар Д.* Основы гамильтоновой механики. — М.: Наука, 1974. — 224 с.

-
- [192] Хайрер Э., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Жесткие и дифференциально-алгебраические задачи. — М.: Мир, 1999.
- [193] Хайрер Э., Любих К., Ваннер Г. (Yairer E., Lubich C., Banner G.) Geometric Numerical Integration. — Berlin: Springer, 2006.
- [194] Хайрер Э., Нерсетт С., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. — М.: Мир, 1990.
- [195] Хакимзянов Г.С., Шокин Ю.И., Барахнин В.Б., Шокина Н.Ю. Численное моделирование течений жидкости с поверхностными волнами. — Новосибирск: Изд-во СО РАН, 2001.
- [196] Хокни Р., Иствуд Дж. Численное моделирование методом частиц. — М.: Мир, 1987.
- [197] Чесбро Г. Открытые инновации. — М.: Изд. Поколение, 2007.
- [198] Четвертушкин Б.Н. Кинетические схемы и квазигазодинамическая система уравнений. — М.: МАКС Пресс, 2004. — 332 с.
- [199] Шайдуров В.В. Многосеточные методы конечных элементов. — М.: Наука, 1989.
- [200] Шенан П., Коснар М., Гордан И. и др. Математика и САПР. Кн. 1. — М.: Мир, 1988.
- [201] Шиликин Г.И. Сеточные аппроксимации сингулярно возмущенных эллиптических и параболических уравнений. — Екатеринбург: УРО РАН, 1992.
- [202] Шокин Ю.И., Лисейкин В.Д., Лебедев А.С., Данаев Н.Т., Китаева И.А. Методы римановой геометрии в задачах построения разностных сеток. — Новосибирск: Наука, 2005.
- [203] Шокин Ю.И., Шурина Э.П., Иткина Н.Б. Современные многосеточные методы. Часть 1. Многомасштабные методы. — Новосибирск: Изд. НГТУ, 2010.
- [204] Шокин Ю.И., Шурина Э.П., Иткина Н.Б. Современные многосеточные методы. Многоуровневые методы. Применение многомасштабных методов. — Новосибирск: Изд. НГТУ, 2012.
- [205] Шокин Ю.И., Яненко Н.Н. Метод дифференциального приближения. Применение к газовой динамике. — Новосибирск: Наука, 1985.

-
- [206] Яненко Н.Н. Метод дробных шагов решения многомерных задач математической физики. — Новосибирск: Наука, 1967.
- [207] Яненко Н.Н., Коновалов А.Н. Современные проблемы математической физики и вычислительной математики. — Новосибирск: Наука, 1982.
- [208] Яненко Н.Н., Коновалов А.Н. Некоторые вопросы теории модульного анализа и параллельного программирования для задач математической физики и механики сплошной среды. — Современные проблемы математической физики и вычислительной математики. — М.: Наука, 1982.
- [209] Яненко Н.Н., Преображенский Н.Г., Разумовский О.С. Методологические проблемы математической физики. — Новосибирск: Наука, 1986.
- [210] Яненко Н.Н., Рычков А.Д. Актуальные проблемы прикладной математики и математического моделирования. — Новосибирск: Наука, 1982.

Предметный указатель

- d -симплекс, 378
- A -скалярное произведение, 173
- Аппроксимация, 173
 - Коллатца, 199
 - Шортли–Уэллера, 202
 - биполиномиальная, 315
 - конформная, 334
 - лагранжевая, 293
 - монотонная, 188
 - первой производной, 177
- Базис, 51
 - иерархический, 302, 318
 - лагранжевый высших порядков, 316
 - эрмитовый кубический, 317
- Барицентрические координаты, 311, 378
- Вариация функционала, 149
- Вектор, 173
 - локальный
 - неизвестного, 286
 - правой части, 287
 - невязки, 148
 - обобщенных координат, 122
- Восполнение, 307
- Гамильтонова электродинамика, 127
- Гомогенизация, 358
- Граница
 - кусочно-гладкая, 63
 - липшицева, 63
- Граничное условие, 173
 - периодичности, 175
- Декомпозиция сеточной расчетной области, 384
- Диагональное преобладание, 173
- Задача, 173
 - Кеплера, 128
 - Стокса, 107
 - вариационная, 147
 - абстрактная, 141
 - вариационного исчисления, 149
 - внешняя краевая, 94
 - внутренняя краевая, 94
 - корректная, 80
 - краевая, 80
 - корректная, 80
 - магнитостатики, 118
 - минимизации абстрактная, 147
 - начально-краевая, 88
 - обратная, 368
 - с пограничными, внутренними слоями, 94
 - сингулярно-возмущенные, 94
 - сопряженная, 162
 - электростатики, 118
- Закон
 - Гаука, 109
- Замыкание, 70
- Замыкание множества, 48
- Изогеометрический анализ, 370
- Изопараметрические преобразования, 323
- Интерполянт решения, 157
- Интерполяция
 - бикубическая эрмитовая, 317
 - кубическая эрмитовая, 299
- Итерационный процесс, 173
- Калибровка Лоренца, 115
- Квадратный корень, 72

- Конденсация, 297
 Конечно-элементный анализ, 370
 Конечные элементы
 иммерсионные, 341
 прямоугольные, 316
 треугольные, 309
 Конечный элемент
 криволинейный, 327
 Корректная разрешимость, 74
 Коэффициент неравномерности сетки, 283
 Кратность
 интерполяционного узла, 354
 собственного значения, 77

 Лагранжиан, 123
 Линейная оболочка, 50

 Масштабирующий множитель, 190
 Матрица, 173
 Гамильтона, 130
 Грама, 56, 160
 баланса, 226
 барицентрическая, 311
 глобальная, 289
 глобальная источника, 275
 глобальная системы, 264
 жесткости, 290
 каноническая структурная, 129
 конвекции, 290
 локальная, 286, 290
 локальная баланса, 264
 локальная источника, 275
 локальная потоков, 274
 масс, 290
 симплектическая, 131
 треугольная, 311
 Мера Лебега, 47
 Метод
 Галеркина, 160
 иммерсионный локаль-
 ный разрывный k -го по-
 рядка, 341
 разрывный, 339
 Ритца, 156
 альтернирующий Шварца,
 391
 декомпозиции
 аддитивный, 391
 декомпозиции областей
 (МДО), 391
 интегро-балансный, 240
 интегро-интерполяционный,
 238
 конечных объемов, 226
 многосеточный, 383
 неопределенных коэффици-
 ентов, 220
 подсеток, 358
 суперэлементов, 358
 энергетический, 147
 Метрика, 47
 Многомасштабные методы, 358
 Многочлен
 Лежандра, 303
 интерполяционный
 Лагранжа, 293
 Эрмита, 356
 Множество
 бикompактное, 49
 всюду плотное, 48
 выпуклое, 51
 замкнутое, 48
 измеримое, 47
 компактное, 49
 линейных функционалов, 60
 ограниченное, 47
 плотное, 48
 Мульти-симплектический фор-
 мализм, 134

 Наилучшее приближение, 55
 Неравенство
 Гельдера, 59

- Коши–Буняковского, 54, 77
 Пуанкаре, 67
 Пуанкаре–Фридрихса, 67
 Фридрихса, 66
 Шварца, 54
 вариационное, 148
 Норма, 60, 173
 Липшица, 62
 евклидовая, 54
 кубическая, 52
 октаэдрическая, 52
 оператора, 69
 сферическая, 52
 эквивалентная, 52
 энергетическая, 82
 Носитель, 47
 Область
 ограниченная, 46
 определения оператора, 68
 расчетная, 371
 Оператор, 68, 173
 Лапласа, 86
 Лапласа пятиточечный раз-
 ностный, 181
 Пуанкаре–Стеклова, 360
 вполне непрерывный, 70
 единичный, 70
 замкнутый, 70
 коммутативный, 72
 линейный, 69
 невырожденный, 70
 непрерывный, 69
 обратный, 70
 ограниченный, 69
 перестановочный, 72
 положительно определен-
 ный, 76
 положительно полуопреде-
 ленный, 76
 проектирования, 72
 самосопряженный, 76
 сеточный, 284
 симметричный, 76
 сопряженный, 75
 Определитель, 56
 Осреднение коэффициентов, 358
 Отношение Рэлея, 79
 Отображение
 симплектическое, 130
 Отображение множества, 68
 Отрезок, 50
 Поверхности Безье, 371
 Погрешность
 аппроксимации, 176
 Подпространство, 50
 Подсетка
 несогласованная, 376
 согласованная, 376
 Полунорма, 51, 67
 Пополнение пространства, 49
 Порядок, 173
 конечного элемента, 285
 погрешности, 176
 Последовательность, 48
 биортогональная, 57
 Постановки задач электромагне-
 тизма, 113
 Потенциал
 двойного слоя, 120
 поверхностный, 119
 простого слоя, 120
 электродинамический, 115
 Преобразование
 Лежандра, 124
 Фурье
 прямое и обратное, 115
 Преобразование(я)
 канонические, 124
 Проекция, 55, 185
 Произведение скалярное энерге-
 тическое, 82
 Производная конормальная, 84
 Производные
 функциональные, или varia-
 ционные, 131

- Пространство
 Лебега, 58
 Равьяра-Тома, 335
 Соболева, 62
 банаховое, 71
 бездивергентных функций, 335
 вещественное, 50
 гильбертово, 54
 двойственное, 60
 евклидово, 53
 комплексное, 50
 метрическое, 47
 нормированное, 51
 ортогональное, 55
 ортонормированное, 55
 основных функций, 61
 полунормированное, 51
 предгильбертово, 54
 рефлексивное, 73
 сепарабельное, 48
 строго нормированное, 55
 топологическое, 49
 унитарное, 54
 фазовое, 130
 энергетическое, 82, 147, 154
- Равенство вариационное, 141
 Радиус спектральный, 79
 Разность
 односторонняя
 направленная, 188
 Разность (конечная)
 назад, 176
 вперед, 176
 второго порядка, 179
 двусторонняя первого порядка, 177
 левая, 176
 односторонняя, 176
 первого порядка, 176
 правая, 176
 трехточечная, 177
 центральная, 177
 четвертого порядка, 180
- Расстояние, 47
 Расширение линейного оператора, 69
 Решение
 разностное, 183
 Решение уравнения
 устойчивое, 83
 Ряд
 Неймана, 72
 Фурье, 79
 абсолютно сходящийся, 53
 сходящийся, 52
- Сетка, 173
 адаптивная, 375
 квадратная, 181
 квазиравномерная, 384
 квазиструктурированная, 376
 кубическая, 196
 неравномерная, 288
 неструктурированная, 376
 параллелепипедоидальная, 195
 прямоугольная, 181, 306
 псевдоравномерная, 283
 равномерная, 181
 регулярная, 377
 симплектическая
 конформная, 378
 неконформная, 378
 структурированная, 376
 точно-ориентированная, 250
 треугольная, 257
 элементно-ориентированная, 250
- Сжатый разреженный формат, 386
 Симметризация, 190
 Симплектическая дискретизация, 378

- Система
 линейно независимая, 50
 ортогонализации Грама–Шмидта, 78
 ортонормальная, 78
 Система уравнений, 173
 симметричная, 189
 абсолютно монотонная, 187
 неразложимая, 187
 положительного типа, 187
 симметризуемая, 190
 условно монотонная, 187
 якобиевая, 190
 Скобки
 Пуассона, 125
 фундаментальные, 126
 Скорость сходимости итераций, 173
 Сумма ряда, 53
 Схема, 173
 балансная, 226
 компактная, 206
 компактная МКЭ, 319
 консервативная разностная, 226
 Сходимость
 абсолютная, 71
 по норме, 71
 равномерная, 71
 сильная, 54, 71
 слабая, 54, 73
 Теорема, 173
 Больцано–Вейерштрасса, 49
 Вишика–Лакса–Мильграма, 144
 Обэна–Ницше, 163
 Пифагора, 55
 Рисса, 60, 146
 Сеа, 161
 Хана–Банаха, 73
 вложения, 64
 о минимальном расстоянии, 55
 о погрешности метода Рунге, 157
 о продолжении функции, 65
 о сжимающем отображении, 144
 о следах, 66
 Тождество, 173
 Якоби, 126
 двумерное интегральное, 236
 интегральное Г. И. Марчука, 228
 Точка
 внутренняя, 48
 предельная, 48
 Триангуляция, 173, 305, 378
 Делоне, 258, 305
 области, 257
 Уравнение(я), 173
 19-точечное, 196
 27-точечное, 196
 девятиточечное Пуассона типа ”ящик”(Микеладзе), 193
 Гельмгольца, 116
 Дарси для течений в пористой среде, 101
 Захарова–Шабата, 134
 Кортевега де Вриза, 132
 Лагранжа, 123
 Ламе, 91
 Лапласа, 86
 Максвелла, 113
 Навье–Стокса, 107
 Пуассона, 86
 Шредингера, 133
 Эйлера, 149, 150
 бигармоническое, 90
 волновое, 88
 гиперболическое, 89
 девятиточечного типа ”большой крест”, 191
 дивергентного вида, 83

- диффузионно-конвективное, 186
диффузионное, 194
неразрывности, 113
параболическое, 89
полигармоническое, 90
положительного типа, 186
равномерно эллиптическое, 89
разностное, 183
семиточечное, 195
скалярное диффузионно-конвективное, 97
сопряженное, 75
теплопроводности, 88
трехточечное, 186
ультрагиперболическое, 89
эллиптическое, 84, 89
эллиптическое в области, 174
Условие
 Стефана, 101
Условие(я)
 Зоммерфельда, 96
Условия
 Зоммерфельда, 117
 главные, 139
 естественные граничные, 139
 смешанные краевые, 84
 сопряжения, 152
Фазовые потоки, 130
Форма
 билинейная, 60
 коэрцитивная, 61
 ограниченная (непрерывная), 60
 симметричная, 60
Формула Грина, 84, 137
Функционал
 линейный, 59, 60
 непрерывный, 59, 60
 ошибок, 155
 целевой, 368
 энергии, 146
Функция
 Куранта, 307
Липшица, 62
абсолютно непрерывная, 62
базисная
 Равьяра-Тома, 335
 внутренняя, 314
 кусочно-линейная, 288, 315
 лагранжевая, 295
 реберная, 313
 узловая, 313
 финитная, 288
билинейная, 315
гармоническая, 94
допустимая, 138
измеримая, 47
интегрируемая по Лебегу, 47
локально интегрируемая, 61
обобщенная, 61
 производная, 61
от оператора, 72
производящая, 125
финитная, 46
характеристическая, 46
Число
 обусловленности оператора, 74
 простое, 77
 степеней свободы, 285
Шар, 51
 замкнутый, 51
Ширина, 173
Элемент
 конечный лагранжевый, 286
 собственный, 77
 стандартный, 302
Энергия
 кинетическая, 122
 потенциальная, 123
Ядро
 оператора, 70
Ячейка Дирихле–Вороного, 258

Оглавление

Введение в предметную область	8
Глава 1. Операторные постановки краевых задач	44
1.1. Основные функциональные пространства: определения и формулировки теорем	46
1.1.1. Метрические и нормированные пространства	47
1.1.2. Пространства функций $L_p(\Omega)$ и $C^p(\Omega)$	58
1.1.3. Линейные функционалы и билинейные формы	59
1.1.4. Обобщенные производные и пространства Соболева	61
1.1.5. Липшицевы границы и теоремы вложения	63
1.1.6. Теоремы о продолжении функции и о следах	65
1.1.7. Пространства $H_0^m(\Omega)$ и $C_0^p(\Omega)$	66
1.2. Операторные формулировки краевых задач	68
1.2.1. Линейные операторы и уравнения	68
1.2.2. Корректность краевой задачи и энергетические пространства	80
1.2.3. Существование, устойчивость и единственность ре- шений краевых задач	83
1.3. Примеры и классификация задач моделирования	85
1.3.1. Уравнения, содержащие оператор Лапласа	86
1.3.2. Классификация дифференциальных уравнений и операторов	88
1.3.3. Типы краевых условий	92
1.3.4. Задачи тепло-массопереноса и фильтрации	97
1.3.5. Задачи гидро-газодинамики	103
1.3.6. Динамические задачи упругости	108
1.3.7. Постановки задач электромагнетизма	113
1.3.8. Интегральные уравнения теории потенциала	118
1.3.9. Гамильтоновы системы уравнений	121
Глава 2. Обобщенные решения и вариационные формулировки краевых задач	135
2.1. Обобщенные решения дифференциальных уравнений	136
2.2. Вариационная постановка краевых задач	140
2.3. Вариационная задача минимизации в симметричном случае	146
2.4. Приближенные методы решения вариационных задач	156
2.4.1. Метод Рунге	156
2.4.2. Метод Галеркина	160
2.4.3. «Несогласованный» метод Галеркина	164

2.4.4. Смешанные вариационные методы	167
Глава 3. Конечно-разностные аппроксимации	173
3.1. Приближения производных и дифференциальных вы- ражений	176
3.1.1. Разностные аппроксимации производной	176
3.1.2. Конечно-разностные выражения второго порядка	179
3.1.3. Двумерный разностный оператор Лапласа	181
3.1.4. Определение погрешности аппроксимации разност- ного выражения	182
3.2. Аппроксимации дифференциальных уравнений	183
3.2.1. Разностные уравнения Пуассона	183
3.2.2. Погрешность аппроксимации разностного уравнения	184
3.2.3. Аппроксимации диффузионно-конвективного урав- нения. Конечно-разностные уравнения положитель- ного типа и монотонные (абсолютно монотонные и условно монотонные) сеточные уравнения	186
3.2.4. Свойства симметричности и симметризуемости раз- ностных уравнений и выражений	189
3.2.5. Разностные аппроксимации двумерного уравнения Пуассона	190
3.2.6. Примеры трехмерных разностных уравнений	195
3.3. Методы аппроксимации граничных условий	197
3.3.1. Аппроксимация условий 1-го рода в одномерных задачах	198
3.3.2. Условия 2-го и 3-го рода (одномерный случай)	199
3.3.3. Аппроксимации граничных условий в двумерных задачах	201
3.3.4. Погрешности аппроксимации разностных краевых задач	204
3.4. Компактные схемы повышенной точности для уравне- ний с переменными коэффициентами	205
3.4.1. Аппроксимации диффузионно-конвективных урав- нений	206
3.4.2. Уравнение Пуассона в цилиндрической системе ко- ординат	212
3.4.3. Аппроксимации систем дифференциальных урав- нений	216
3.4.4. Аппроксимация граничных условий 2-го и 3-го родов	217
3.5. Метод неопределенных коэффициентов	220
Глава 4. Методы конечных объемов	226
4.1. Интегральное тождество Г. И. Марчука и его применение	227
4.1.1. Одномерный случай	227

4.1.2. Двумерный случай	235
4.2. Интегро-балансные аппроксимации	238
4.2.1. Одномерные уравнения	238
4.2.2. Двумерные уравнения	244
4.3. Аппроксимации на элементно-ориентированных сетках	250
4.3.1. Аппроксимация одномерного уравнения диффузии	250
4.3.2. Аппроксимация уравнения Пуассона на прямоуголь- ной элементно-ориентированной сетке	254
4.4. Треугольные конечные объемы	257
4.5. Поэлементная аппроксимация на прямоугольной сетке	264
4.5.1. Постановка задачи и обозначения	265
4.5.2. Методы аппроксимации интегралов	266
4.5.3. Основные свойства алгебраической системы	270
4.5.4. Локальные матрицы и технология сборки алгебра- ической системы	273
4.5.5. Аппроксимация уравнения диффузии с перемен- ными коэффициентами	276
Глава 5. Конечно-элементные методы и технологии	281
5.1. Принцип поэлементного формирования алгебраической системы	283
5.2. Построение одномерных уравнений МКЭ	288
5.2.1. Кусочно-линейная аппроксимация	288
5.2.2. Лангражевые аппроксимации высших порядков	293
5.2.3. Эрмитовы базисные функции	299
5.2.4. Иерархические базисы	302
5.3. Двумерные конечные элементы	305
5.3.1. Конечно-элементные аппроксимации на треугольниках	306
5.3.2. Прямоугольные конечные элементы	314
5.3.3. Изопараметрические преобразования и криволи- нейные элементы	323
5.3.4. Элементы Равьяра–Тома	331
5.4. Разрывные методы Галеркина	339
5.4.1. Одномерные иммерсионные алгоритмы	339
5.4.2. Многомерное квазилинейное эллиптическое урав- нение	344
5.4.3. Нестационарная диффузионно-конвективная задача	348
5.5. Общие замечания о построении матриц МКЭ	353
5.6. Многомасштабные методы и технологии	357

Глава 6. Алгоритмы и технологии построения сеток	363
6.1. Формальные аспекты математических постановок	366
6.2. Геометрические и функциональные объекты	369
6.3. Сеточные объекты и операции	374
6.3.1. Основные понятия и спецификации сеток	374
6.3.2. Некоторые подходы к генерации сеток	384
6.3.3. Вопросы формирования поузловой информации	385
6.3.4. Поэлементные технологии	387
6.4. Примеры построения адаптивных квазиструктурированных сеток	388
6.5. Задача декомпозиции сеточных областей	391
6.6. Принципы построения сеточных генераторов	401
Литература	405
Предметный указатель	420

Научное издание

Валерий Павлович Ильин

Математическое моделирование. Часть I
Непрерывные и дискретные модели

Ответственный редактор

Оформлено в системе L^AT_EX, макрос $\mathcal{NCS}$

Редактор ??

Технический редактор ??

Компьютерная графика *П.В. Ильин*

Лицензия ИД № 02202 от 30.06.2000 г. Подписано в печать
17.07.2006 г.

Формат $60 \times 84 \frac{1}{8}$. Гарнитура Times New Roman.
Усл. печ. л. ?. Уч.-изд. л. ?. Тираж 500 экз. Заказ № ?.

Издательство СО РАН
630090, Новосибирск, 90, Морской просп., 2