

Власова, Виктория Михайловна

Приближенные методы математической физики: Учебник для студентов высших технических учебных заведений/ Георгий Николаевич Кувыркин, Владимир Степанович Зарубин; Зарубин, Владимир Степанович, Крищенко, А.П.. - М.: МГТУ им. Н.Э. Баумана, 2001.- 700 с. : ил.

ISBN 5-7038-1768-4 (Вып. XIII) ,5-7038-1270-4, 3000 экз.

Книга является тринадцатым выпуском серии учебников,, Математика в техническом университете". Последовательно изложены математические модели физических процессов, элементы прикладного функционального анализа и приближенные аналитические методы решения задач математической физики, а также широко применяемые в научных исследованиях и инженерной практике численные методы конечных разностей, конечных и граничных элементов. Рассмотрены примеры использования этих методов в прикладных задачах. Содержание учебника соответствует курсам лекций, которые авторы читают в МГТУ им. Н.Э. Баумана. Для студентов технических университетов. Может быть полезен преподавателям, аспирантам и инженерам.

Численные методы решения уравнений

Высшая школа

ББК 22.193

УДК 517.1(075.8)

Оглавление

Предисловие - 5-11с.

Основные обозначения - 12-16с.

ЧАСТЬ I МАТЕМАТИЧЕСКИЕ МОДЕЛИ ФИЗИЧЕСКИХ ПРОЦЕССОВ - 17-134с.

1 Основные физические субстанции - 19-37с.

1.1 Особенности постановки задач математической физики - 19-21с.

1.2 Плотность физических субстанций - 21-25с.

1.3 Перенос физических субстанций - 26-30с.

Д. 1.1 Некоторые формулы векторного анализа - 30-37с.

Вопросы и задачи - 37с.

2 Законы сохранения физических субстанций - 38-84с.

2.1 Закон сохранения массы - 38-45с.

2.2 Дивергентная форма уравнения неразрывности - 46-57с.

2.3 Законы сохранения электрического заряда и тепловой энергии - 57-72с.

2.4 Закон сохранения количества движения - 72-81с.

Д. 2.1 Формулы векторного анализа в случае неоднородной среды - 81-84с.

Вопросы и задачи - 84с.

3 Математические модели некоторых сред - 85-134с.

3.1 Модели идеальной жидкости (газа) - 85-91с.

3.2 Модели вязкой жидкости - 91-98с.

3.3 Упругое твердое тело - 98-102с.

3.4 Уравнение переноса энергии в среде - 103-108с.

3.5 Уравнения Максвелла - 108-117с.

3.6 Электромагнитные процессы в медленно движущейся среде - 117-122с.

Д. 3.1 Поверхности разрыва в электромагнитном поле - 122-125с.

Д. 3.2 Примеры задач, описываемых интегральными уравнениями - 125-133с.

Вопросы и задачи - 134с.

ЧАСТЬ II ЭЛЕМЕНТЫ ФУНКЦИОНАЛЬНОГО АНАЛИЗА И ПРИБЛИЖЕННЫЕ АНАЛИТИЧЕСКИЕ МЕТОДЫ - 135-398с.

4 Нормированные пространства и операторы - 137-208с.

4.1 Нормированные пространства - 137-145с.

4.2 Операторы в нормированных пространствах - 146-161с.

4.3 Линейные операторы - 161-164с.

4.4 Линейные ограниченные функционалы - 164-171с.

4.5 Нормированное пространство линейных операторов - 171-185с.

4.6 Спектр линейного оператора - 186-196с.

4.7 Пополнение нормированного пространства - 196-204с.

Вопросы и задачи - 204-208с.

5 Операторы в гильбертовых пространствах - 209-282с.

5.1 Гильбертово пространство - 209-215с.

5.2 Операторы и функционалы в гильбертовом пространстве - 215-226с.

5.3 Энергетическое пространство - 226-233с.

5.4 Однородное операторное уравнение - 234-246с.

5.5 Уравнения с вполне непрерывными симметрическими операторами - 246-258с.

Д.5.1 Сопряженные пространства и сопряженные операторы - 258-268с.

Д.5.2 Критерий базисности системы функций - 268-269с.

Д.5.3 Положительная определенность эллиптического оператора - 270-281с.

Вопросы и задачи - 281-282с.

6 Приближенные аналитические методы - 283-398с.

6.1 Общая схема построения приближенных методов - 283-293с.

6.2 Погрешности приближенных методов - 293-297с.

6.3 Метод малого параметра - 297-303с.

6.4 Общий случай метода малого параметра - 303-323с.

- 6.5 Метод ортогональных проекций - 323-326с.
- 6.6 Коллокации в подобластях и в точках - 326-343с.
- 6.7 Метод наименьших квадратов - 343-351с.
- 6.8 Методы Бубнова—Галеркина и Ритца - 351-357с.
- 6.9 Задачи на собственные значения - 357-371с.
- 6.10 Особенности выбора базисных функций - 371-389с.
- Д.6.1 Проекционный метод - 389-395с.
- Вопросы и задачи - 395-398с.

ЧАСТЬ III СЕТОЧНЫЕ МЕТОДЫ - 399-522с.

7 Основы метода конечных разностей - 401-424с.

- 7.1 Понятие о сеточных методах - 401-403с.
- 7.2 Аппроксимация производных конечными разностями - 404-407с.
- 7.3 Метод баланса - 407-414с.
- 7.4 Пример простейшей разностной схемы - 414-424с.
- Вопросы и задачи - 424с.

8 Одномерные краевые задачи - 425-490с.

- 8.1 Разностные схемы для стационарных задач - 425-440с.
- 8.2 Задача Штурма — Лиувилля - 441-452с.
- 8.3 Нестационарная задача теплопроводности - 452-467с.
- 8.4 Некоторые динамические задачи - 467-472с.
- Д.8.1 Модификации метода прогонки - 472-489с.
- Вопросы и задачи - 490с.

9 Многомерные задачи - 491-522с.

- 9.1 Особенности решения многомерных задач - 491-492с.
- 9.2 Двумерная и трехмерная задачи теплопроводности - 492-500с.
- 9.3 Различные многомерные задачи - 500-514с.

Д.9.1 Алгоритмы матричной и ортогональной прогонок - 514-521с.

Вопросы и задачи - 521-522с.

ЧАСТЬ IV МЕТОДЫ КОНЕЧНЫХ И ГРАНИЧНЫХ ЭЛЕМЕНТОВ - 523-683с.

10 Основы метода конечных элементов - 525-568с.

10.1 Одномерная краевая задача - 525-537с.

10.2 Типы конечных элементов - 538-562с.

10.3 Матричная форма представления функций - 563-567с.

Вопросы и задачи - 567-568с.

11 Прикладные задачи - 569-621с.

11.1 Особенности применения метода конечных элементов - 569-572с.

11.2 Задачи теплопроводности в твердом теле - 572-591с.

11.3 Двумерное течение вязкой жидкости - 592-602с.

11.4 Задачи теории упругости - 602-616с.

11.5 Электромагнитное поле в цилиндрическом волноводе - 616-621с.

Вопросы и задачи - 621с.

12 Введение в метод граничных элементов - 622-683с.

12.1 Граничные интегральные уравнения - 622-635с.

12.2 Способы аппроксимации функций на границе - 635-647с.

12.3 Учет анизотропии и неоднородности - 647-655с.

12.4 Нестационарные задачи - 656-663с.

12.5 Статическая задача теории упругости - 663-674с.

12.6 Сравнение методов граничных и конечных элементов - 674-676с.

Д. 12.1 Особенности решения осесимметричных задач - 676-683с.

Вопросы и задачи - 683с.

Список рекомендуемой литературы - 584-688с.

Предметный указатель - 689-695с.

ПРЕДИСЛОВИЕ

Точное аналитическое решение задач математической физики обычно требует интегрирования дифференциальных уравнений с частными производными, включающих искомые функции. Эти уравнения в общем случае необходимо проинтегрировать в некоторой пространственно-временной области, на границе которой искомые функции подчинены заданным краевым условиям.

Реализация такого подхода связана обычно с большими и не всегда преодолимыми трудностями. Но с прикладной точки зрения наряду с аналитической формой точного решения задачи не меньшее значение имеет получение приближенного аналитического решения или приближенных числовых значений искомых величин. Математические модели ряда физических процессов содержат интегральные или интегро-дифференциальные уравнения, в которых искомые функции входят и под знак интеграла. Точное аналитическое решение таких уравнений возможно лишь в редких случаях, что также подчеркивает значимость приближенных методов решения.

Дифференциальные уравнения, описывающие законы сохранения и переноса физических субстанций и используемые при постановке задач математической физики (они приведены в первой части книги), можно рассматривать как операторные, действующие в тех или иных функциональных пространствах. В связи с этим во второй части кратко изложены необходимые сведения из функционального анализа и рассмотрены свойства некоторых операторов, характерных для таких задач. Эти сведения использованы затем при изложении приближенных аналитических и численных (в третьей и четвертой частях) методов решения задач математической физики.

Значительное число примеров, иллюстрирующих рассматриваемые методы, связаны с процессами теплопроводности в твердом теле. В силу прозрачности физической постановки таких задач и предсказуемости качественного характера решения они являются хорошим „полигоном“ для проверки эффективности методов и сопоставления различных подходов к получению количественного решения краевых задач, описываемых дифференциальными уравнениями эллиптического и параболического типов. Методы, применяемые для решения краевых задач, описываемых уравнениями гиперболического типа, проиллюстрированы на примерах, связанных с волновым уравнением.

Ссылкой в тексте на конкретный выпуск комплекса учебников „Математика в техническом университете“ является его номер, записанный римскими цифрами. Например, [I-7.5] означает ссылку на пятый параграф седьмой главы в первом выпуске, тогда как (см. 1.2) — отсылает читателя ко второму параграфу первой главы этой книги, а (см. Д.4.1) — к первому дополнению четвертой главы этой книги. Ссылки в тексте на номера формул и рисунков набраны обычным шрифтом (например, (2.1) — первая формула в главе 2, рис. 1.5 — пятый рисунок в главе 1).

За предисловием следует перечень основных обозначений, где наряду с их краткой расшифровкой даны ссылки на разделы этого или других выпусков серии, в которых можно найти более подробное объяснение каждого обозначения. После этого перечня приведены написание и русское произношение букв латинского и греческого алфавитов.

В конце книги помещены список рекомендуемой литературы и предметный указатель, в который входят в алфавитном порядке (по существительному в именительном падеже) все выделенные в тексте **полужирным курсивом** термины с указанием страницы, где они строго определены или описаны. Выделение термина *светлым курсивом* означает, что в данном параграфе он отнесен к ключевым словам и читателю для понимания излагаемого материала должно быть известно значение

этого термина. Читатель может уточнить это значение, найдя при помощи предметного указателя необходимую страницу, на которой используемый термин определен или описан. Если термин введен в другом выпуске, то указан его номер римской цифрой. Курсивом в предметном указателе даны ссылки, указывающие на дополнительную информацию о термине.

Изучение приближенных методов математической физики опирается на знание практически всех разделов общего курса высшей математики. Поэтому перед чтением этой книги необходимо в целях самоконтроля выполнить следующие несложные задания. В конце каждого задания приведена ссылка на тот выпуск, в котором при возникновении затруднений можно найти все необходимые сведения. Значения терминов, выделенных в тексте этих заданий **прямым полужирным** шрифтом, далее будем считать известными (в основном тексте книги эти термины не выделены и не входят в предметный указатель).

Задания для самопроверки

1. Как определить множества **целых $\mathbb{Z}$** и **рациональных $\mathbb{Q}$** чисел при помощи множества **$\mathbb{N}$ натуральных чисел**? Что такое **абсолютное значение** действительного числа и **модуль комплексного числа**? Каковы свойства **точных верхней и нижней граней** подмножества множества **действительных чисел $\mathbb{R}$** ? [I]

2. Из каких этапов состоят **доказательства от противного** и по **методу математической индукции**? [I]

3. Что такое **объединение**, **пересечение**, **разность** и **прямое (декартово) произведение** множеств (подмножеств), **дополнение** множества? [I]

4. Какие **точки** множества в **$\mathbb{R}^n$** называют **внутренними**, **граничными**, **предельными**, **изолированными**? Что такое **открытое**, **замкнутое**, **ограниченное**, **компактное** множества в **n -мерном евклидовом арифметическом пространстве $\mathbb{R}^n$** ? [V]

5. Запишите с помощью неравенств условия принадлежности точки x промежуткам числовой прямой: отрезку $[a, b]$, интервалу (a, b) , полуинтервалу $(a, b]$, бесконечному интервалу $(-\infty, b)$ и бесконечному полуинтервалу $[a, +\infty)$. [I]

6. Изобразите на числовой прямой окрестности конечной и бесконечной точек расширенной числовой прямой. В чем отличие этих окрестностей от проколотых окрестностей и полуокрестностей? [I]

7. Какими свойствами обладает взаимно однозначное отображение двух множеств? Чему равна композиция прямого и обратного отображений двух множеств? [I]

8. Приведите примеры составной и периодической действительных функций действительного переменного и укажите их области определения (существования) и значений. Как расположены относительно начала координат графики четной и нечетной функции? [I]

9. Сколько нулей имеет многочлен степени n ? В чем различие между простым и кратным нулем? [I]

10. Сформулируйте определения предела, производной и дифференциала действительной функции действительного переменного в точке. Всякая ли функция, непрерывная в точке, является дифференцируемой в этой точке? Приведите примеры функций, имеющих точки: а) устранимого разрыва; б) разрыва первого рода; в) разрыва второго рода. Каковы свойства функции, непрерывной (дифференцируемой) на отрезке? Как вычислить производную сложной функции? Что называют вектор-функцией? [I], [II]

11. Является ли сходящаяся числовая последовательность ограниченной? В чем различие между монотонной и строго монотонной последовательностью? Сформулируйте признак Вейерштрасса сходимости последовательности. [I]

12. В каких точках числовой оси функции $\sin x$, $1/x$ являются бесконечно малыми, а функции x^2 , $\operatorname{ctg} x$ — бесконечно большими? [I]

13. В чем различие между **монотонной** и **строго монотонной** в некотором промежутке функциями? Каковы условия существования в нем непрерывной и строго монотонной функции, **обратной** заданной функции? Изобразите графики **возрастающей, убывающей, невозрастающей и неубывающей** в промежутке функций. [I]

14. Приведите примеры бесконечно малых при $x \rightarrow a$ функций: а) **одного порядка**; б) **первого порядка малости** одной относительно другой; в) **несравнимых**; г) **эквивалентных**. Каковы свойства эквивалентных бесконечно малых функций? В каком случае **главную часть** функции, бесконечно малой при $x \rightarrow a$, можно представить **степенной функцией**? Каков смысл символов „**о малое**“ и „**О большое**“? [I]

15. Что такое **абсолютная и относительная погрешности**? [II]

16. Каковы правила вычисления **скалярного, векторного и смешанного произведений векторов**? Что такое **нулевой, единичный и коллинеарные векторы, радиус-вектор точки в пространстве**? [III], [IV]

17. Что такое **квадратная, прямоугольная, блочная, симметрическая, нулевая, единичная, верхняя (нижняя) треугольная, вырожденная матрица**? Что представляют собой **транспонированная и обратная матрицы** по отношению к данной матрице? Как вычислить **ранг матрицы**? Чему равен **определитель диагональной матрицы**? При каком условии **однородная квадратная система линейных алгебраических уравнений (СЛАУ)** имеет **ненулевое решение**? [III]

18. Перечислите **аксиомы линейного пространства**, которым подчиняются **линейные операции** в этом пространстве. Что называют **базисом и размерностью линейного**

пространства? Каковы аксиомы расстояния между элементами метрического пространства и нормы элементов нормированного пространства? В чем состоит отличие этих пространств? Дайте геометрическую интерпретацию неравенства треугольника в трехмерном евклидовом арифметическом пространстве $\mathbb{R}^3$. Что такое ортонормированный базис в $\mathbb{R}^n$? [I], [IV]

19. Как найти собственные значения и собственные векторы линейного оператора в конечномерном линейном пространстве? Что такое собственное подпространство линейного оператора, характеристические многочлен и уравнение матрицы? Какую квадратичную форму называют положительно (отрицательно) определенной? Сформулируйте критерий Сильвестра. [IV]

20. Запишите выражение для тривиальной линейной комбинации n векторов (функций) и сформулируйте определение линейно зависимой (независимой) системы векторов (функций). [III], [IV], [VIII]

21. Что называют векторной функцией многих переменных и ее координатной функцией, частной производной и производной по направлению? Запишите формулу Тейлора для скалярной функции многих переменных. При каком условии смешанная производная не зависит от последовательности дифференцирования? Что понимают под внешней нормалью к кусочно гладкой кривой, ограничивающей область на плоскости? Что понимают под внешней нормалью к кусочно гладкой поверхности, ограничивающей область в пространстве? [V]

22. Что такое неопределенный интеграл? Напишите формулу интегрирования по частям и формулу Ньютона — Лейбница. Чему равна производная определенного интеграла по переменному верхнему пределу? Что такое абсолютно сходящийся несобственный интеграл? В чем различие интегралов Римана и Лебега? [VI], [IX]

23. Что называют кратным (в том числе двойным и тройным), криволинейным и поверхностным интегралами, скалярным и векторным, безвихревым, потенциальным, соленоидальным и лапласовым полями, потоком и циркуляцией вектора, телесным углом? Напишите выражения для операторов Гамильтона и Лапласа в декартовой прямоугольной системе координат. Как при помощи оператора Гамильтона записать операции дивергенции, градиента и ротора? Что понимают под элементарным объемом? [VII]

24. Что такое сходящийся числовой ряд? При каких условиях функциональный ряд сходится на данном множестве поточечно, равномерно, абсолютно, условно? Что такое радиус сходимости степенного ряда? Запишите частичную сумму и остаток ряда Тейлора. В чем различие между ортогональной и ортонормированной системами функций? Как вычислить коэффициенты Фурье разложения функции в ряд Фурье? Каков смысл выражения почти всюду? Что понимают под замыканием множества? [IX]

25. Что называют линейным дифференциальным оператором и линейным обыкновенным дифференциальным уравнением (линейным ОДУ), нормальной системой ОДУ, уравнениями с частными производными эллиптического, параболического и гиперболического типов? Напишите уравнения Лапласа, Пуассона, теплопроводности и волновое уравнение. В чем различие между задачей Коши и краевой задачей? Что такое начальные и граничные условия? [VIII], [XII]

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ

- ◀ и ▶ — начало и окончание доказательства
- # — окончание примера, замечания, теоремы, следствия
- $a \in A$ — элемент a принадлежит множеству A I-1.1
- $A = \{a, b, c\}$ — множество A состоит из элементов a, b, c I-1.1
- $\emptyset$ — пустое множество I-1.1
- $A \subset B, B \supset A$ — множество A включено в множество B I-1.2
- $\mathbb{N}$ — множество натуральных чисел I-1.3
- $\mathbb{Z}$ — множество целых чисел I-1.3
- $\mathbb{Q}$ — множество рациональных чисел I-1.3
- $\mathbb{R}$ — множество действительных чисел I-1.3
- $\mathbb{C}$ — множество комплексных чисел I-4.3, X
- $X \times Y$ — декартово произведение множеств X и Y I-2.5
- $\mathbb{R}^n$ — произведение (декартово) n множеств действительных чисел (n -я декартова степень множества $\mathbb{R}$); n -мерное линейное арифметическое пространство I-2.5, IV
- $[a, b], (a, b)$ — отрезок, интервал с концами в точках a и b I-1.3
- $[a, b), (a, b]$ — полуинтервалы с концами в точках a и b I-1.3
- $|x|$ — абсолютная величина (модуль) числа x I-1.3
- ∞ — объединение бесконечных точек $+\infty$ и $-\infty$ I-1.3
- $A \Rightarrow B$ — из высказывания A следует высказывание B (A — достаточное условие B , а B — необходимое условие A) I-1.5
- $A \Leftrightarrow B$ — высказывания A и B равносильны I-1.5

$\exists$ — квантор существования ($\exists x$ — существует такое x , что ...) I-1.5

$\forall$ — квантор всеобщности ($\forall x$ — для любого x) I-1.5

$\sum_{k=1}^n a_k$ — сумма n слагаемых $a_1, a_2, \dots, a_n$ I-2.6

$\prod_{m=1}^n a_m$ — произведение n сомножителей $a_1, a_2, \dots, a_n$ I-2.6

$k = \overline{1, n}$ — число k принимает последовательно все значения из множества натуральных чисел от 1 до n включительно I-2.6

$\sup X, \sup_{x \in X} x$ — точная верхняя грань множества X I-2.7

$\inf X, \inf_{x \in X} x$ — точная нижняя грань множества X I-2.7

ρ и φ — полярные координаты (радиус и угол) точки на плоскости I-4.3

$\lim_{x \rightarrow a} f(x)$ — предел функции $f(x)$ в точке a (при $x \rightarrow a$) I-7.1

δx и $\delta y = \delta f(x)$ — приращения аргумента x и функции $y = f(x)$ I-9.1

$f'(a)$ — производная функции $f(x)$ в точке a II

$y'(x), y'_x, dy/dx, y'$ — производная функции $y = f(x)$ II

dx и $dy = df(x)|_{x=a}$ — дифференциалы аргумента x и функции $y = f(x)$ в точке a II

$f^{(n)}(a)$ — производная n -го порядка (n -я производная) функции $f(x)$ в точке a II

$\int$ — знак интеграла VI

$\overline{U}$ — замыкание области U IX

x_1, x_2, x_3 — координаты точки в декартовой прямоугольной системе координат $Ox_1x_2x_3$ 1.2

e_1, e_2, e_3 — орты декартовой прямоугольной системы координат $Ox_1x_2x_3$ 1.2

- ∇ — дифференциальный оператор Гамильтона VII, 1.3
- $\nabla^2 = \Delta$ — дифференциальный оператор Лапласа VII, 2.1
- 0 — нулевой элемент линейного пространства IV, 4.1
- $[\cdot]$ — скачок значений функции при переходе через поверхность разрыва 2.2
- $C[a, b]$ — нормированное пространство функций, непрерывных на отрезке $[a, b]$ IX
- $C^k[a, b]$ — нормированное пространство функций, непрерывно дифференцируемых k раз на отрезке $[a, b]$ 4.1
- $L_2[a, b]$ — гильбертово пространство функций, суммируемых с квадратом на отрезке $[a, b]$ IX
- $L_2(\Omega)$ — гильбертово пространство функций, суммируемых с квадратом в области Ω 5.1
- $\{u_n\}_N$ — система N функций $u_n, n = \overline{1, N}$ 4.1
- $\{u_n\}$ — последовательность функций $u_n, n \in \mathbb{N}$ 4.1
- $\|\cdot\|$ — норма элемента или оператора в нормированном пространстве 4.1
- $\rho(u, v)$ — расстояние между элементами u и v в метрическом пространстве $\mathcal{U}$ 4.1
- $D(f), R(f)$ — область определения и область значений функции (оператора, функционала) f 4.2, 4.4
- $A \circ B, AB$ — композиция, произведение операторов A и B 4.2
- $I_{\mathcal{U}}$ — тождественный оператор $I: \mathcal{U} \rightarrow \mathcal{U}$ 4.2
- A^{-1} — оператор, обратный к оператору A 4.2
- $\mathcal{L}(\mathcal{U}, \mathcal{W})$ — линейное пространство линейных операторов, действующих из линейного пространства $\mathcal{U}$ в линейное пространство $\mathcal{W}$ IV, 4.5
- $\ker A$ — ядро линейного оператора A IV, 4.5
- $\langle u, v \rangle$ — скалярное произведение элементов u и v евклидова (гильбертова) пространства 5.1

- $u \perp v$ — элементы u и v евклидова (гильбертова) пространства ортогональны 5.1
- $w \perp M$ — элемент w в евклидовом (гильбертовом) пространстве ортогонален подпространству M 5.1

Единицы измерения физических величин

- кг, м, с --- килограмм, метр и секунда — основные единицы измерения массы, расстояния и времени 1.2
- Дж, Кл, Н — джоуль, кулон и ньютон --- единицы измерения энергии, электрического заряда и силы 1.2
- Вт, А, В, Ом — ватт, ампер, вольт и ом - - единицы измерения мощности, силы электрического тока, электрических напряжения и сопротивления 1.3
- К --- кельвин — основная единица измерения температуры 1.3
- Па — паскаль — единица измерения давления и механического напряжения 1.3

Используемые сокращения

- ГИУ — граничное интегральное уравнение 12.1
- ГЭ — граничный элемент 12.2
- КЭ — конечный элемент 10
- МГЭ -- метод граничных элементов 12
- МКР — метод конечных разностей 7.1
- МКЭ — метод конечных элементов 10
- ОДУ — обыкновенное дифференциальное уравнение 4.2
- СЛАУ — система линейных алгебраических уравнений 6.1

Буквы латинского алфавита

Начертание	Произношение	Начертание	Произношение
A a A a	а	N n N n	эн
B b B b	бэ	O o O o	о
C c C c	цэ	P p P p	пэ
D d D d	дэ	Q q Q q	ку
E e E e	е	R r R r	эр
F f F f	эф	S s S s	эс
G g G g	же	T t T t	тэ
H h H h	аш	U u U u	у
I i I i	и	V v V v	вэ
J j J j	йот	W w W w	дубль-вэ
K k K k	ка	X x X x	икс
L l L l	эль	Y y Y y	игрек
M m M m	эм	Z z Z z	зэт

Представлен наиболее употребительный (но не единственный) вариант произношения (в частности, вместо „йот“ иногда говорят „жи“).

Буквы греческого алфавита

Начертание	Произношение	Начертание	Произношение	Начертание	Произношение
A α	альфа	I ι	йота	P ρ	ро
B β	бета	K κ	каппа	Σ σ	сигма
Γ γ	гамма	Λ λ	ламбда	Τ τ	тау
Δ δ	дельта	Μ μ	ми	Υ υ	ипсилон
Ε ε	эпсилон	Ν ν	ни	Φ φ	фи
Z ζ	дзета	Ξ ξ	кси	Χ χ	хи
Η η	эта	Ο ο	омикрон	Ψ ψ	пси
Θ θ θ	тэта	Π π	пи	Ω ω	омега

Наряду с указанным произношением также говорят „лямбда“, „мю“ и „ню“.

ЧАСТЬ I

Математические модели физических процессов

1. ОСНОВНЫЕ ФИЗИЧЕСКИЕ СУБСТАНЦИИ

1.1. Особенности постановки задач математической физики

Предметом изучения математической физики является теория *математических моделей* физических явлений. Эта теория находится на стыке математики и физики, поскольку такие модели описывают конкретные физические процессы, а методы построения и исследования этих моделей являются математическими.

Однако методы, характерные для математической физики, применимы для изучения не только физических явлений. Эти методы в настоящее время используют в химии, геологии, биологии, экологии, экономике. Находят они применение и в технике при *математическом моделировании* различных технических систем и устройств, под которым понимают адекватную замену исследуемого технического объекта его математической моделью и ее последующее изучение методами вычислительной математики с привлечением современных вычислительных средств.

Столь широкое распространение методов математической физики связано с большой общностью математических моделей, базирующихся на фундаментальных законах природы, в том числе на законах сохранения таких физических субстанций, как масса, энергия, заряд, количество движения и момент количества движения*. Это приводит, в частности, к тому,

* Довольно часто в физике и механике наряду с термином „количество движения“ используют термин „импульс“

что одни и те же математические модели описывают явления различной природы.

Для математической физики характерно исследование процессов в системе, представляющей обычно некоторую пространственную область, заполненную непрерывной материальной средой (так называемой *сплошной средой*). В такой области величины, описывающие состояние среды и протекающие в ней физические процессы, зависят, как правило, от пространственных координат и времени. Теоретической основой построения математических моделей таких процессов являются механика и электродинамика сплошных сред. В общем случае модели математической физики описывают поведение системы на трех уровнях: взаимодействие системы в целом с внешней средой; взаимодействие между элементарными объемами системы и свойства отдельно взятого элементарного объема.

Взаимодействие на первом уровне находит отражение в формулировке *краевых условий*, т.е. условий на границе области решения задачи, включающих в общем случае граничные и начальные условия. Второму уровню отвечает описание взаимодействия элементарных объемов на основе законов сохранения физических субстанций и их переноса в пространстве, что позволяет построить математические модели процессов переноса этих субстанций. Наконец, третий уровень соответствует установлению *уравнений состояния* среды, т.е. построению математических моделей поведения среды в элементарном объеме.

В упомянутые математические модели входят характерные для задач математической физики уравнения с частными производными [XII], а в некоторых случаях — *интегральные* и *интегро-дифференциальные уравнения*. Такие *уравнения* относят к классу *функциональных*, решением которых являются функции (в отличие от уравнений, решением которых являются числа). Эти модели могут содержать также интегральные функционалы от искомых функций [XV]. На допустимом множестве функций функционал достигает так называемого ста-

ционарного значения, которому отвечает *стационарная точка функционала*, т.е. искомое решение рассматриваемой задачи. В некоторых случаях это значение может соответствовать экстремуму функционала, что обычно связано с формулировкой задачи на основе некоторого вариационного принципа, имеющего определенное физическое содержание [XV].

1.2. Плотность физических субстанций

Количественной характеристикой любой физической субстанции является ее *объемная плотность*, т.е. количество этой субстанции в единице объема. Рассмотрим окрестность точки $M \in \Omega \subset \mathbb{R}^3$, имеющую диаметр d и объем ΔV . Пусть в этом объеме находится масса Δm некоторой *сплошной среды*. Предел

$$\lim_{d \rightarrow 0} \frac{\Delta m}{\Delta V} = \rho(M) \quad (1.1)$$

называют *плотностью среды* в той точке M пространства, к которой стягивается рассматриваемая окрестность при $d \rightarrow 0$. Основной (стандартной) единицей измерения плотности среды является кг/м^3 .

Аналогично можно ввести понятие *объемной плотности энергии* как предел

$$\varepsilon(M) = \lim_{d \rightarrow 0} \frac{\Delta E}{\Delta V}, \quad (1.2)$$

где ΔE — количество энергии в объеме ΔV . Основной единицей измерения объемной плотности энергии является Дж/м^3 . Если объем ΔV содержит электрический заряд ΔQ , то предел

$$\rho_e(M) = \lim_{d \rightarrow 0} \frac{\Delta Q}{\Delta V} \quad (1.3)$$

называют *объемной плотностью электрического заряда*, основной единицей измерения которой является Кл/м^3 .

Напомним [VII], что если функции $\rho(M)$, $\varepsilon(M)$ и $\rho_e(M)$ ограничены в ограниченной замкнутой области $\Omega \subset \mathbb{R}^3$ и непрерывны в Ω всюду, кроме, быть может, некоторого множества точек *объема нуль*, то эти функции интегрируемы в области Ω . В дальнейшем пространственную область Ω будем обозначать так же, как и ее объем V . Тогда для массы m , энергии E и электрического заряда Q в этой области можно записать

$$m = \int_V \rho(M) dV, \quad E = \int_V \varepsilon(M) dV, \quad Q = \int_V \rho_e(M) dV. \quad (1.4)$$

Понятие объемной плотности применимо не только к физическим субстанциям, выражаемым скалярными величинами (массе, энергии, заряду), но и к субстанциям, выражаемым векторными величинами. Пусть векторное поле скорости движения среды задано векторной функцией $\mathbf{v} = \mathbf{v}(t, \mathbf{x})$, зависящей в общем случае от времени t и координат x_i , $i = 1, 2, 3$, радиус-вектора $\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3$, определяющего положение точки пространства относительно прямоугольной системы координат $Ox_1x_2x_3$ с *ортами* $\mathbf{e}_i$, (рис. 1.1). Тогда произведение $\rho \mathbf{v}$ можно назвать вектором **объемной плотности количества движения среды**. Аналогично векторное произведение $\mathbf{x} \times (\rho \mathbf{v})$ назовем вектором **объемной плотности момента количества движения**

среды относительно начала координат. Основными единицами измерения модулей этих векторов являются $\frac{\text{кг}}{\text{м}^2 \cdot \text{с}}$ и $\frac{\text{кг}}{\text{м} \cdot \text{с}}$, или $\frac{\text{Н} \cdot \text{с}}{\text{м}^3}$ и $\frac{\text{Н} \cdot \text{с}}{\text{м}^2}$ соответственно.

Если $\rho \mathbf{v}$ — непрерывная функция координат в пространственной области объемом V всюду, кроме, быть может, некоторого множест-

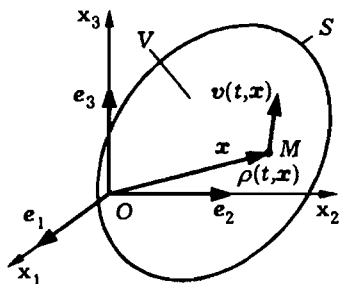


Рис. 1.1

ва точек объема нуль, то для находящейся в этой области среды векторы количества движения $\mathbf{J}$ и момента количества движения $\mathbf{K}$ можно представить в виде

$$\mathbf{J} = \int_V \rho \mathbf{v} dV, \quad \mathbf{K} = \int_V \mathbf{x} \times \rho \mathbf{v} dV. \quad (1.5)$$

Замечание 1.1. Подынтегральные функции в (1.4) и (1.5) могут в общем случае зависеть от времени t . Тогда значения соответствующих интегралов также могут изменяться во времени. Пусть действительная функция $C = C(t, \mathbf{x})$ непрерывно дифференцируема по t и задает в текущий момент времени зависимость от пространственных координат объемной плотности некоторой физической субстанции. Количество $\bar{C}(t)$ этой субстанции в фиксированный момент времени t в объеме V выражает интеграл

$$\bar{C}(t) = \int_V C(t, \mathbf{x}) dV, \quad (1.6)$$

зависящий от параметра t [VII].

Пусть точки поверхности S , ограничивающей объем V , движутся вместе с частицами среды, находящимися в этих точках, с заданной скоростью $\mathbf{v}(t, \mathbf{x})$ относительно системы координат $Ox_1x_2x_3$ (рис. 1.2). Тогда в момент времени $t + \Delta t$ поверхность

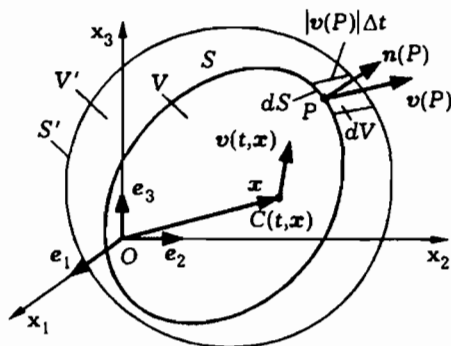


Рис. 1.2

переместится в положение S' и будет ограничивать объем V' . Сначала рассмотрим случай, когда $V \subset V'$. Предполагая, что интеграл в (1.6) дифференцируем по параметру t , найдем скорость изменения количества субстанции. Согласно определению производной [11],

$$\begin{aligned} \frac{d\bar{C}(t)}{dt} &= \frac{d}{dt} \int_V C(t, \mathbf{x}) dV = \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \left(\int_{V'} C(t + \Delta t, \mathbf{x}) dV - \int_V C(t, \mathbf{x}) dV \right) = \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \left(\int_V (C(t + \Delta t, \mathbf{x}) - C(t, \mathbf{x})) dV + \int_{V^*} C(t + \Delta t, \mathbf{x}) dV \right) = \\ &= \int_V \left(\lim_{\Delta t \rightarrow 0} \frac{C(t + \Delta t, \mathbf{x}) - C(t, \mathbf{x})}{\Delta t} \right) dV + \\ &\quad + \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{V^*} C(t + \Delta t, \mathbf{x}) dV, \quad (1.7) \end{aligned}$$

где $V^* = V' \setminus V$ — объем, который омекает поверхность S при перемещении в положение S' . Этот объем представим состоящим из элементарных цилиндров, каждый из которых имеет основание dS и образующую длиной $|\mathbf{v}(P)|\Delta t$, параллельную вектору $\mathbf{v}(P) = \mathbf{v}(t, \mathbf{x}(P))$ скорости среды в точке $P \in S$, причем $\mathbf{n}(P)$ — единичный вектор внешней нормали к поверхности S в точке $P \in S$ (см. рис. 1.2). Тогда элементарный объем во втором интеграле в правой части (1.7) запишем в виде $dV = \mathbf{v}(t, \mathbf{x})\mathbf{n}(P)\Delta t dS \geq 0$, $P \in S$, и получим

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{V^*} C(t + \Delta t, \mathbf{x}) dV &= \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_S C(t + \Delta t, \mathbf{x}) \mathbf{v}(t, \mathbf{x}) \mathbf{n} \Delta t dS = \int_S C(t, \mathbf{x}) \mathbf{v}(t, \mathbf{x}) \mathbf{n} dS, \end{aligned}$$

поскольку $C(t + \Delta t, \mathbf{x}) \rightarrow C(t, \mathbf{x})$ при $\Delta t \rightarrow 0$ в силу непрерывности функции $C(t, \mathbf{x})$.

Предполагая непрерывную дифференцируемость по координатам функций C и $\mathbf{v}$, преобразуем интеграл в правой части последнего равенства по формуле Остроградского — Гаусса. В итоге, учитывая, что подынтегральная функция первого интеграла в правой части (1.7) равна частной производной $\frac{\partial C(t, \mathbf{x})}{\partial t}$, $\mathbf{x} \in V$, получаем формулу дифференцирования по времени t интеграла, взятого по изменяемому во времени объему:

$$\begin{aligned} \frac{d\bar{C}(t)}{dt} &= \frac{d}{dt} \int_V C(t, \mathbf{x}) dV = \\ &= \int_V \left(\frac{\partial C(t, \mathbf{x})}{\partial t} + \nabla(C(t, \mathbf{x})\mathbf{v}(t, \mathbf{x})) \right) dV, \quad (1.8) \end{aligned}$$

где ∇ — дифференциальный оператор Гамильтона

$$\nabla = \frac{\partial}{\partial x_1} \mathbf{e}_1 + \frac{\partial}{\partial x_2} \mathbf{e}_2 + \frac{\partial}{\partial x_3} \mathbf{e}_3, \quad (1.9)$$

$\mathbf{e}_i, i = 1, 2, 3$, — орты системы координат $Ox_1x_2x_3$ (см. рис. 1.2). Напомним, что действие этого оператора на векторную функцию соответствует операции дивергенции.

Отметим, что к (1.8) можно прийти и в случае, когда $V' \subset V$, но при этом во втором интеграле в правой части (1.7) $dV = \mathbf{v}(t, \mathbf{x})\mathbf{n}(P)\Delta t dS \leq 0$, $P \in S$. Можно показать*, что (1.8) применимо при произвольном изменении объема V во времени. Для фиксированного объема V_0 , ограниченного неподвижной относительно системы координат $Ox_1x_2x_3$ поверхностью S_0 , вместо (1.8) будем иметь

$$\frac{d\bar{C}(t)}{dt} = \frac{d}{dt} \int_{V_0} C(t, \mathbf{x}) dV = \int_{V_0} \frac{\partial C(t, \mathbf{x})}{\partial t} dV. \quad (1.10)$$

*См.: Седов Л.И.

1.3. Перенос физических субстанций

Большинство изучаемых в математической физике процессов связано с переносом в пространстве конкретных физических субстанций: массы, энергии, электрического заряда, количества движения или его момента. Интенсивность переноса физической субстанции определяют **плотностью потока**, равной количеству субстанции, переносимой в единицу времени через единичную площадку, перпендикулярную направлению переноса. Среди механизмов переноса выделяют конвективный (или молярный) и диффузионный (или молекулярный).

Первый из них связан с движением *сплошной среды*, определяемым векторным полем ее скорости $\mathbf{v} = \mathbf{v}(t, M)$, $M \in \mathbb{R}^3$, в момент времени t . Для физической субстанции, выражаемой скалярной величиной, плотность потока конвективного переноса является вектором, коллинеарным вектору скорости $\mathbf{v}$ и равным произведению $\mathbf{v}$ и *объемной плотности* этой субстанции. Так, направление и интенсивность конвективного переноса массы определяет вектор **плотности потока массы** $\rho \mathbf{v}$, где ρ — *плотность среды*. Модуль этого вектора равен количеству массы, переносимой в единицу времени через единичную площадку, перпендикулярную направлению вектора скорости. Основная единица измерения плотности потока массы (как и модуля вектора *объемной плотности количества движения среды*) — $\frac{\text{кг}}{\text{м}^2 \cdot \text{с}}$.

Направление и интенсивность конвективного переноса энергии и заряда определяют векторами $\epsilon \mathbf{v}$ и $\rho_e \mathbf{v}$ **плотности потока энергии** и **плотности электрического тока**, где ϵ и ρ_e — *объемные плотности энергии и электрического заряда* соответственно. Модули этих векторов измеряют в $\text{Вт}/\text{м}^2$ и $\text{А}/\text{м}^2$ соответственно.

Диффузионный перенос физических субстанций может происходить и при отсутствии направленного движения среды (например, при хаотическом молекулярном движении в жидкости,

газе или плазме и тепловом движении ионов, атомов и молекул в твердом теле). При неравномерном пространственном распределении в среде объемной плотности C некоторой физической субстанции хаотическое движение микрочастиц среды постепенно приводит к выравниванию этого распределения. В изотропной среде, свойства которой одинаковы во всех направлениях, диффузионный перенос физической субстанции, вызванный неравномерным пространственным распределением скалярной величины C , происходит в направлении убывания объемной плотности, т.е. в направлении, противоположном направлению градиента ∇C скалярного поля, задаваемого в пространстве в текущий момент времени t функцией $C = C(t, M)$.

При построении математических моделей в математической физике широко используют эмпирический закон диффузионного переноса

$$j^{(C)} = -K^{(C)} \nabla C, \quad (1.11)$$

где $j^{(C)}$ — вектор плотности потока физической субстанции при диффузионном переносе; $K^{(C)}$ — эмпирический коэффициент диффузионного переноса этой субстанции. Функцию $C = C(t, \mathbf{x})$ обычно предполагают непрерывно дифференцируемой необходимое число раз по всем ее аргументам. Она выполняет роль *потенциальной функции* по отношению к векторному полю плотности потока этой субстанции при ее диффузионном переносе.

Например, функция $C = C(t, M)$ может задавать распределение в среде объемной плотности некоторого вещества (примеси в жидкости или газе, ионов в плазме, легирующего элемента в сплаве). В этом случае $K^{(C)}$ называют коэффициентом диффузии данного вещества в этой среде, а (1.11) выражает известный в физике закон Фика*.

Интенсивность диффузионного переноса физической субстанции не всегда связывают с градиентом скалярного поля объемной плотности этой субстанции. Так, в математических

*А. Э. Фик (1829–1901) — немецкий физик.

моделях процесса распространения в среде теплоты как одной из форм энергии в качестве потенциальной функции вместо ε используют функцию $T = T(t, M)$ распределения в пространстве в текущий момент времени t температуры, характеризующей при определенных условиях объемную плотность тепловой энергии среды. Это приводит к установленному французским математиком и физиком Ж.Б.Ж. Фурье (1768–1830) эмпирическому закону теплопроводности [XII]

$$\mathbf{q} = -\lambda \nabla T, \quad (1.12)$$

где $\mathbf{q}$ — вектор плотности теплового потока, λ — теплопроводность среды.

Линейную связь вектора плотности потока физической субстанции с градиентом некоторой потенциальной функции используют и в тех случаях, когда перенос этой субстанции происходит путем движения микрочастиц под действием внешнего поля. Так, под действием электростатического поля, описываемого потенциальной функцией $U = U(M)$ электрического напряжения, в электропроводящей среде возникает электрический ток, вектор плотности которого равен

$$\mathbf{j}^{(e)} = -\sigma \nabla U, \quad (1.13)$$

где σ — электрическая проводимость среды. Если ввести вектор $\mathbf{E} = -\nabla U$ напряженности электростатического поля, то из (1.13) получим уравнение

$$\mathbf{j}^{(e)} = \sigma \mathbf{E}, \quad (1.14)$$

обобщающее закон Ома* на случай сплошной среды (модули векторов $\mathbf{j}^{(e)}$ и $\mathbf{E}$ измеряют в А/м² и В/м соответственно).

При неравномерном распределении давления, заданном функцией $p = p(t, M)$, через пористую среду может просачиваться

*Г.С. Ом (1787–1854) — немецкий физик.

жидкость или газ. Тогда вектор скорости частиц жидкости или газа подчиняется закону Дарси*

$$\mathbf{v} = -\kappa \nabla p, \quad (1.15)$$

где κ --- коэффициент фильтрации.

Потенциальная функция в соотношении (1.11) может зависеть от пространственных распределений нескольких физических величин. Например, для многокомпонентной смеси химически реагирующих веществ диффузионный перенос физических субстанций связан с выравниванием неравномерного пространственного распределения так называемого химического потенциала, который зависит от концентрации этих веществ, температуры и давления. В этом случае вектор плотности потока конкретной субстанции будет линейной комбинацией векторов, коллинеарных градиентам концентрации, температуры и давления соответственно. Тогда говорят о концентрационной диффузии субстанции, ее термо- и бародиффузии.

Не затрагивая особенностей различных процессов диффузионного переноса, ограничимся лишь констатацией того, что в большинстве случаев вектор $\mathbf{j}^{(C)}$ плотности потока конкретной физической субстанции можно считать линейно зависящим от градиента некоторой потенциальной функции $\Phi(t, M)$, которая не всегда совпадает с объемной плотностью C этой субстанции. Эту зависимость можно записать в общей форме

$$\mathbf{j}^{(C)} = -K^{(\Phi)} \nabla \Phi, \quad (1.16)$$

где $K^{(\Phi)}$ — коэффициент пропорциональности. Если среда анизотропна, т.е. ее свойства различны в различных направлениях, то вместо (1.16) используют соотношения

$$j_i^{(C)} = - \sum_{j=1}^3 K_{ij}^{(\Phi)} \frac{\partial \Phi}{\partial x_j}, \quad i = 1, 2, 3, \quad (1.17)$$

*А. Дарси (1803–1858) — французский инженер.

где $j_i^{(C)}$ — проекции вектора $j^{(C)}$ на координатные оси Ox_i прямоугольной системы координат $Ox_1x_2x_3$; $K_{ij}^{(\Phi)}$ — компоненты тензора второго ранга коэффициентов переноса конкретной субстанции.

Представленные соотношения характерны для диффузионного переноса субстанций, объемная плотность которых является скалярной величиной. Объемные плотности количества движения и момента количества движения являются векторными величинами, что усложняет выражения для плотностей потоков этих субстанций при диффузионном переносе.

Дополнение 1.1. Некоторые формулы векторного анализа

При постановке задач математической физики и преобразовании описывающих их уравнений часто приходится использовать формулы векторного анализа, в том числе *формулу Остроградского — Гаусса*. Напомним, что она связывает между собой интеграл по пространственной области $V \subset \mathbb{R}^3$ от дивергенции векторного поля, задаваемого векторной функцией $u(M)$, $M \in V$, и интеграл по ограничивающей V кусочно гладкой поверхности S от проекции $u_n(P) = u(P)n(P)$ вектора $u(P)$ на направление единичного вектора $n(P)$ внешней нормали к S в точке $P \in S$:

$$\int_V \nabla u dV = \int_S u_n dS = \int_S u n dS. \quad (1.18)$$

При этом полагают, что векторная функция $u(x)$ непрерывно дифференцируема на замыкании $\bar{V} = V \cup S$ области V по всем координатам x_i , $i = 1, 2, 3$, радиус-вектора $x \in \mathbb{R}^3$, определяющего положение точки пространства относительно прямоугольной системы координат $Ox_1x_2x_3$ (см. рис. 1.1). Отметим, что поверхностный интеграл в (1.18) берется по всем участкам поверхности S , ограничивающей область V . Ясно, что в

двумерном случае V соответствует плоской области (не обязательно односвязной), ограниченной замкнутыми плоскими кривыми (контурами), составляющими ее границу S .

М.В. Остроградский в 1834 году получил формулу (1.18) для произвольного многомерного случая. Пусть ограниченная область $\Omega \subset \mathbb{R}^m$ имеет границу $\partial\Omega$, заданную уравнением $f(\tilde{M}) = 0$, $\tilde{M} \in \Omega$, причем функция f непрерывно дифференцируема по координатам $x_i(\tilde{M})$, $i = \overline{1, m}$, точки $\tilde{M}$. При этом частные производные $\frac{\partial f}{\partial x_i}$ не обращаются одновременно в нуль ни в одной точке на $\partial\Omega$, т.е.

$$|\tilde{\nabla} f| = \sqrt{\sum_{i=1}^m \left(\frac{\partial f}{\partial x_i}\right)^2} \neq 0,$$

где $\tilde{\nabla}$ — дифференциальный оператор

$$\tilde{\nabla} = \sum_{i=1}^m e_i \frac{\partial}{\partial x_i}. \quad (1.19)$$

Здесь e_i — орты ортонормированного базиса в $\mathbb{R}^m$. Поэтому в любой точке $\tilde{P} \in \partial\Omega$ границы $\partial\Omega$ существует единичный вектор

$$\tilde{n} = \frac{\tilde{\nabla} f}{|\tilde{\nabla} f|} \quad (1.20)$$

внешней нормали к $\partial\Omega$, если знак функции f выбрать так, чтобы было $f < 0$ для точек, принадлежащих области Ω . Поскольку частные производные функции f непрерывны, то направляющие косинусы

$$n_i = \frac{1}{|\tilde{\nabla} f|} \frac{\partial f}{\partial x_i}, \quad i = \overline{1, m},$$

вектора $\tilde{n}(\tilde{P})$ относительно ортов e_i , также непрерывно зависят от координат $x_i(\tilde{P})$, $i = \overline{1, m}$, точки $\tilde{P} \in \partial\Omega$.

Представим интеграл по области Ω от непрерывной на ее замыкании $\bar{\Omega} = \Omega \cup \partial\Omega$ действительной функции $w(\tilde{M})$, $\tilde{M} \in \bar{\Omega}$, в виде m -кратного интеграла:

$$\int_{\Omega} w d\Omega = \int_{\Omega} \dots \int_{\Omega} w dx_1 dx_2 \dots dx_m. \quad (1.21)$$

Если область Ω в $\mathbb{R}^m$ является прямоугольным параллелепипедом, т.е. $x_i \in (a_i, b_i)$, $i = \overline{1, m}$, то интеграл в (1.21) можно найти последовательным вычислением определенных интегралов:

$$\int_{\Omega} w d\Omega = \int_{a_1}^{b_1} dx_1 \int_{a_2}^{b_2} dx_2 \dots \int_{a_m}^{b_m} w(x_1, x_2, \dots, x_m) dx_m.$$

В более общем случае предположим, что точки, принадлежащие области $\Omega \subset \mathbb{R}^m$, образуют в линейном пространстве $\mathbb{R}^m$ *выпуклое подмножество*, т.е. при произвольно выбранных двух точках $\tilde{M}', \tilde{M}'' \in \bar{\Omega}$ с координатами $x_i(\tilde{M}'')$ и $x_i(\tilde{M}')$, $i = \overline{1, m}$, соответственно точка $\tilde{M}$ с координатами

$$x_i(\tilde{M}) = \alpha x_i(\tilde{M}') + (1 - \alpha) x_i(\tilde{M}''), \quad i = \overline{1, m}, \quad \alpha \in [0, 1],$$

также принадлежит замыканию $\bar{\Omega}$ области Ω и для этой точки $f(\tilde{M}) \leq 0$ при любых $\alpha \in [0, 1]$. Геометрически это означает, что прямая, параллельная, например, орту e_m и проходящая через любую точку $\tilde{M} \in \bar{\Omega}$, пересечет границу $\partial\Omega$ не более чем в двух точках $\tilde{M}_*, \tilde{M}^* \in \partial\Omega$ (рис. 1.3).

Иными словами, если координаты $x_1, x_2, \dots, x_{m-1}$ фиксированы, то при движении точки $\tilde{M} \in \Omega$ по этой прямой координата x_m пробегает интервал (x'_m, x''_m) , где $x'_m = f_*(x_1, x_2, \dots, x_{m-1})$ и $x''_m = f^*(x_1, x_2, \dots, x_{m-1})$. Ясно, что подстановка x'_m и x''_m вместо x_m в уравнение $f(\tilde{x}) = 0$ границы $\partial\Omega$ обращает его в

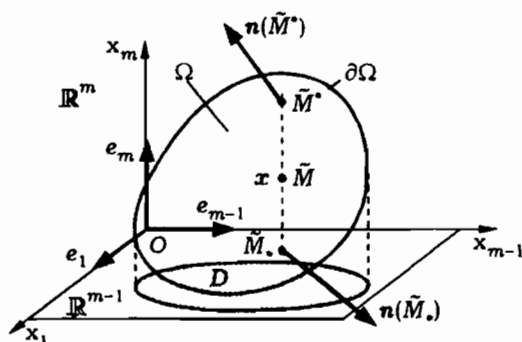


Рис. 1.3

тождество. Тогда по аналогии с вычислением тройных интегралов вместо (1.21) можно записать

$$\int_{\Omega} w d\Omega = \int_D \cdots \int_{f_*(x_1, x_2, \dots, x_{m-1})}^{f^*(x_1, x_2, \dots, x_{m-1})} dx_1 dx_2 \dots dx_{m-1} \int w dx_m, \quad (1.22)$$

где $D \subset \mathbb{R}^{m-1}$ — область, которая является проекцией области $\Omega \subset \mathbb{R}^m$ на гиперплоскость $\mathbb{R}^{m-1}$, ортогональную орту e_m (см. рис. 1.3).

Пусть $w = \frac{\partial u_m}{\partial x_m}$, где $u_m(x_1, x_2, \dots, x_m)$ — функция, непрерывная вместе со своей производной $\frac{\partial u}{\partial x_m}$ на замыкании $\bar{\Omega}$. В этом случае (1.22) примет вид

$$\begin{aligned} \int_{\Omega} \frac{\partial u_m}{\partial x_m} d\Omega &= \int_D dD \int_{f_*(x_1, x_2, \dots, x_{m-1})}^{f^*(x_1, x_2, \dots, x_{m-1})} \frac{\partial u_m}{\partial x_m} dx_m = \\ &= \int_D u_m(x_1, x_2, \dots, x_{m-1}, f^*(x_1, x_2, \dots, x_{m-1})) dD - \\ &\quad - \int_D u_m(x_1, x_2, \dots, x_{m-1}, f_*(x_1, x_2, \dots, x_{m-1})) dD. \end{aligned}$$

Так как аргументы подынтегральной функции u_m интегралов в правой части этого равенства являются координатами точек $\widetilde{M}^*$ и $\widetilde{M}_*$ границы $\partial\Omega$, то в первом из этих интегралов $dD = n_m(\widetilde{M}^*)d(\partial\Omega(\widetilde{M}^*))$, а во втором интеграле $dD = -n_m(\widetilde{M}_*)d(\partial\Omega(\widetilde{M}_*))$, где n_m является направляющим косинусом (относительно орта e_m) единичного вектора $\tilde{n}$ внешней нормали к $\partial\Omega$ в соответствующих точках, а $d(\partial\Omega)$ — элементарным участком границы $\partial\Omega$ в окрестности каждой из этих точек. Поэтому эти интегралы можно заменить одним интегралом по границе $\partial\Omega$ и в итоге записать

$$\int_{\Omega} \frac{\partial u_m}{\partial x_m} d\Omega = \int_{\partial\Omega} u_m n_m d(\partial\Omega). \quad (1.23)$$

Ясно, что произвольную область $\Omega \subset R^n$ можно разбить внутренними границами на такие подобласти, для которых сохраняет силу (1.23). Объединяя интегралы по этим подобластям и учитывая, что интегралы по участкам внутренних границ взаимно уничтожатся, поскольку направляющие косинусы внешних нормалей в точках таких границ имеют противоположные знаки, заключаем, что (1.23) справедливо для произвольной области с гладкой границей. Но интеграл в правой части (1.23) не изменит своего значения, если множество точек, в которых подынтегральная функция терпит разрыв, имеет меру Лебега, равную нулю. Поэтому (1.23) верно и в случае кусочно гладкой границы $\partial\Omega$.

Аналогично (1.23) для любой функции $u_i(x_1, x_2, \dots, x_m)$, $i = \overline{1, m}$, непрерывной вместе со своей производной $\frac{\partial u_i}{\partial x_i}$ на замыкании $\overline{\Omega}$, можно записать

$$\int_{\Omega} \frac{\partial u_i}{\partial x_i} d\Omega = \int_{\partial\Omega} u_i n_i d(\partial\Omega), \quad i = \overline{1, m}. \quad (1.24)$$

Суммируя по i эти равенства, получаем **формулу Остроградского**

$$\int_{\Omega} \sum_{i=1}^m \frac{\partial u_i}{\partial x_i} d\Omega = \int_{\partial\Omega} \sum_{i=1}^m u_i n_i d(\partial\Omega) \quad (1.25)$$

для области $\Omega \subset \mathbb{R}^m$ в m -мерном пространстве $\mathbb{R}^m$. Обозначая $\tilde{\mathbf{u}} = u_1 \mathbf{e}_1 + u_2 \mathbf{e}_2 + \dots + u_m \mathbf{e}_m$ и учитывая (1.19) и правила скалярного умножения в евклидовом арифметическом пространстве $\mathbb{R}^m$, вместо (1.25) приходим к формуле Остроградского в виде

$$\int_{\Omega} \tilde{\nabla} \tilde{\mathbf{u}} d\Omega = \int_{\partial\Omega} \tilde{\mathbf{u}} \tilde{\mathbf{n}} d(\partial\Omega), \quad (1.26)$$

где $\tilde{\mathbf{n}} = n_1 \mathbf{e}_1 + n_2 \mathbf{e}_2 + \dots + n_m \mathbf{e}_m$.

Пусть в (1.26) $\tilde{\mathbf{u}} = \tilde{v} \tilde{\nabla} \tilde{u}$, где $\tilde{u}$, $\tilde{v}$ — действительные функции, дважды непрерывно дифференцируемые на замыкании $\bar{\Omega}$. Тогда, выполняя в соответствии с (1.19) дифференцирование, находим $\tilde{\nabla} \tilde{\mathbf{u}} = \tilde{v} \tilde{\nabla} (\tilde{\nabla} \tilde{u}) + (\tilde{\nabla} \tilde{v}) \tilde{\nabla} \tilde{u} = \tilde{v} \tilde{\nabla}^2 \tilde{u} + (\tilde{\nabla} \tilde{v}) \tilde{\nabla} \tilde{u}$ и после подстановки в (1.26) получаем

$$\int_{\Omega} (\tilde{v} \tilde{\nabla}^2 \tilde{u} + (\tilde{\nabla} \tilde{v}) \tilde{\nabla} \tilde{u}) d\Omega = \int_{\partial\Omega} \tilde{v} (\tilde{\nabla} \tilde{u}) \tilde{\mathbf{n}} d(\partial\Omega),$$

или

$$\int_{\Omega} \tilde{v} \tilde{\nabla}^2 \tilde{u} d\Omega = \int_{\partial\Omega} \tilde{v} (\tilde{\nabla} \tilde{u}) \tilde{\mathbf{n}} d(\partial\Omega) - \int_{\Omega} (\tilde{\nabla} \tilde{v}) \tilde{\nabla} \tilde{u} d\Omega. \quad (1.27)$$

Меняя в (1.27) местами функции $\tilde{u}$ и $\tilde{v}$ и вычитая полученный результат из (1.27), имеем

$$\int_{\Omega} (\tilde{v} \tilde{\nabla}^2 \tilde{u} - \tilde{u} \tilde{\nabla}^2 \tilde{v}) d\Omega = \int_{\partial\Omega} (\tilde{v} \tilde{\nabla} \tilde{u} - \tilde{u} \tilde{\nabla} \tilde{v}) \tilde{\mathbf{n}} d(\partial\Omega). \quad (1.28)$$

В трехмерном случае ($m = 3$) оператор $\tilde{\nabla}$ соответствует дифференциальному оператору Гамильтона ∇ , а $\tilde{\nabla}^2$ — оператору Лапласа ∇^2 , часто обозначаемому Δ . Обозначив в (1.27) и (1.28) трехмерную область Ω символом V , а поверхность $\partial\Omega$ символом S , приходим к известным *первой и второй формулам Грина*

$$\int_V v \nabla^2 u \, dV = \int_S v (\nabla u) \mathbf{n} \, dS - \int_V (\nabla v) \nabla u \, dV, \quad (1.29)$$

$$\int_V (v \nabla^2 u - u \nabla^2 v) \, dV = \int_S (v \nabla u - u \nabla v) \mathbf{n} \, dS, \quad (1.30)$$

установленным английским математиком и физиком Дж. Гринном (1793–1841) в 1828 году в связи с его исследованиями по теории потенциала. Здесь u и v — действительные функции, дважды непрерывно дифференцируемые на замыкании $\bar{V}$ области V ; $\mathbf{n}$ — единичный вектор внешней нормали к S . Равенство (1.29) иногда называют предварительной формулой Грина. Таким образом, (1.27) и (1.28) можно рассматривать как обобщение первой и второй формул Грина на m -мерный случай.

Отметим, что при приближенном решении задач математической физики нередко возникает необходимость в преобразовании интегралов по области V , содержащих вместо $\nabla^2 u$ выражение вида $\nabla(\lambda \nabla u)$, где $\lambda(M)$ — действительная функция, имеющая на замыкании $\bar{V}$ кусочно непрерывные производные по координатам точки $M \in \bar{V}$. В этом случае вместо формул (1.29) и (1.30) будем иметь

$$\int_V v \nabla(\lambda \nabla u) \, dV = \int_S \lambda v (\nabla u) \mathbf{n} \, dS - \int_V \lambda (\nabla v) \nabla u \, dV, \quad (1.31)$$

$$\int_V v \nabla(\lambda \nabla u) = \int_V u \nabla(\lambda \nabla v) \, dV + \int_S \lambda (v \nabla u - u \nabla v) \mathbf{n} \, dS. \quad (1.32)$$

Ясно, что при $\lambda \equiv \text{const}$ эти формулы переходят в формулы (1.29) и (1.30).

Вопросы и задачи

1.1. Доказать, что если функция $f(r)$ дважды непрерывно дифференцируема и $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$, то $\nabla f(r) = f'(r)\mathbf{x}/r$ и $\nabla^2 f(r) = f''(r) + 2f'(r)/r$, где $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3$ и $\mathbf{e}_i$, $i = 1, 2, 3$, — орты прямоугольной системы координат $Ox_1x_2x_3$.

1.2. Доказать, что объем V , ограниченный поверхностью S , можно вычислить по формуле

$$V = \frac{1}{6} \int_S (\nabla \mathbf{x}^2) \mathbf{n} dS,$$

где $\mathbf{x}$ — радиус-вектор точки поверхности, а $\mathbf{n}$ — единичный вектор внешней нормали к поверхности в этой точке.

1.3. Доказать справедливость формул (1.31), (1.32).

2. ЗАКОНЫ СОХРАНЕНИЯ ФИЗИЧЕСКИХ СУБСТАНЦИЙ

2.1. Закон сохранения массы

Пусть скалярное поле *плотности среды* и векторное поле скорости движения этой среды заданы в некоторой области $\Omega \subset \mathbb{R}^3$ непрерывно дифференцируемыми функциями $\rho = \rho(t, M)$ и $\mathbf{v} = \mathbf{v}(t, M)$ времени t и координат x_i , $i = 1, 2, 3$, точки $M \in \Omega$ в прямоугольной системе координат $Ox_1x_2x_3$. Рассмотрим в этой области произвольную подобласть V , объем которой обозначим тем же символом V . В этой подобласти в некоторый момент времени t находится в соответствии с (1.4) масса *сплошной среды*

$$m = \int_{V(t)} \rho(t, M) dV.$$

Пусть точки поверхности S , ограничивающей объем V , движутся вместе с частицами среды. Тогда масса среды в этом изменяющемся во времени объеме $V(t)$ остается постоянной, поскольку частицы среды не пересекают поверхность S . Поэтому в соответствии с (1.8) имеем

$$\frac{dm}{dt} = \frac{d}{dt} \int_{V(t)} \rho(t, M) dV = \int_{V(t)} \left(\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) \right) dV = 0. \quad (2.1)$$

Это равенство является одной из интегральных форм закона сохранения массы.

В фиксированный момент времени t выделим некоторый объем $V_n(t) \subset V(t)$, ограниченный поверхностью S_n , точки которой также движутся вместе с частицами среды. Поэтому

(2.1) сохраняет силу и для объема $V_n(t)$. Обозначим через d_n диаметр области, соответствующий этому объему. Пусть последовательность $\{V_n(t)\}$ объемов $V_n(t) \subset V(t)$, $n \in \mathbb{N}$, такова, что при $d_n \rightarrow 0$ они стягиваются в точку $M \in \Omega$. Тогда при условии непрерывности подынтегральной функции в (2.1), согласно теореме о среднем значении для тройного интеграла [VII], из (2.1) следует, что

$$\begin{aligned} \lim_{d_n \rightarrow 0} \frac{1}{V_n(t)} \int_{V_n(t)} \left(\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) \right) dV = \\ = \frac{\partial \rho(t, M)}{\partial t} + \nabla(\rho(t, M) \mathbf{v}(t, M)) = 0. \end{aligned}$$

Отсюда получаем локальную форму закона сохранения массы

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0. \quad (2.2)$$

В системе координат, движущейся вместе с частицей среды, изменение плотности во времени описывается полной производной $d\rho/dt = \frac{\partial \rho}{\partial t} + \mathbf{v}(\nabla \rho)$, равной сумме локальной $\frac{\partial \rho}{\partial t}$ и конвективной (или переносной) производных. Так как $\nabla(\rho \mathbf{v}) = \rho \nabla \mathbf{v} + \mathbf{v}(\nabla \rho)$, то вместо (2.2) можно написать

$$\frac{d\rho}{dt} + \rho \nabla \mathbf{v} = 0. \quad (2.3)$$

В механике сплошной среды (2.2) и (2.3) называют **уравнениями неразрывности**.

Если среда несжимаема, то ее плотность неизменна, хотя в случае неоднородной среды в различных точках пространства и в различные моменты времени она может иметь разную плотность. Но в системе координат, движущейся вместе с любой частицей несжимаемой среды, $d\rho/dt = 0$. В этом случае уравнение неразрывности принимает наиболее простой вид

$$\nabla \mathbf{v} = 0. \quad (2.4)$$

Ясно, что (2.4) справедливо и для однородной несжимаемой среды при $\rho = \text{const}$, что следует и из (2.2).

Отметим, что в случае несжимаемой среды операции интегрирования по изменяемому во времени объему и дифференцирования по времени можно менять местами. Действительно, для объемной плотности C некоторой физической субстанции, используя (1.8) и выражение для полной производной, имеем

$$\frac{d}{dt} \int_V C dV = \int_V \left(\frac{\partial C}{\partial t} + \nabla(Cv) \right) dV = \int_V \left(\frac{dC}{dt} + C \nabla v \right) dV. \quad (2.5)$$

Отсюда в случае несжимаемой среды с учетом (2.4) заключаем, что

$$\frac{d}{dt} \int_V C dV = \int_V \frac{dC}{dt} dV.$$

Если в (2.5) положить $C' \equiv 1$, то в случае сжимаемой среды с учетом формулы Остроградского — Гаусса будем иметь

$$\frac{dV}{dt} = \int_V \nabla v dV = \int_S v n dS.$$

Таким образом, дивергенция $\nabla v(t, M)$ является локальной скоростью относительного увеличения элементарного объема в окрестности точки M в момент времени t . Очевидно, что для несжимаемой среды в соответствии с (2.4) эта скорость равна нулю, т.е. объем области, занятой средой, остается постоянным, но область может изменять свою форму.

Пусть $C = \rho C_m$, где $C_m = C_m(t, M)$ — количество физической субстанции, приходящейся на единицу массы среды, находящейся в момент времени t в окрестности точки M . Тогда, заменяя в (2.5) C' на ρC_m и учитывая (2.3), в случае непрерывной дифференцируемости по времени t функций $\rho(t, M)$ и

$C_m(t, M)$ получаем

$$\begin{aligned} \frac{d}{dt} \int_V \rho C_m dV &= \int_V \left(\frac{d(\rho C_m)}{dt} + \rho C_m \nabla \mathbf{v} \right) dV = \\ &= \int_V \rho \frac{dC_m}{dt} dV + \int_V C_m \left(\frac{d\rho}{dt} + \rho \nabla \mathbf{v} \right) dV = \int_V \rho \frac{dC_m}{dt} dV. \quad (2.6) \end{aligned}$$

Пример 2.1. Пусть скорость среды, заполняющей внешность $\Omega \subset \mathbb{R}^3$ сферы радиуса r_0 с центром в начале прямоугольной системы координат $Ox_1x_2x_3$, задана векторной функцией

$$\mathbf{v} = v_0(t) \frac{r_0^2 \mathbf{x}}{|\mathbf{x}|^3} \quad (2.7)$$

времени t и радиус-вектора $\mathbf{x}$ точки $M \in \Omega$ области Ω , причем $v_0(t)$ имеет смысл модуля вектора скорости на поверхности этой сферы. Так как вектор $\mathbf{v}$ скорости среды в любой точке $M \in \Omega$ сонаправлен радиус-вектору $\mathbf{x}$ этой точки, а модуль $|\mathbf{v}|$ зависит лишь от расстояния точки $M \in \Omega$ от начала координат, то векторное поле скорости среды является центральным [VII].

С учетом того, что проекция вектора скорости на ось Ox_i , $i = 1, 2, 3$, равна $v_i = v_0(t) r_0^2 x_i / |\mathbf{x}|^3$ и $|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + x_3^2}$, найдем дивергенцию этого поля:

$$\begin{aligned} \nabla \mathbf{v} &= \sum_{i=1}^3 \frac{\partial v_i}{\partial x_i} = v_0(t) r_0^2 \sum_{i=1}^3 \frac{\partial (x_i / |\mathbf{x}|^3)}{\partial x_i} = \\ &= \frac{3v_0(t) r_0^2}{|\mathbf{x}|^3} - \frac{3v_0(t) r_0^2}{2|\mathbf{x}|^5} \sum_{i=1}^3 2x_i^2 = 0. \end{aligned}$$

Итак, в любой точке области функция $\mathbf{v} = v_0(t) r_0^2 \mathbf{x} / |\mathbf{x}|^3$ удовлетворяет (2.4), т.е., согласно (2.3), среда является несжимаемой (но не обязательно однородной).

Нетрудно убедиться, что ротор этого поля скорости равен

$$\nabla \times \mathbf{v} = \left(\frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \right) \mathbf{e}_1 + \left(\frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \right) \mathbf{e}_2 + \left(\frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \right) \mathbf{e}_3 = \mathbf{0},$$

т.е. векторное поле является не только соленоидальным (в силу $\nabla \mathbf{v} \equiv 0$), но и безвихревым, а так как область Ω поверхностно односвязна, то и потенциальным [VII]. По определению потенциального поля существует такая действительная функция $\varphi(t, \mathbf{x})$, что $\mathbf{v}(t, \mathbf{x}) = \nabla \varphi(t, \mathbf{x})$. Эту функцию называют *скалярным потенциалом векторного поля скорости*, или коротко *потенциалом скорости*. В этом случае (2.4) переходит в уравнение Лапласа

$$\nabla^2 \varphi = 0, \quad (2.8)$$

где $\nabla^2 = \Delta$ — дифференциальный оператор Лапласа. Граничное условие для (2.8) на поверхности сферы $|\mathbf{x}| = r_0$ с учетом (2.7) имеет вид

$$(\nabla \varphi) \mathbf{n} = \mathbf{v} \mathbf{n} = v_0(t) \left(\frac{r_0^2 \mathbf{x}}{|\mathbf{x}|^3} \frac{\mathbf{x}}{|\mathbf{x}|} \right) \Big|_{|\mathbf{x}|=r_0} = v_0(t), \quad (2.9)$$

где $\mathbf{n} = \mathbf{x}/|\mathbf{x}|$ — единичный вектор внешней нормали к сфере. Таким образом, (2.8) и (2.9) в сочетании с условием ограниченности φ при $|\mathbf{x}| \rightarrow \infty$ составляют вытекающую из закона сохранения массы математическую формулировку краевой задачи для нахождения потенциала скорости. Несложно проверить, что задаче (2.8), (2.9) удовлетворяет функция $\varphi = -v_0(t)r_0^2/|\mathbf{x}| + C_0$, где $C_0 = \text{const}$.

Если несжимаемая среда неоднородна, то зависимость ее плотности $\rho(t, \mathbf{x})$ от времени t в точке пространства с радиус-вектором $\mathbf{x}$ описывается уравнением неразрывности (2.2), которое можно записать в виде $\frac{\partial \rho}{\partial t} = -\nabla(\rho \nabla \varphi)$. При установившемся движении такой среды, т.е. при $\frac{\partial \rho}{\partial t} = 0$, векторное поле плотности $\rho \nabla \varphi = \rho \mathbf{v}$ потока массы является соленоидальным. Если же поле скорости потенциально, но среда сжимаема, то потенциал φ поля скорости не удовлетворяет (2.8), а (2.2) и (2.3) переходят в равенство

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \nabla \varphi) = \frac{d\rho}{dt} + \rho \nabla^2 \varphi = 0. \quad \#$$

Закон сохранения массы можно применить не только к однородной среде, но и к смеси n разнородных веществ, заполняющих некоторый объем V . Для каждого k -го вещества, $k = \overline{1, n}$, можно ввести плотность $\rho^{(k)}$, характеризующую его объемную концентрацию в смеси, а при помощи векторной функции $\mathbf{v}^{(k)} = \mathbf{v}^{(k)}(t, \mathbf{x})$ задать векторное поле его скорости. Тогда для плотности смеси получим

$$\rho = \sum_{k=1}^n \rho^{(k)}, \quad (2.10)$$

а из условия, что вектор $\rho \mathbf{v}$ плотности потока смеси при конвективном переносе равен сумме векторов $\rho^{(k)} \mathbf{v}^{(k)}$, $k = \overline{1, n}$, плотностей потоков отдельных веществ, найдем вектор средней скорости смеси

$$\mathbf{v} = \frac{1}{\rho} \sum_{k=1}^n \rho^{(k)} \mathbf{v}^{(k)}. \quad (2.11)$$

Пусть в смеси не происходит превращения веществ. Тогда для каждого k -го вещества справедлив закон сохранения массы в виде (2.1):

$$\frac{dm_k}{dt} = \int_{V(t)} \left(\frac{\partial \rho^{(k)}}{\partial t} + \nabla(\rho^{(k)} \mathbf{v}^{(k)}) \right) dV = 0. \quad (2.12)$$

Рассуждения, аналогичные проведенным при получении (2.2) и (2.3), в случае непрерывности в (2.12) подынтегрального выражения позволяют записать при $k = \overline{1, n}$

$$\frac{\partial \rho^{(k)}}{\partial t} + \nabla(\rho^{(k)} \mathbf{v}^{(k)}) = 0, \quad \text{или} \quad \frac{d\rho^{(k)}}{dt} + \rho^{(k)} \nabla \mathbf{v}^{(k)} = 0. \quad (2.13)$$

Ясно, что, суммируя по k первые равенства в (2.13) и учитывая (2.10) и (2.11), получаем уравнение неразрывности смеси в виде (2.2).

Если же в смеси происходит превращение веществ за счет химических реакций или ионизации, то для каждого k -го вещества этот процесс характеризуется скоростью $\dot{m}_V^{(k)}$ изменения массы этого вещества в единицу времени в единице объема, причем из условия сохранения массы смеси следует, что

$$\sum_{k=1}^n \dot{m}_V^{(k)} = 0. \quad (2.14)$$

В этом случае вместо (2.12) получим выражение, соответствующее закону сохранения массы k -го вещества:

$$\frac{dm_k}{dt} = \int_{V(t)} \left(\frac{\partial \rho^{(k)}}{\partial t} + \nabla(\rho^{(k)} \mathbf{v}^{(k)}) \right) dV = \int_{V(t)} \dot{m}_V^{(k)} dV.$$

Отсюда приходим к локальной форме закона сохранения массы k -го вещества:

$$\frac{\partial \rho^{(k)}}{\partial t} + \nabla(\rho^{(k)} \mathbf{v}^{(k)}) = \dot{m}_V^{(k)}, \quad k = \overline{1, n}. \quad (2.15)$$

Суммирование (2.15) по k с учетом (2.10), (2.11) и (2.14) снова приведет к (2.2).

Величина $\rho^{(k)} \mathbf{v}$ является вектором плотности потока k -го вещества при конвективном переносе, определяемом движением смеси в целом, а величину $\mathbf{J}^{(k)} = \rho^{(k)}(\mathbf{v}^{(k)} - \mathbf{v})$ можно рассматривать как вектор плотности потока этого вещества при диффузионном переносе, вызванном отличием скорости k -го вещества от средней скорости смеси. Тогда (2.15) можно представить в виде

$$\frac{\partial \rho^{(k)}}{\partial t} + \nabla \rho^{(k)} \mathbf{v} = \dot{m}_V^{(k)} - \nabla \mathbf{J}^{(k)}, \quad k = \overline{1, n}. \quad (2.16)$$

Принимая во внимание (2.10) и (2.11), нетрудно установить, что сумма по k всех векторов $\mathbf{J}^{(k)}$ равна нулевому вектору $\mathbf{0}$.

Скорости отдельных веществ в смеси обычно неизвестны. Но для описания диффузионного переноса k -го вещества в

смеси можно использовать эмпирический закон Фика (см. 1.3). Обозначим объемную концентрацию k -го вещества через C , т.е. $C = \rho^{(k)}$. Тогда с учетом (1.11) вместо (2.16) получим

$$\frac{\partial C}{\partial t} + \nabla(Cv) = \nabla(K^{(C)} \nabla C) + I_V^{(C)}. \quad (2.17)$$

Здесь $K^{(C)}$ — коэффициент диффузии данного вещества в смеси, который в общем случае может зависеть от C , времени и пространственных координат, а $I_V^{(C)} = \dot{m}_V^{(k)}$. Если средняя скорость смеси равна нулю, т.е. $v = 0$, а $K^{(C)}$ не зависит от C и пространственных координат, то из (2.17) следует, что

$$\frac{\partial C}{\partial t} = K^{(C)} \nabla^2 C + I_V^{(C)}. \quad (2.18)$$

Для нахождения объемной плотности $C = C(t, \mathbf{x})$ в объеме V_0 , ограниченном неподвижной поверхностью S_0 , необходимо решить (2.18) при заданных *краевых условиях*. В эти условия должны войти функция $C_0(\mathbf{x}) = C(0, \mathbf{x})$, задающая в V распределение C в момент времени, принимаемый за начальный, т.е. начальные условия, и граничные условия на S .

Если $I_V^{(C)}$ не зависит от C или же зависит линейно, то (2.18) является линейным уравнением параболического типа. При $\frac{\partial C}{\partial t} \equiv 0$ и линейной зависимости $I_V^{(C)}$ от C из (2.18) следуют **уравнение Гельмгольца***, а если $\frac{\partial C}{\partial t} \equiv 0$ и $I_V^{(C)}$ не зависит от C , то — уравнение Пуассона**, относящиеся к уравнениям эллиптического типа. Наконец, при $\frac{\partial C}{\partial t} \equiv 0$ и $I_V^{(C)} \equiv 0$ имеем уравнение Лапласа. Подчеркнем, что все эти варианты уравнений следуют из локальной формы закона сохранения массы некоторого вещества в смеси.

*Г.Л.Ф. Гельмгольц (1821–1894) — немецкий физик, математик, физиолог и психолог.

**С.Д. Пуассон (1781–1840) — французский механик, физик и математик.

2.2. Дивергентная форма уравнения неразрывности

Пусть в каждой точке трехмерного пространства, определяемой радиус-вектором $\mathbf{x}$ в прямоугольной системе координат $Ox_1x_2x_3$ (см. рис. 1.1), плотность $\rho(\mathbf{x})$ среды и вектор $\mathbf{v}(\mathbf{x})$ ее скорости не зависят от времени t , т.е. процесс переноса массы является установившимся. Тогда $\frac{\partial \rho}{\partial t} \equiv 0$ и из (2.2) получим так называемую **дивергентную форму**

$$\nabla(\rho \mathbf{v}) = 0 \quad (2.19)$$

уравнения неразрывности. Используя выражение (1.9) для дифференциального оператора Гамильтона и представление $\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + v_3 \mathbf{e}_3$ вектора скорости среды в координатной форме, где v_i — проекция вектора $\mathbf{v}$ на ось Ox_i , $i = 1, 2, 3$, а $\mathbf{e}_i$ — единичный вектор, задающий направление этой оси, запишем (2.19) в координатной форме:

$$\frac{\partial(\rho v_1)}{\partial x_1} + \frac{\partial(\rho v_2)}{\partial x_2} + \frac{\partial(\rho v_3)}{\partial x_3} = 0. \quad (2.20)$$

Рассмотрим произвольный объем $V \in \mathbb{R}^3$, ограниченный поверхностью S (рис. 2.1) и заполненный сплошной средой. В окрестности каждой точки этого объема справедливо равенство (2.19). Интегрируя это равенство по объему V и используя

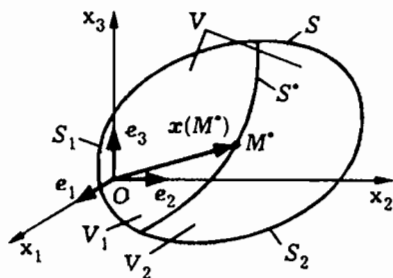


Рис. 2.1

формулу Остроградского — Гаусса, запишем

$$\int_S \rho \mathbf{v} \mathbf{n} dS = \int_V \nabla(\rho \mathbf{v}) dV = 0, \quad (2.21)$$

где $\mathbf{n}$ — единичный вектор внешней нормали к поверхности S .

Если в области V существуют такие точки, в которых функции $\rho(\mathbf{x})$ и $\mathbf{v}(\mathbf{x})$ не имеют непрерывных производных по пространственным координатам или не являются непрерывными, то в этих точках (2.19) не имеет смысла, а формула Остроградского — Гаусса не применима к области V в целом и первое равенство в (2.21) в общем случае теряет силу. Вместе с тем следует отметить, что если множество таких точек в $V \in \mathbb{R}^3$ имеет меру Лебега, равную нулю, в частности образует в V линию или поверхность, то значение интеграла по V в (2.21), равное нулю, остается неизменным.

Пусть поверхность S^* , в точках которой функция $\rho \mathbf{v}$ не является непрерывной и которую называют **поверхностью разрыва**, делит область V на две подобласти V_1 и V_2 так, что $V_1 \cup V_2 \cup S^* = V$ и $V_1 \cap V_2 \cap S^* = \emptyset$ (см. рис. 2.1). Выясним, при каких условиях на поверхности S^* сохраняет силу (2.21). Обозначим через S_1 и S_2 части поверхности $S = S_1 \cup S_2$, ограничивающие вместе с S^* объемы V_1 и V_2 соответственно. Предположим, что в V_1 и V_2 функция $\rho \mathbf{v}$, которую обозначим $\rho_1 \mathbf{v}_1$ и $\rho_2 \mathbf{v}_2$ соответственно, непрерывно дифференцируема по пространственным координатам. Тогда к каждой из областей V_1 и V_2 можно применить формулу Остроградского — Гаусса и написать

$$\int_{S_1} \rho_1 \mathbf{v}_1 \mathbf{n} dS + \int_{S^*} \rho_1 \mathbf{v}_1 \mathbf{n}^* dS = \int_{V_1} \nabla(\rho_1 \mathbf{v}_1) dV = 0,$$

$$\int_{S_2} \rho_2 \mathbf{v}_2 \mathbf{n} dS + \int_{S^*} \rho_2 \mathbf{v}_2 (-\mathbf{n}^*) dS = \int_{V_2} \nabla(\rho_2 \mathbf{v}_2) dV = 0,$$

где $\mathbf{n}^*$ — единичный вектор нормали к поверхности S^* , направленный в сторону области V_2 . Складывая почленно эти равенства и учитывая, что S^* является в V множеством точек объема нуль, получаем

$$\begin{aligned} \int_{S_1} \rho_1 \mathbf{v}_1 \mathbf{n} dS + \int_{S_2} \rho_2 \mathbf{v}_2 \mathbf{n} dS + \int_{S^*} \rho_1 \mathbf{v}_1 \mathbf{n}^* dS + \int_{S^*} \rho_2 \mathbf{v}_2 (-\mathbf{n}^*) dS = \\ = \int_S \rho \mathbf{v} \mathbf{n} dS + \int_{S^*} (\rho_1 \mathbf{v}_1 - \rho_2 \mathbf{v}_2) \mathbf{n}^* dS = \\ = \int_{V_1} \nabla(\rho_1 \mathbf{v}_1) dV + \int_{V_2} \nabla(\rho_2 \mathbf{v}_2) dV = \int_V \nabla(\rho \mathbf{v}) dV = 0. \end{aligned}$$

Следовательно, (2.21) как интегральная форма закона сохранения массы остается в силе при условии

$$\int_{S^*} (\rho_1 \mathbf{v}_1 - \rho_2 \mathbf{v}_2) \mathbf{n}^* dS = 0. \quad (2.22)$$

Пусть V является окрестностью некоторой точки $M^* \in S^*$ с радиус-вектором $\mathbf{x}(M^*)$ (см. рис. 2.1). Стягивая V к точке M^* , из (2.22) получаем, что

$$\rho_1^* \mathbf{v}_1^* \mathbf{n}^* = \rho_2^* \mathbf{v}_2^* \mathbf{n}^* \text{ на } S^*, \quad (2.23)$$

где ρ_1^* , $\mathbf{v}_1^*$ и ρ_2^* , $\mathbf{v}_2^*$ — пределы функций $\rho(\mathbf{x})$ и $\mathbf{v}(\mathbf{x})$ при стремлении к точке $M^* \in S^*$ точек $M_1 \in V_1$ и $M_2 \in V_2$ соответственно. Следовательно, (2.23) устанавливает условия в точках на поверхности разрыва, не противоречащие закону сохранения массы.

Таким образом, векторная функция $\rho \mathbf{v}$ может не быть непрерывной в точках поверхности S^* , но должна сохранять непрерывность проекции на нормаль $\mathbf{n}^*$ при переходе через эту поверхность. При этом допустим разрыв проекции этой функции на плоскость, касательную к S^* в точках разрыва.

Отметим, что это свойство любого соленоидального векторного поля, дивергенция которого равна нулю. В частности, в точках поверхности, ограничивающей *векторную трубку* в таком поле, допустим разрыв векторной функции, но ее проекция на нормаль к этой поверхности равна нулю.

Условие вида (2.23) позволяет строить так называемые негладкие и разрывные решения задач математической физики, не удовлетворяющие (2.19) во всех точках рассматриваемой области, путем „сшивания“ гладких решений, полученных в подобластях, где (2.19) сохраняет силу. Такие задачи возникают при изучении процессов переноса массы, например, в неоднородной несжимаемой среде с резким или скачкообразным изменением плотности, в сжимаемой среде со скачками уплотнения и ударными волнами.

Эти соображения вызывают естественное стремление попытаться использовать дивергентную форму уравнения неразрывности для общего случая неустановившегося во времени движения сплошной среды. В этом случае функции ρ и v в уравнении неразрывности (2.2) зависят не только от пространственных координат, но и от времени t .

Аналогично (2.20) запишем (2.2) в координатной форме:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v_1)}{\partial x_1} + \frac{\partial(\rho v_2)}{\partial x_2} + \frac{\partial(\rho v_3)}{\partial x_3} = 0. \quad (2.24)$$

Наличие в левой части (2.24) четырех частных производных позволяет придать ей компактную форму, которую удобно использовать в дальнейшем при рассмотрении ряда вопросов.

В четырехмерном евклидовом арифметическом пространстве $\mathbb{R}^4$ с ортонормированным базисом $e = (e_t, e_1, e_2, e_3)$ определим вектор

$$\tilde{r} = \rho(v_0 e_t + v_1 e_1 + v_2 e_2 + v_3 e_3) = \rho(v_0 e_t + v), \quad (2.25)$$

где v_0 — произвольная константа, имеющая размерность скорости (эту константу можно принять равной 1 м/с), и введем

дифференциальную операцию дивергенции при помощи оператора

$$\tilde{\nabla} = \frac{1}{v_0} \mathbf{e}_t \frac{\partial}{\partial t} + \mathbf{e}_1 \frac{\partial}{\partial x_1} + \mathbf{e}_2 \frac{\partial}{\partial x_2} + \mathbf{e}_3 \frac{\partial}{\partial x_3} = \frac{1}{v_0} \mathbf{e}_t \frac{\partial}{\partial t} + \nabla, \quad (2.26)$$

который аналогичен дифференциальному векторному оператору (1.19) (см. Д.1.1). В соответствии с правилом *стандартного скалярного умножения*, учитывая (2.25) и (2.26), вычислим

$$\tilde{\nabla} \tilde{\mathbf{r}} = \frac{\partial \rho}{\partial t} + \frac{\partial \rho v_1}{\partial x_1} + \frac{\partial \rho v_2}{\partial x_2} + \frac{\partial \rho v_3}{\partial x_3}.$$

Сравнивая правую часть этого равенства с (2.24), приходим к выводу, что уравнение неразрывности в пространстве $\mathbb{R}^4$ можно записать в виде

$$\tilde{\nabla} \tilde{\mathbf{r}} = 0. \quad (2.27)$$

Такую *форму* представления *уравнения неразрывности* будем называть *обобщенной дивергентной*.

Пусть V — произвольный изменяющийся во времени объем, ограниченный движущейся поверхностью S (см. рис. 1.1), точки которой перемещаются со скоростью $\mathbf{v}(t, \mathbf{x})$ среды. В таком объеме масса m среды будет неизменной во времени (см. 2.1). Обозначим через $\Theta \subset [t_0, t_1)$ промежуток времени t , в течение которого происходит движение среды, и рассмотрим область $\Omega = \Theta \times V \subset \mathbb{R}^4$ с границей $\partial\Omega$ (рис. 2.2). Интегрируя (2.27) по области Ω и применяя в $\mathbb{R}^4$ *формулу* (1.26) *Остроградского* (см. Д.1.1), получаем

$$\int_{\partial\Omega} \tilde{\mathbf{r}} \tilde{\mathbf{n}} d(\partial\Omega) = \int_{\Omega} \tilde{\nabla} \tilde{\mathbf{r}} d\Omega = 0, \quad (2.28)$$

где $\tilde{\mathbf{n}}$ — единичный вектор внешней нормали к $\partial\Omega$ (в данном случае $d\Omega = dt dV$ и $d(\partial\Omega) = dt dS$).

Пусть в области Ω существуют точки $\tilde{M}(t; x_1; x_2; x_3) \in \Omega$, в которых функция $\tilde{\mathbf{r}}(\tilde{M})$ не является непрерывной или же

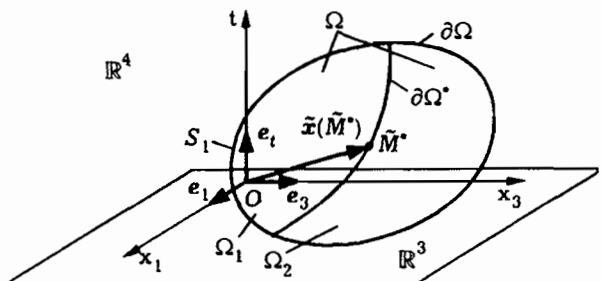


Рис. 2.2

не имеет непрерывных частных производных по координатам этих точек. Тогда интеграл по области Ω в (2.28) аналогично интегралу по трехмерной области V в (2.21) сохранил свое значение, равное нулю, если такие точки образуют множество $\Omega^* \subset \Omega$ с мерой Лебега, равной нулю. Но при этом первое равенство в (2.28) может не иметь места. Пусть такое множество $\partial\Omega^* \subset \Omega$ является в $\mathbb{R}^4$ гладкой поверхностью [V] и делит область Ω на две подобласти Ω_1 и Ω_2 так, что $\Omega_1 \cup \Omega_2 \cup \partial\Omega^* = \Omega$ и $\Omega_1 \cap \Omega_2 \cap \partial\Omega^* = \emptyset$.

Множество $\partial\Omega^* \subset \mathbb{R}^4$ зададим уравнением

$$\varphi(t, \mathbf{x}) = 0, \quad (2.29)$$

где $\varphi(t, \mathbf{x})$ — действительная функция, непрерывно дифференцируемая по времени t и пространственным координатам x_i , $i = 1, 2, 3$, радиус-вектора $\mathbf{x}$ точки $M \in \mathbb{R}^3$, причем частные производные этой функции по пространственным координатам не обращаются одновременно в нуль в любой точке $\tilde{M}^*(t; x_1; x_2; x_3) \in \partial\Omega^*$, т.е. $|\nabla\varphi| \neq 0$. Тогда по аналогии с выражением для вектора нормали к поверхности в $\mathbb{R}^3$ [V] получим с учетом (2.26) выражение

$$\tilde{\mathbf{n}}^* = \frac{1}{v_0} \frac{\partial\varphi}{\partial t} \mathbf{e}_t + \frac{\partial\varphi}{\partial x_1} \mathbf{e}_1 + \frac{\partial\varphi}{\partial x_2} \mathbf{e}_2 + \frac{\partial\varphi}{\partial x_3} \mathbf{e}_3 = \tilde{\nabla}\varphi, \quad (2.30)$$

определяющее в $\mathbb{R}^4$ вектор, который можно рассматривать как вектор нормали к $\partial\Omega^*$ в произвольной точке $\tilde{M}^* \in \partial\Omega^*$.

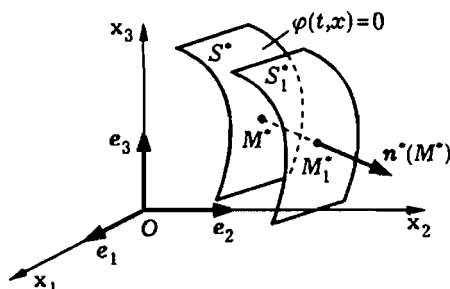


Рис. 2.3

Если через $\tilde{r}_1^*$ и $\tilde{r}_2^*$ обозначить пределы функции $\tilde{r}(\tilde{M})$ при стремлении к точке $\tilde{M}^* \in \partial\Omega^*$ точек $\tilde{M}_1 \in \Omega_1$ и $\tilde{M}_2 \in \Omega_2$ соответственно, то, используя ту же процедуру, что и в случае установившегося переноса массы, в произвольной точке $\tilde{M}^* \in \partial\Omega^*$ приходим к условию

$$\tilde{r}_1^* \tilde{n}^* = \tilde{r}_2^* \tilde{n}^* \quad \text{на } \partial\Omega^*, \quad (2.31)$$

при выполнении которого сохраняет силу первое равенство в (2.28).

Теперь в трехмерном пространстве $\mathbb{R}^3$ с системой координат $Ox_1x_2x_3$ рассмотрим подвижную поверхность, заданную тем же уравнением (2.29). Пусть в момент времени t эта поверхность занимает положение S^* и за время Δt переходит в положение S_1^* (рис. 2.3). Вектор скорости $v^*(t, M^*)$ перемещения поверхности S^* в точке $M^* \in S^*$ определим как предел

$$v^*(t, M^*) = \lim_{\Delta t \rightarrow 0} \frac{x(M_1^*) - x(M^*)}{\Delta t}, \quad (2.32)$$

где $M_1^* \in S_1^*$ — точка пересечения S_1^* с прямой, лежащей на направлении единичного вектора

$$n^*(M^*) = \left. \frac{\nabla \varphi(t, x)}{|\nabla \varphi(t, x)|} \right|_{x=x(M^*)} \quad (2.33)$$

нормали к S^* в точке $M^* \in S^*$, а $\mathbf{x}(M_1^*)$ и $\mathbf{x}(M^*)$ — радиус-векторы точек M_1^* и M^* соответственно. Так как $\varphi(t, \mathbf{x}(M^*)) = 0$ и $\varphi(t + \Delta t, \mathbf{x}(M_1^*)) = 0$, то в линейном приближении с учетом (2.33) получаем

$$\begin{aligned} \varphi(t + \Delta t, \mathbf{x}(M_1^*)) - \varphi(t, \mathbf{x}(M^*)) &= \left. \frac{\partial \varphi(t, \mathbf{x})}{\partial t} \right|_{\mathbf{x}=\mathbf{x}(M^*)} \Delta t + \\ &+ \left. \nabla \varphi(t, \mathbf{x}) \right|_{\mathbf{x}=\mathbf{x}(M^*)} (\mathbf{x}(M_1^*) - \mathbf{x}(M^*)) = \left. \frac{\partial \varphi(t, \mathbf{x})}{\partial t} \right|_{\mathbf{x}=\mathbf{x}(M^*)} \Delta t + \\ &+ \left. \left| \nabla \varphi(t, \mathbf{x}) \right| \right|_{\mathbf{x}=\mathbf{x}(M^*)} \mathbf{n}^*(M^*) (\mathbf{x}(M_1^*) - \mathbf{x}(M^*)) = 0. \end{aligned}$$

Отсюда в силу того, что векторы $\mathbf{x}(M_1^*) - \mathbf{x}(M^*)$ и $\mathbf{n}^*(M^*)$ являются коллинеарными, имеем

$$(\mathbf{x}(M_1^*) - \mathbf{x}(M^*)) = - \frac{1}{\left| \nabla \varphi(t, \mathbf{x}) \right|} \left. \frac{\partial \varphi(t, \mathbf{x})}{\partial t} \right|_{\mathbf{x}=\mathbf{x}(M^*)} \mathbf{n}^*(M^*) \Delta t$$

и после подстановки в (2.32) находим

$$\mathbf{v}^*(t, M^*) = - \frac{\partial \varphi(t, \mathbf{x})}{\partial t} \left. \frac{1}{\left| \nabla \varphi(t, \mathbf{x}) \right|} \right|_{\mathbf{x}=\mathbf{x}(M^*)} \mathbf{n}^*(M^*). \quad (2.34)$$

Учитывая, что в (2.30) $\tilde{\nabla} \varphi = \frac{1}{v_0} \frac{\partial \varphi}{\partial t} \mathbf{e}_t + \nabla \varphi$, и используя (2.33) и (2.34), запишем

$$\tilde{\mathbf{n}}^*(\tilde{M}^*) = - \frac{\mathbf{v}^*(t, M^*)}{v_0} \mathbf{e}_t + \mathbf{n}^*(M^*). \quad (2.35)$$

Тогда вместо (2.31) с учетом (2.25) и обозначения $[\cdot]$ скачка значений функции при переходе через поверхность разрыва S^* в направлении, противоположном вектору $\tilde{\mathbf{n}}^*$, в любой точке

$M^* \in S^*$ имеем

$$\begin{aligned} [\rho(t, \mathbf{x}(M^*)) \mathbf{v}(t, \mathbf{x}(M^*)) \mathbf{n}^*(M^*)] = \\ = [\rho(t, \mathbf{x}(M^*))] \mathbf{v}^*(t, M^*) \mathbf{n}^*(M^*). \end{aligned}$$

Отсюда следует, что если поверхность S^* неподвижна относительно системы координат $Ox_1x_2x_3$, т.е. функция φ не зависит от t и $|\mathbf{v}^*(M^*)| = 0$, то для любой точки $M^* \in S^*$ верно условие (2.23). Если же для каждой точки $M^* \in S^*$ в фиксированный момент времени t ввести свою *сопутствующую систему координат*, начало которой имеет скорость, определяемую вектором $\mathbf{v}^*(t, M^*)$, то в такой системе координат будем иметь

$$[\rho(t, \mathbf{x}(M^*)) \mathbf{w}(t, \mathbf{x}(M^*)) \mathbf{n}^*(M^*)] = 0, \quad M^* \in S^*, \quad (2.36)$$

где $\mathbf{w} = \mathbf{v} - \mathbf{v}^*$ — вектор скорости движения среды относительно точки $M^* \in S^*$.

Пример 2.2. Пусть среда разнородна и состоит из смеси n веществ, плотность и скорость которых описывают функции $\rho^{(k)}(t, \mathbf{x})$ и $\mathbf{v}^{(k)}(t, \mathbf{x})$, $k = \overline{1, n}$, непрерывно дифференцируемые в области $\Omega \subset \mathbb{R}^4$ по всем своим аргументам. Тогда для каждого k -го вещества вместо (2.15) получим обобщенную дивергентную форму закона сохранения массы в виде

$$\tilde{\nabla} \tilde{\mathbf{r}}^{(k)} = \dot{m}_V^{(k)}, \quad (2.37)$$

где $\dot{m}_V^{(k)}(t, \mathbf{x})$ — непрерывная в области $\Omega \subset \mathbb{R}^4$ функция, характеризующая скорость образования этого вещества в единице объема за счет превращения из других веществ, а $\tilde{\mathbf{r}}^{(k)}$ отличается от $\tilde{\mathbf{r}}$ в (2.25) лишь верхним индексом k у функций ρ и $\mathbf{v}$. Интегрируя (2.37) по области Ω и используя формулу Остроградского в виде (1.26), получаем

$$\int_{\partial\Omega} \tilde{\mathbf{r}}^{(k)} \tilde{\mathbf{n}} d(\partial\Omega) = \int_{\Omega} \tilde{\nabla} \tilde{\mathbf{r}}^{(k)} d\Omega = \int_{\Omega} \dot{m}_V^{(k)} d\Omega, \quad (2.38)$$

где, по-прежнему, $\partial\Omega$ — граница области Ω , а $\tilde{n}$ — единичный вектор внешней нормали к $\partial\Omega$.

Пусть теперь множество $\partial\Omega^* \subset \Omega$ точек, в которых функция $\tilde{r}^{(k)}$ не имеет непрерывных производных или разрывна, делит область Ω на две подобласти Ω_1 и $\Omega_2 = (\Omega \setminus \partial\Omega^*) \setminus \Omega_1$, причем на $\partial\Omega^*$ может происходить образование k -го вещества за счет превращения из других веществ со скоростью $\dot{m}_S^{(k)}$. Если $\partial\Omega^*$ является в Ω множеством, мера Лебега которого равна нулю, то интеграл по Ω в (2.38) не изменит своего значения и второе равенство в (2.38) сохранит силу, но первое равенство в общем случае не будет иметь места. Действительно, из условия сохранения k -го вещества в области Ω теперь следует, что

$$\int_{\partial\Omega} \tilde{r}^{(k)} \tilde{n} d(\partial\Omega) = \int_{\Omega} \dot{m}_V^{(k)} d\Omega + \int_{\partial\Omega^*} \dot{m}_S^{(k)} d(\partial\Omega^*). \quad (2.39)$$

Обозначим через $\partial\Omega_1$ и $\partial\Omega_2 = \partial\Omega \setminus \partial\Omega_1$ части границы $\partial\Omega$ области Ω , ограничивающие извне подобласти Ω_1 и Ω_2 соответственно, и предположим, что в Ω_1 и Ω_2 функция $\tilde{r}^{(k)}$, которую обозначим $r_1^{(k)}$ и $r_2^{(k)}$ соответственно, непрерывно дифференцируема по всем своим аргументам. Тогда к каждой из этих подобластей можно применить формулу Остроградского (1.26) и написать

$$\begin{aligned} \int_{\partial\Omega_1} \tilde{r}_1^{(k)} \tilde{n} d(\partial\Omega) + \int_{\partial\Omega^*} \tilde{r}_1^{(k)} \tilde{n}^* d(\partial\Omega^*) &= \int_{\Omega_1} \tilde{\nabla} \tilde{r}_1^{(k)} d\Omega, \\ \int_{\partial\Omega_2} \tilde{r}_2^{(k)} \tilde{n} d(\partial\Omega) + \int_{\partial\Omega^*} \tilde{r}_2^{(k)} (-\tilde{n}^*) d(\partial\Omega^*) &= \int_{\Omega_2} \tilde{\nabla} \tilde{r}_2^{(k)} d\Omega, \end{aligned}$$

где $\tilde{n}^*$ — единичный вектор нормали к $\partial\Omega^*$, направленный в сторону подобласти Ω_2 . Складывая почленно эти равенства и считая $\partial\Omega^* \subset \Omega$ множеством, имеющим равную нулю меру

Лебега, с учетом второго равенства в (2.38) получаем

$$\begin{aligned}
 & \int_{\partial\Omega_1} \tilde{\mathbf{r}}_1^{(k)} \tilde{\mathbf{n}} d(\partial\Omega) + \int_{\partial\Omega_2} \tilde{\mathbf{r}}_2^{(k)} \tilde{\mathbf{n}} d(\partial\Omega) + \\
 & + \int_{\partial\Omega^*} \tilde{\mathbf{r}}_1^{(k)} \tilde{\mathbf{n}}^* d(\partial\Omega^*) + \int_{\partial\Omega^*} \tilde{\mathbf{r}}_2^{(k)} (-\tilde{\mathbf{n}}^*) d(\partial\Omega^*) = \\
 & = \int_{\partial\Omega} \tilde{\mathbf{r}}^{(k)} \tilde{\mathbf{n}} d(\partial\Omega) + \int_{\partial\Omega^*} (\tilde{\mathbf{r}}_1^{(k)} - \tilde{\mathbf{r}}_2^{(k)}) \tilde{\mathbf{n}}^* d(\partial\Omega^*) = \\
 & = \int_{\Omega_1} \tilde{\nabla} \tilde{\mathbf{r}}_1^{(k)} d\Omega + \int_{\Omega_2} \tilde{\nabla} \tilde{\mathbf{r}}_2^{(k)} d\Omega = \int_{\Omega} \tilde{\nabla} \tilde{\mathbf{r}}^{(k)} d\Omega = \int_{\Omega} \dot{m}_V^{(k)} d\Omega.
 \end{aligned}$$

Отсюда с учетом (2.39) следует, что второе равенство в (2.38) сохраняет силу при выполнении условия

$$\int_{\partial\Omega^*} (\tilde{\mathbf{r}}_1^{(k)} - \tilde{\mathbf{r}}_2^{(k)}) \tilde{\mathbf{n}}^* d(\partial\Omega^*) + \int_{\partial\Omega^*} \dot{m}_S^{(k)} d(\partial\Omega^*) = 0. \quad (2.40)$$

Рассматривая область Ω как окрестность некоторой точки $\tilde{M}^* \in \partial\Omega^*$ и стягивая эту окрестность к данной точке, из (2.40) заключаем, что

$$(\tilde{\mathbf{r}}_1^* - \tilde{\mathbf{r}}_2^*) \tilde{\mathbf{n}}^* + \dot{m}_S^{(k)} = 0 \quad \text{на } S^*, \quad (2.41)$$

где $\tilde{\mathbf{r}}_1^*$ и $\tilde{\mathbf{r}}_2^*$ — пределы функций $\tilde{\mathbf{r}}_1^{(k)}$ и $\tilde{\mathbf{r}}_2^{(k)}$ при стремлении к точке $\tilde{M}^* \in \partial\Omega^*$ точек $\tilde{M}_1 \in \Omega_1$ и $\tilde{M}_2 \in \Omega_2$ соответственно. Если множество $\partial\Omega^*$ задано в $\mathbb{R}^4$ уравнением (2.29), то точке $\tilde{M}^* \in \partial\Omega^*$ в момент времени t соответствует точка $M^* \in S^*$ на поверхности разрыва S^* , заданной тем же уравнением в $\mathbb{R}^3$ (см. рис. 2.1). Тогда с учетом (2.35) и обозначения скачка функций при переходе через поверхность разрыва S^* в любой точке $M^* \in S^*$ получаем

$$\begin{aligned}
 & [\rho^{(k)}(t, \mathbf{x}(M^*)) \mathbf{v}^{(k)}(t, \mathbf{x}(M^*)) \mathbf{n}^*(M^*)] + \dot{m}_S^{(k)}(t, \mathbf{x}(M^*)) = \\
 & = [\rho^{(k)}(t, \mathbf{x}(M^*))] \mathbf{v}^*(t, M^*) \mathbf{n}^*(M^*). \quad (2.42)
 \end{aligned}$$

Для записанного в виде (2.17) закона сохранения k -го вещества, имеющего в движущейся со скоростью v среде объемную концентрацию $C = \rho^{(k)}$, вместо (2.37) будем иметь

$$\tilde{\nabla} \tilde{C} = I_V^{(C)}, \quad (2.43)$$

где $C = C v_0 e_t + C v - K^{(C)} \nabla C$ ($K^{(C)}$ — коэффициент диффузии); $I_V^{(C)} = \dot{m}_V^{(k)}$. В этом случае вместо (2.42) получаем

$$\begin{aligned} & [(C(t, x(M^*))v(t, x(M^*)) - K^{(C)} \nabla C(t, x(M^*)))n^*(M^*)] + \\ & + I_S^{(C)}(t, x(M^*)) = [C(t, x(M^*))v^*(t, M^*)n^*(M^*), \end{aligned} \quad (2.44)$$

где $I_S^{(C)} = \dot{m}_S^{(k)}$. Отсюда в сопутствующей для точки $M^* \in S^*$ системе координат приходим к условию

$$\begin{aligned} & [C(t, x(M^*))w(t, x(M^*))n^*(M^*)] + I_S^{(C)}(t, x(M^*)) = \\ & = [K^{(C)} \nabla C(t, x(M^*))n^*(M^*)]. \end{aligned}$$

2.3. Законы сохранения электрического заряда и тепловой энергии

Пусть в области $V_0 \subset \mathbb{R}^3$, ограниченной поверхностью S_0 , неподвижной относительно прямоугольной системы координат $Ox_1x_2x_3$ (рис. 2.4), существуют источники электрического заряда интенсивностью $I_V^{(e)}$, равной изменению его объемной плотности ρ_e в единицу времени. Через $j_n^{(e)}$ обозначим проекцию вектора $j^{(e)}$ плотности электрического тока в точках поверхности S_0 на направление в этих точках единичного вектора n внешней нормали к S_0 . Тогда, учитывая (1.4), условие сохранения заряда в объеме V_0 можно представить в виде

$$\frac{dQ}{dt} = \frac{d}{dt} \int_{V_0} \rho_e dV = \int_{V_0} I_V^{(e)} dV - \int_{S_0} j_n^{(e)} dS.$$

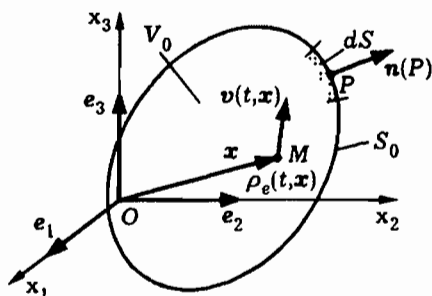


Рис. 2.4

Отсюда, используя (1.10) и формулу Остроградского — Гаусса, получаем интегральную форму закона сохранения электрического заряда:

$$\int_{V_0} \left(\frac{\partial \rho_e}{\partial t} + \nabla \cdot \mathbf{j}^{(e)} - I_V^{(e)} \right) dV = 0. \quad (2.45)$$

Так как объем V_0 произволен, то при условии непрерывности подынтегральной функции из (2.45) находим локальную форму этого закона:

$$\frac{\partial \rho_e}{\partial t} + \nabla \cdot \mathbf{j}^{(e)} = I_V^{(e)}. \quad (2.46)$$

Выделив конвективную и диффузионную составляющие переноса электрических зарядов (см. 1.3), запишем $\mathbf{j}^{(e)} = \rho_e \mathbf{v} + \sigma \mathbf{E}$, где $\mathbf{v}$ и σ — вектор скорости и электрическая проводимость среды; $\mathbf{E}$ — напряженность электрического поля. Тогда вместо (2.46) получим

$$\frac{\partial \rho_e}{\partial t} + \nabla \cdot (\rho_e \mathbf{v}) = I_V^{(e)} - \nabla \cdot (\sigma \mathbf{E}). \quad (2.47)$$

В отличие от (2.18), куда входит лишь одна неизвестная функция C , в (2.47) помимо неизвестной функции ρ_e входит еще и напряженность $\mathbf{E}$, которая, вообще говоря, зависит от ρ_e

(см. 3.5). Поэтому даже при заданных функциях $I_V^{(e)}$ и v , как правило, формулы (2.47) недостаточно для получения полной формулировки задачи.

В четырехмерном евклидовом арифметическом пространстве $\mathbb{R}^4$ с ортонормированным базисом $e = (e_t, e_1, e_2, e_3)$ определим вектор $\tilde{r}_e = \rho_e v_0 e_t + \rho_e v + \sigma E$, где, как и в (2.25), v_0 — произвольная константа, имеющая размерность скорости. Тогда *обобщенная дивергентная форма* закона сохранения заряда примет вид $\tilde{\nabla} \tilde{r}_e = I_V^{(e)}$, причем дифференциальная операция дивергенции задана в $\mathbb{R}^4$ при помощи оператора (2.26). Пусть векторная функция $\tilde{r}_e(t, M)$ не имеет непрерывных производных или не является непрерывной в точках $M^* \in S^*$ некоторой поверхности S^* , делящей область V_0 на подобласти V_1 и V_2 . Тогда по аналогии с (2.44) в любой точке $M^* \in S^*$ получим для этой поверхности разрыва условие

$$[(\rho_e(t, M^*)v(t, M^*) + \sigma E(t, M^*))n^*(M^*)] + \\ + I_S^{(e)}(t, M^*) = [\rho_e(t, M^*)]v^*(t, M^*)n^*(M^*). \quad (2.48)$$

Здесь $I_S^{(e)}$ — поверхностная интенсивность источников заряда, равная изменению заряда в единицу времени на единице площади поверхности S^* , $v^*(t, M^*)$ — вектор скорости перемещения точки $M^* \in S^*$ вместе с поверхностью S^* , а символ $[\cdot]$ обозначает скачок значений функции при переходе в точке $M^* \in S^*$ через поверхность S^* в направлении, противоположном единичному вектору $n^*(M^*)$ к этой поверхности.

Условие (2.48) позволяет „сшивать“ на поверхностях разрыва так называемые гладкие решения, полученные в отдельных подобластях и удовлетворяющие закону сохранения заряда (2.47). Необходимость в этом возникает, например, при переносе заряда в неоднородной среде и при возникновении источников заряда на границах контакта разнородных материалов.

Перейдем к рассмотрению закона сохранения тепловой энергии в неподвижной несжимаемой среде, занимающей неизменяе-

мый во времени объем V_0 , ограниченный поверхностью S_0 , фиксированной относительно системы координат $Ox_1x_2x_3$. Пусть в среде действуют источники тепловой энергии объемной мощностью $I_V^{(q)}$. Через q_n обозначим проекцию вектора $\mathbf{q}$ плотности теплового потока в точках поверхности S_0 на направление в этих точках единичного вектора $\mathbf{n}$ внешней нормали к S_0 . Тогда условие сохранения тепловой энергии в объеме V_0 примет вид

$$\frac{d}{dt} \int_{V_0} \varepsilon_T dV = \int_{V_0} I_V^{(q)} dV - \int_{S_0} q_n dS. \quad (2.49)$$

Здесь ε_T — объемная плотность тепловой энергии.

Применяя (1.10) и формулу (1.18) Остроградского — Гаусса, из (2.49) получаем интегральную форму закона сохранения тепловой энергии в виде

$$\int_V \left(\frac{\partial \varepsilon_T}{\partial t} + \nabla \mathbf{q} - I_V^{(q)} \right) dV = 0. \quad (2.50)$$

В силу произвольности объема V_0 из (2.50) при условии непрерывности подынтегральной функции следует локальная форма этого закона:

$$\frac{\partial \varepsilon_T}{\partial t} + \nabla \mathbf{q} = I_V^{(q)}. \quad (2.51)$$

Примем, что ε_T в (2.51) зависит лишь от температуры T , причем

$$\varepsilon_T = \int_0^T c dT, \quad (2.52)$$

где $c > 0$ — теплоемкость единицы объема среды. Тогда, подставляя (1.12) и (2.52) в (2.51) и используя правило дифференцирования интеграла по верхнему пределу, получаем нелинейное (в общем случае) дифференциальное уравнение нестационарной

теплопроводности

$$\frac{\partial T}{\partial t} = \frac{1}{c} \nabla(\lambda \nabla T) + \frac{I_V^{(q)}}{c}. \quad (2.53)$$

Если теплопроводность среды λ не зависит от температуры и пространственных координат, то вместо (2.53) имеем аналогичное (2.18) уравнение

$$\frac{\partial T}{\partial t} = \tilde{a} \nabla^2 T + \frac{I_V^{(q)}}{c}, \quad (2.54)$$

где $\tilde{a} = \lambda/c$ — температуропроводность среды. Примем, что $\tilde{a}$ не зависит от температуры. Тогда если $I_V^{(q)}$ не зависит от T или зависит линейно, то (2.54) будет линейным уравнением параболического типа. Если $\frac{\partial T}{\partial t} \equiv 0$ и $I_V^{(q)}/c$ не зависит от T , то из (2.54) следует дифференциальное уравнение стационарной теплопроводности, являющееся уравнением Пуассона. Наконец, при $\frac{\partial T}{\partial t} \equiv 0$ и $I_V^{(q)} \equiv 0$ (2.54) переходит в уравнение Лапласа. Отметим, что все перечисленные типы уравнений получены из локальной формы закона сохранения тепловой энергии неподвижной несжимаемой среды.

В уравнение (2.53) входит единственная неизвестная функция $T = T(t, M)$, $M \in V_0$, описывающая изменение во времени t распределения температуры в области V_0 (рис. 2.5). Для решения уравнения (2.53) необходимо задать *краевые условия*. Начальным условием будет распределение в V_0 температуры

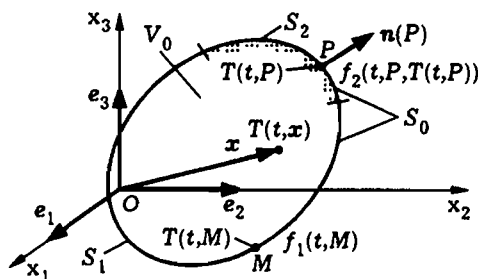


Рис. 2.5

$T_0(M) = T(0, M)$ в момент времени $t = 0$, принимаемый за начальный, а на ограничивающей V_0 неподвижной поверхности S_0 должны быть заданы граничные условия. Рассмотрим некоторые типы граничных условий.

Если в точках $M \in S_1$ на участках $S_1 \subset S_0$ поверхности S_0 в любой момент времени $t > 0$ функция $f_1(t, M)$ задает значения температуры этой поверхности (см. рис. 2.5), то говорят о граничных условиях I рода [XII] и записывают

$$T(t, M) = f_1(t, M), \quad M \in S_1. \quad (2.55)$$

На остальных участках $S_2 = S_0 \setminus S_1$ поверхности S_0 граничные условия в общем случае могут иметь вид

$$\lambda(\nabla T(t, P))\mathbf{n}(P) = f_2(t, P, T(t, P)), \quad P \in S_2, \quad (2.56)$$

где $\mathbf{n}(P)$ -- единичный вектор внешней нормали к S_0 в точке $P \in S_2$. Функция $f_2(t, P, T(t, P))$ имеет смысл проекции на направление $\mathbf{n}(P)$ вектора плотности теплового потока, подводимого извне к поверхности S_2 в окрестности точки $P \in S_2$. Если эта функция не зависит от $T(t, P)$, то (2.55) называют граничными условиями II рода [XII], а если

$$f_2(t, P, T(t, P)) = \alpha(t, P)(T_c(t, P) - T(t, P)), \quad P \in S_2, \quad (2.57)$$

то граничными условиями III рода, причем T_c и $\alpha \geq 0$ -- заданные температура внешней среды и коэффициент теплообмена с этой средой. При $\alpha = 0$ имеем $f_2 = 0$, что соответствует частному случаю граничных условий II рода на идеально теплоизолированной поверхности. Поскольку значения f_2 по физическому смыслу задачи конечны, то при увеличении интенсивности теплообмена ($\alpha \rightarrow \infty$) имеем $T(t, P) \rightarrow T_c(t, P)$. В этом случае граничные условия III рода переходят в граничные условия I рода.

Пример 2.3. Рассмотрим прямой стержень, поперечное сечение которого является правильным шестиугольником со

стороной l (рис. 2.6). Примем, что объемная мощность $I_V^{(q)}$ источников энергии в стержне, коэффициент теплообмена α на его поверхности и температура T_c внешней среды не изменяются по длине L стержня.

Если $l \ll L$, то поперечные сечения стержня, достаточно удаленные от его торцов, будут находиться примерно в одинаковых условиях, так что переносом тепловой энергии вдоль стержня можно пренебречь. Поэтому процесс переноса тепловой энергии в каждом из таких сечений можно рассматривать независимо от процесса в соседних сечениях, а задачу переноса тепловой энергии сформулировать как двумерную в системе координат Ox_1x_2 . Тогда искомой будет функция $T = T(t, x_1, x_2)$, описывающая нестационарное температурное поле в поперечном сечении стержня. В этом случае (2.53) примет вид

$$c \frac{\partial T}{\partial t} = \frac{\partial}{\partial x_1} \left(\lambda \frac{\partial T}{\partial x_1} \right) + \frac{\partial}{\partial x_2} \left(\lambda \frac{\partial T}{\partial x_2} \right) + I_V^{(q)}. \quad (2.58)$$

На контуре шестиугольного поперечного сечения имеем граничные условия III рода

$$\lambda(\nabla T) \mathbf{n}(P) = \alpha(T_c - T), \quad (2.59)$$

причем в данном случае

$$(\nabla T) \mathbf{n}(P) = \frac{\partial T}{\partial x_1} n_1 + \frac{\partial T}{\partial x_2} n_2,$$

где n_1 и n_2 — направляющие косинусы единичного вектора внешней нормали к контуру. В математическую формулировку задачи должно входить начальное условие в виде заданного начального распределения температуры в рассматриваемом поперечном сечении стержня:

$$T(0, x_1, x_2) = T_0(x_1, x_2). \quad (2.60)$$

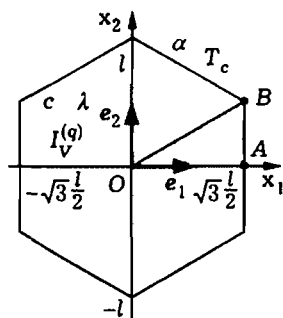


Рис. 2.6

Если $T_0(x_1, x_2) = \text{const}$, а $I_V^{(q)}$, α , T_c , c и λ не зависят от координат, то в силу симметрии поперечного сечения стержня решение задачи достаточно найти для треугольника OAB с идеально теплоизолированными сторонами OA и OB (см. рис. 2.6). В этом случае вместо (2.59) имеем

$$\begin{cases} \lambda \frac{\partial T}{\partial x_1} = \alpha(T_c - T) \text{ на } AB; \\ \lambda \frac{\partial T}{\partial x_2} = 0 \text{ на } OA; \\ \lambda \left(\frac{\partial T}{\partial x_1} n'_1 + \frac{\partial T}{\partial x_2} n'_2 \right) = 0 \text{ на } OB, \end{cases} \quad (2.61)$$

причем $n'_1 = -1/2$ и $n'_2 = \sqrt{3}/2$. #

В процессе теплопроводности возможно возникновение поверхности, для точек которой не будет выполнено условие непрерывности функции температуры или производных этой функции. Рассмотрим сначала эту ситуацию на примерах.

Пример 2.4. Пусть твердое тело объемом V_0 ограничено поверхностью S_0 и разделено на две части V_1 и V_2 поверхностью S^* (рис. 2.7), причем λ_1 и λ_2 — коэффициенты теплопроводности материала частей V_1 и V_2 соответственно. В точках $P \in S_0$ заданы постоянные во времени значения температуры $T(P) = f_1(P)$, а в каждой из частей V_1 и V_2 тела установившееся

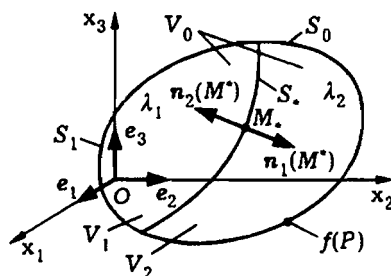


Рис. 2.7

распределение температуры при отсутствии объемных источников тепловой энергии удовлетворяет уравнениям

$$\begin{cases} \nabla(\lambda_1 \nabla T_1(M_1)) = 0, & M_1 \in V_1, \\ \nabla(\lambda_2 \nabla T_2(M_2)) = 0, & M_2 \in V_1, \end{cases} \quad (2.62)$$

которые следуют из (2.53) при $\frac{\partial T}{\partial t} \equiv 0$ и $I_V^{(q)} \equiv 0$.

Предположим, что в точках $M^* \in S^*$ тепловой контакт частей V_1 и V_2 тела является идеальным, т.е. $T_1(M^*) = T_2(M^*)$ и функция

$$T(M) = \begin{cases} T_1(M), & M \in V_1; \\ T_2(M), & M \in V_2; \\ T_1(M) = T_2(M), & M \in S^*, \end{cases} \quad (2.63)$$

непрерывна в точках $M \in S^*$. Рассмотрим баланс тепловых потоков, проходящих через элементарный участок dS^* поверхности S^* в окрестности некоторой точки $M^* \in S^*$ (см. рис. 2.7). Из V_1 к этому участку в соответствии с (1.12) (законом Фурье) поступает тепловой поток $dQ_1 = -\lambda_1(\nabla T_1)\mathbf{n}_1(M^*)dS^*$, где $\mathbf{n}_1(M^*)$ — единичный вектор нормали в точке $M^* \in S^*$, внешней по отношению к части V_1 тела. Аналогично из части V_2 тела к данному участку поступает тепловой поток $dQ_2 = -\lambda_2(\nabla T_2)\mathbf{n}_2(M^*)dS^*$. В сумме эти тепловые потоки равны нулю: $dQ_1 + dQ_2 = 0$. Отсюда имеем

$$\lambda_1(\nabla T_1)\mathbf{n}_1(M^*) + \lambda_2(\nabla T_2)\mathbf{n}_2(M^*) = 0,$$

или, учитывая, что $\mathbf{n}_1 = -\mathbf{n}_2$, получаем

$$\lambda_1 \frac{\partial T_1}{\partial n_1} = \lambda_2 \frac{\partial T_2}{\partial n_1}. \quad (2.64)$$

Таким образом, если через поверхность передается тепловой поток и $\lambda_1 \neq \lambda_2$, то $\frac{\partial T_1}{\partial n_1} \neq \frac{\partial T_2}{\partial n_1}$, т.е. производная $\frac{\partial T}{\partial n_1}$ по направлению вектора $\mathbf{n}_1(M^*)$ функции (2.63), непрерывной в точке

$M^* \in S^*$, терпит в этой точке разрыв. Равенство $T_1(M^*) = T_2(M^*)$ и (2.64) составляют граничные условия IV рода, или граничные условия идеального теплового контакта.

Отличие теплового контакта между частями V_1 и V_2 тела в окрестности точки $M^* \in S^*$ от идеального можно характеризовать некоторым конечным значением $\alpha_k \geq 0$ коэффициента контактной проводимости (идеальному тепловому контакту соответствует $\alpha_k \rightarrow \infty$, а значение $\alpha_k = 0$ отвечает идеально теплоизолированному участку поверхности S^*). Тогда вместо (2.64) для точки $M^* \in S^*$ имеем

$$\lambda_1 \frac{\partial T_1}{\partial n_1} = \alpha_k (T_1 - T_2) = \lambda_2 \frac{\partial T_2}{\partial n_1}. \quad (2.65)$$

Отсюда следует, что при конечном значении α_k передачу через поверхность S^* теплового потока сопровождает скачок температуры $[T(M^*)] = T_1(M^*) - T_2(M^*)$, $M^* \in S^*$, в направлении, противоположном вектору $n_1(M^*)$, где под $T_1(M^*)$ и $T_2(M^*)$ в данном случае надо понимать пределы функций $T_1(M_1)$ и $T_2(M_2)$ при сближении с M^* точек M_1 и M_2 соответственно. #

Если при переходе через некоторую поверхность рассматриваемая функция не является непрерывной, то говорят о **поверхности сильного разрыва** по отношению к этой функции, а если функция непрерывна, а разрывна хотя бы одна из ее частных производных, то говорят о **поверхности слабого разрыва**. При идеальном тепловом контакте частей тела с различными значениями λ_1 и λ_2 контактная поверхность S^* в примере 2.4 является поверхностью слабого разрыва по отношению к функции (2.63) температуры $T(M)$, $M \in V$. В случае неидеального теплового контакта S^* будет поверхностью сильного разрыва по отношению к функции температуры.

Пример 2.5. Рассмотрим двухслойную пластину, слои которой имеют различные значения коэффициентов теплопроводности $\lambda_1 \neq \lambda_2$, и примем, что температурное поле в пластине одномерно, т.е. изменяется лишь в направлении координатной

оси Ox_1 (рис. 2.8). Пусть при установившемся процессе теплопроводности через пластину в положительном направлении этой оси проходит тепловой поток плотностью q . Тогда в случае идеального теплового контакта между слоями в соответствии с (2.64) в плоскости контакта ($x_1 = 0$) имеем

$$q = -\lambda_1 \frac{dT_1}{dx_1} = -\lambda_2 \frac{dT_2}{dx_1}. \quad (2.66)$$

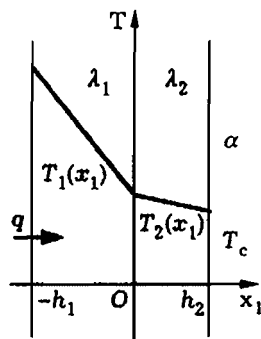


Рис. 2.8

Ясно, что при постоянных значениях λ_1 и λ_2 температура по толщине h_1 первого слоя и по толщине h_2 второго слоя будет изменяться линейно (на рис. 2.8 изображено распределение температуры для случая $\lambda_1 < \lambda_2$). При заданных значениях q , температуры T_c внешней среды и коэффициента α теплообмена на поверхности пластины при $x_1 = h_2$ из (2.66) несложно найти, что $T_2(h_2) = T_c + \frac{q}{\alpha}$, $T_2(0) = T_1(0) = T_2(h_2) + \frac{qh_2}{\lambda_2}$ и $T_1(-h_1) = T_1(0) + \frac{qh_1}{\lambda_1}$. Отсюда получаем

$$\Delta T_1 = T_1(-h_1) - T_1(0) = \frac{qh_1}{\lambda_1} \quad \text{и} \quad \Delta T_2 = T_2(0) - T_2(h_2) = \frac{qh_2}{\lambda_2}.$$

При $\frac{h_2}{\lambda_2} \ll \frac{h_1}{\lambda_1}$ перепадом ΔT_2 температуры по толщине второго слоя можно пренебречь по сравнению с перепадом ΔT_1 по толщине первого слоя и принять температуру второго слоя однородной по его толщине, т.е. $T_2(x_1) = \text{const}$. Таким образом, существенное различие термических сопротивлений $\frac{h_1}{\lambda_1}$ и $\frac{h_2}{\lambda_2}$ слоев позволяет упростить математическую модель процесса установившейся теплопроводности в двухслойной пластине. Возникающая при этом погрешность пропорциональна $\frac{h_2 \lambda_1}{h_1 \lambda_2}$. Если при сопоставимых толщинах слоев их теплопроводности отличаются на два и более порядков (это характерно,

например, для металлов и теплоизолирующих материалов), то погрешность вычисления температур не превышает процента от ΔT_1 .

Принятое упрощение можно использовать и в случае нестационарного процесса теплопроводности, описываемого в слоях пластины уравнениями

$$\frac{\partial T_1}{\partial t} = \tilde{a}_1 \frac{\partial^2 T_1}{\partial x_1^2}, \quad \frac{\partial T_2}{\partial t} = \tilde{a}_2 \frac{\partial^2 T_2}{\partial x_1^2}, \quad (2.67)$$

где $\tilde{a}_1 = \lambda_1/c_1$ и $\tilde{a}_2 = \lambda_2/c_2$ — температуропроводности пластин, а c_1 и c_2 — их объемные теплоемкости. Помимо начальных распределений $T_1(0, x_1) = T_1^0(x_1)$ и $T_2(0, x_1) = T_2^0(x_1)$ температуры в слоях в момент времени $t = 0$ для решения (2.67) необходимо задать граничные условия:

$$\begin{cases} -\lambda_1 \frac{\partial T_1}{\partial x_1} = q, & x_1 = -h_1; \\ T_1 = T_2 \text{ и } \lambda_1 \frac{\partial T_1}{\partial x_1} = \lambda_2 \frac{\partial T_2}{\partial x_1}, & x_1 = 0; \\ \lambda_2 \frac{\partial T_2}{\partial x_1} = \alpha(T_c - T_2), & x_1 = h_2. \end{cases} \quad (2.68)$$

Интегрируя второе уравнение (2.67) по толщине h_2 второго слоя и учитывая (2.68), получаем

$$\begin{aligned} \int_0^{h_2} \frac{\partial T_2(t, x_1)}{\partial t} dx_1 &= \tilde{a}_2 \int_0^{h_2} \frac{\partial^2 T_2(t, x_1)}{\partial x_1^2} dx_1 = \frac{\lambda_2}{c_2} \frac{\partial T_2(t, x_1)}{\partial x_1} \Big|_0^{h_2} = \\ &= \frac{\alpha}{c_2} (T_c - T_2(t, h_2)) - \frac{\lambda_1}{c_2} \frac{\partial T_1(t, 0)}{\partial x_1}. \end{aligned}$$

Отсюда, полагая при $h_2/\lambda_2 \ll h_1/\lambda_1$ температуру по толщине этого слоя однородной, т.е. $T_2(t, x_1) = T_2(t) = T_1(t, 0)$, находим

$$h_2 c_2 \frac{\partial T_1}{\partial t} = \alpha(T_c - T_1) - \lambda_1 \frac{\partial T_1}{\partial x_1} \quad \text{при } x_1 = 0. \quad (2.69)$$

Это равенство и первое равенство (2.68) являются в принятом приближении граничными условиями для первого уравнения (2.67). Отметим, что мы, пренебрегая термическим сопротивлением второго слоя пластины, учитываем его теплоемкость. Характер распределения температуры по толщине пластины в некоторый момент времени t в этом приближении показан на рис. 2.9.

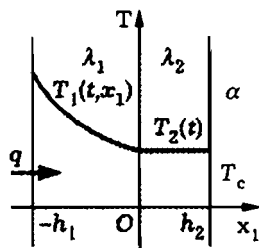


Рис. 2.9

Таким образом, плоскость $x_1 = 0$ остается и при принятом приближении поверхностью слабого разрыва по отношению как к установившемуся, так и к нестационарному распределениям температуры в рассматриваемой двухслойной пластине.

Пример 2.6. Поверхность слабого разрыва по отношению к распределению температуры может возникнуть при фазовом переходе в среде (например, в процессах плавления, испарения, сублимации, затвердевания или конденсации), причем эта поверхность может перемещаться во времени t относительно частиц среды. Примем, что переход среды из одного агрегатного состояния в другое происходит при фиксированной температуре $T^* = \text{const}$ и сопровождается поглощением единиц массы среды тепловой энергии r (величину r называют теплотой плавления, испарения или сублимации и для этих процессов обычно $r > 0$).

Рассмотрим в неподвижной среде одномерное поле температуры, изменяющейся лишь в направлении координатной оси Ox_1 . Тогда плоскость $x_1 = x^*$, в которой температура среды равна T^* , будет поверхностью слабого разрыва относительно распределения температуры в среде, если в процессе фазового перехода тепловой контакт между различными агрегатными состояниями этой среды сохраняется идеальным, т.е. $T_1(t, x^*) = T_2(t, x^*) = T^*$ (рис. 2.10). Эта плоскость разделяет два агрегатных состояния среды с теплопроводностями λ_1 и λ_2 ,

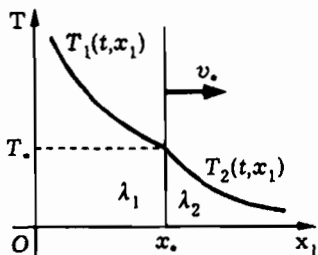


Рис. 2.10

существующие при температурах соответственно выше и ниже T^* . Скорость $v^* = dx^*/dt$ движения поверхности слабого разрыва в данном случае можно найти из баланса тепловой энергии при $x_1 = x^*$. С учетом (1.12) запишем

$$-\lambda_1 \frac{T_1}{\partial x_1} + \lambda_2 \frac{T_2}{\partial x_1} = \rho r v^*, \quad (2.70)$$

где ρ — плотность среды перед изменением ее агрегатного состояния. Ясно, что $v^* > 0$ при $r > 0$, если левая часть (2.70) положительна, и наоборот. Таким образом, движущаяся плоскость фазового перехода при $r \neq 0$ является поверхностью сильного разрыва по отношению к функции плотности теплового потока. #

Условия на поверхности S^* разрыва в процессах теплопроводности можно сформулировать в более общем виде. Для этого в четырехмерном евклидовом арифметическом пространстве $\mathbb{R}^4$ с ортонормированным базисом $e = (e_t, e_1, e_2, e_3)$ введем вектор $\tilde{q} = \varepsilon_T v_0 e_t + q$, причем, как и в (2.25), v_0 — произвольная константа, имеющая размерность скорости. При этом локальная форма (2.51) закона сохранения тепловой энергии перейдет в обобщенную дивергентную форму $\tilde{\nabla} \tilde{q} = I_V^{(q)}$.

Пусть векторная функция $\tilde{q}(t, M)$ не имеет непрерывных производных или не является непрерывной в точках $M^* \in S^*$ поверхности S^* , делящей область V_0 на подобласти V_1 и V_2 . Примем, что эта поверхность может перемещаться, причем ее точки $M^* \in S^*$ имеют скорость $v^*(t, M^*)$. Тогда аналогично (2.44) в любой точке $M^* \in S^*$ получим

$$\begin{aligned} [q(t, M^*) n^*(M^*)] + I_S^{(q)}(t, M^*) = \\ = [\varepsilon_T(t, M^*)] v^*(t, M^*) n^*(M^*), \quad (2.71) \end{aligned}$$

где $I_S^{(q)}$ — поверхностная интенсивность источников тепловой энергии. Если поверхность S^* неподвижна, то $|v^*| = 0$ и из (2.71) следует условие

$$[q_n(t, x(M^*))] + I_S^{(q)}(t, x(M^*)) = 0 \quad (2.72)$$

для скачка значений $q_n = \mathbf{q}(t, M^*) \mathbf{n}^*(M^*)$ проекции вектора $\mathbf{q}$ на направление единичного вектора $\mathbf{n}^*(M^*)$ нормали к S^* при переходе в точке $M^* \in S^*$ через S^* в противоположном направлении.

В примере 2.5 плоскость $x_1 = 0$, разделяющая слои двухслойной пластины, неподвижна. Поэтому при отсутствии в этой плоскости источников тепловой энергии равенство (2.66) в случае установившегося процесса теплопроводности и второе равенство при $x_1 = 0$ в (2.68) в случае нестационарного процесса соответствуют условию (2.72) при $I_S^{(q)} \equiv 0$. Если же в нестационарном процессе теплопроводности пренебречь термическим сопротивлением второго слоя пластины и учитывать лишь его теплоемкость, то это равносильно появлению в плоскости $x_1 = 0$ источника тепловой энергии с поверхностной интенсивностью $I_S^{(q)} = -h_2 c_2 \frac{\partial T_1}{\partial t}$. В этом случае граничное условие (2.69) также соответствует (2.72), если учесть, что $q_n = -\alpha(T_c - T_1)$ перед переходом через плоскость $x_1 = 0$ в направлении, противоположном положительному направлению оси Ox_1 .

Пусть плоскость $x_1 = x^*$ фазового перехода, в которой при фиксированной температуре $T^* = \text{const}$ происходит изменение агрегатного состояния среды, имеет в момент времени t скорость v^* (см. пример 2.6). В этой плоскости при $v^* > 0$ происходит скачок $[\epsilon_T] = \rho r$ объемной плотности тепловой энергии ϵ_T , что в соответствии с (2.71) при $I_S^{(q)} \equiv 0$ связано со скачком $[q_n]$ плотности теплового потока, учитываемым условием (2.70). В сопутствующей системе координат, движущейся вместе с плоскостью фазового перехода, трактовка условия (2.71) будет иной. Теперь относительно системы координат, движущейся

вместе с этой плоскостью, не происходит затрат тепловой энергии на изменение значения ε_T , но в этой плоскости действует источник тепловой энергии с поверхностной интенсивностью $-\rho v^*$, что также приводит к (2.70).

2.4. Закон сохранения количества движения

Согласно второму закону Ньютона, скорость изменения количества движения любого тела равна сумме всех сил, действующих на это тело. Для *сплошной среды*, находящейся в объеме V , ее количество движения в соответствии с (1.5) равно

$$\mathbf{J} = \int_V \rho \mathbf{v} dV, \quad (2.73)$$

где ρ и $\mathbf{v}$ — *плотность* и вектор скорости *среды*. Действующие на среду силы распределены по ее объему и ограничивающей его поверхности. Такие силы называют соответственно *объемными* и *поверхностными*. Примером объемной силы является сила тяжести, а поверхностной — сила давления.

Объемные силы будем характеризовать вектором *объемной плотности*

$$\mathbf{b} = \lim_{d \rightarrow 0} \frac{\mathbf{R}}{\Delta V}, \quad (2.74)$$

где $\mathbf{R}$ — вектор равнодействующей сил, распределенных по объему ΔV среды, а d — диаметр области, занятый этим объемом. Аналогично вектором плотности поверхностных сил будет предел

$$\mathbf{p} = \lim_{d' \rightarrow 0} \frac{\mathbf{P}}{\Delta S}, \quad (2.75)$$

где $\mathbf{P}$ — вектор равнодействующей сил, распределенных по участку ΔS поверхности $\hat{S}$, ограничивающей рассматриваемый объем V среды, а d' — диаметр области, соответствующий этому участку поверхности. Модули векторов $\mathbf{b}$ и $\mathbf{p}$ измеряют в Н/м^3 и Н/м^2 (или в Па — паскалях) соответственно.

Пример 2.7. Если среда находится в однородном поле тяготения, характеризуемом вектором $\mathbf{g}$ ускорения свободного падения, то на массу Δm в объеме ΔV будет действовать сила тяжести $\mathbf{R} = \mathbf{g}\Delta m$. Тогда, согласно (1.1) и (2.74), найдем

$$\mathbf{b} = \lim_{\Delta V \rightarrow 0} \frac{\mathbf{g}\Delta m}{\Delta V} = \mathbf{g} \lim_{\Delta V \rightarrow 0} \frac{\Delta m}{\Delta V} = \rho \mathbf{g}.$$

При действии на плоский участок поверхности ΔS внешнего давления p равнодействующая будет равна $\mathbf{P} = -p\Delta S\mathbf{n}$, где $\mathbf{n}$ — единичный вектор нормали к участку ΔS , внешней по отношению к объему, занятому средой. В этом случае из (2.75) следует, что

$$\mathbf{p} = \lim_{\Delta S \rightarrow 0} \frac{-p\Delta S\mathbf{n}}{\Delta S} = -p\mathbf{n}.$$

Можно показать, что и в случае произвольной криволинейной поверхности при действии внешнего давления вектор плотности поверхностных сил в точке поверхности с единичным вектором внешней нормали $\mathbf{n}$ будет равен $-p\mathbf{n}$. #

Равнодействующие сил, распределенных по объему V и ограничивающей его поверхности S , будут соответственно

$$\int_V \mathbf{b} dV \quad \text{и} \quad \int_S \mathbf{p} dS,$$

поскольку силы, действующие между отдельными частями среды внутри объема V , взаимно уравновешены. Тогда, согласно второму закону Ньютона и равенству (2.73), получим интегральную форму закона сохранения количества движения

$$\frac{d}{dt} \int_V \rho \mathbf{v} dV = \int_V \mathbf{b} dV + \int_S \mathbf{p} dS. \quad (2.76)$$

Ясно, что если в некоторых точках объема V и поверхности S приложены сосредоточенные силы, то в правую часть (2.76) должна войти также и равнодействующая таких сил.

Заменяя в (2.6) C_m на проекцию v_i вектора $\mathbf{v}$ на ось Ox_i , $i = 1, 2, 3$, прямоугольной системы координат $Ox_1x_2x_3$ с ортами $\mathbf{e}_i$, имеем

$$\frac{d}{dt} \int_V \rho v_i dV = \int_V \rho \frac{dv_i}{dt} dV.$$

Поскольку $\mathbf{v} = v_1 \mathbf{e}_1 + v_2 \mathbf{e}_2 + v_3 \mathbf{e}_3$, то

$$\frac{d}{dt} \int_V \rho \mathbf{v} dV = \int_V \rho \frac{d\mathbf{v}}{dt} dV,$$

так что вместо (2.76) для произвольного объема V , ограниченного поверхностью S , можно написать

$$\int_V \left(\rho \frac{d\mathbf{v}}{dt} - \mathbf{b} \right) dV = \int_S \mathbf{p} dS. \quad (2.77)$$

Понятие плотности поверхностных сил применимо к любой точке $M \in V$ объема V среды, если в этой точке задать единичный вектор $\mathbf{n}$ и провести через нее перпендикулярно этому вектору элементарную площадку ΔS . В этом случае вектор $\mathbf{p}$, определяемый согласно (2.75), принято называть **вектором напряжений**. При этом сила $\mathbf{p} \Delta S$ будет действовать со стороны части среды на ту ее часть, для которой вектор $\mathbf{n}$ является внешней нормалью к площадке ΔS .

Ясно, что вектор $\mathbf{p}$ в фиксированной точке среды зависит от ориентации площадки ΔS , задаваемой единичным вектором $\mathbf{n} = n_1 \mathbf{e}_1 + n_2 \mathbf{e}_2 + n_3 \mathbf{e}_3$ с направляющими косинусами n_1, n_2, n_3 относительно координатных осей Ox_1, Ox_2, Ox_3 соответственно (рис. 2.11). Рассмотрим пирамиду с треугольным основанием ΔS , боковые грани которой параллельны координатным плоскостям. Объем такой пирамиды равен $\Delta V = h \Delta S / 3$, где h — высота, опущенная из вершины пирамиды на основание ΔS в точку $M \in \Delta S$, а площади боковых граней, перпендикулярных координатным осям, равны соответственно

$$\Delta S_1 = \Delta S n_1, \quad \Delta S_2 = \Delta S n_2, \quad \Delta S_3 = \Delta S n_3.$$

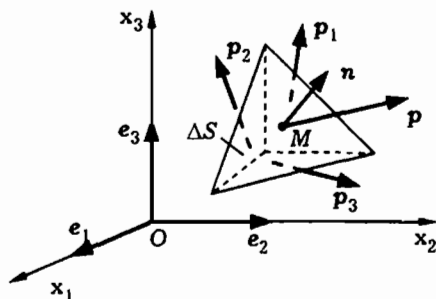


Рис. 2.11

Применяя теоремы о среднем значении для тройного и двойного интегралов к интегралам по объему ΔV и ограничивающей его поверхности [VII], в соответствии с (2.77) имеем

$$\rho \frac{dv}{dt} \frac{h\Delta S}{3} - b \frac{h\Delta S}{3} = p\Delta S + p_1\Delta S n_1 + p_2\Delta S n_2 + p_3\Delta S n_3,$$

где p_1, p_2, p_3 — векторы напряжений на боковых гранях пирамиды, перпендикулярных соответствующим координатным осям (чертой сверху обозначены средние значения параметров в объеме и на соответствующих участках поверхности). Сокращая это равенство на ΔS и затем стягивая объем пирамиды в точку M так, что при $h \rightarrow 0$ пирамида, сохраняя ориентацию относительно координатных осей, остается подобной сама себе, получаем

$$p + p_1 n_1 + p_2 n_2 + p_3 n_3 = 0. \quad (2.78)$$

Равенство (2.78) должно быть выполнено для любой площадки, проходящей через любую точку в объеме V или на ограничивающей его поверхности S . Рассмотрим квадратную матрицу (σ_{ij}) третьего порядка с такими ее элементами σ_{ij} , $i, j = 1, 2, 3$, что

$$\begin{cases} -p_1 = \sigma_{11}e_1 + \sigma_{12}e_2 + \sigma_{13}e_3, \\ -p_2 = \sigma_{21}e_1 + \sigma_{22}e_2 + \sigma_{23}e_3, \\ -p_3 = \sigma_{31}e_1 + \sigma_{32}e_2 + \sigma_{33}e_3. \end{cases} \quad (2.79)$$

Тогда для проекций вектора $\mathbf{p} = p_1\mathbf{e}_1 + p_2\mathbf{e}_2 + p_3\mathbf{e}_3$ на координатные оси, согласно (2.78), имеем

$$\begin{cases} p_1 = n_1\sigma_{11} + n_2\sigma_{21} + n_3\sigma_{31}, \\ p_2 = n_1\sigma_{12} + n_2\sigma_{22} + n_3\sigma_{32}, \\ p_3 = n_1\sigma_{13} + n_2\sigma_{23} + n_3\sigma_{33}. \end{cases} \quad (2.80)$$

Эти соотношения, устанавливающие линейную зависимость проекций вектора напряжения в площадке от проекций вектора нормали к этой площадке, составляют содержание **фундаментальной леммы Коши**.

Матрица $(\sigma_{ji}) = (\sigma_{ij})^T$ преобразует проекции n_j , $j = 1, 2, 3$, вектора $\mathbf{n}$ в проекции p_i , $i = 1, 2, 3$, вектора $\mathbf{p}$ в соответствии с соотношениями (2.80), которые можно записать в виде

$$p_i = \sum_{j=1}^3 n_j \sigma_{ji}, \quad i = 1, 2, 3,$$

или, используя правило суммирования слагаемых по повторяющимся индексам (здесь по индексу j), в виде

$$p_i = n_j \sigma_{ji}, \quad i, j = 1, 2, 3. \quad (2.81)$$

Известно [IV], что такая матрица соответствует **тензору** второго ранга. В данном случае его называют **тензором напряжений**. Из (2.81) видно, что элемент σ_{ji} матрицы (σ_{ji}) — компонента тензора напряжений — является проекцией на координатную ось Ox_i вектора $\mathbf{p}$ напряжения в площадке, единичный вектор внешней нормали к которой совпадает с ортом $\mathbf{e}_j$, т.е. $n_j = 1$. Диагональные элементы σ_{11} , σ_{22} , σ_{33} этой матрицы называют **нормальными напряжениями**, а внедиагональные — **касательными** к площадке, причем первый индекс элемента σ_{ji} указывает на отмеченную выше ориентацию площадки. Если $\sigma_{ji} > 0$, то считают, что соответствующее напряжение действует в положительном направлении оси Ox_i ,

$i = 1, 2, 3$, а при $\sigma_{ji} < 0$ — в противоположном. На рис. 2.12 показаны направления действия нормальных и касательных напряжений для случая $\sigma_{ji} > 0$, $i, j = 1, 2, 3$.

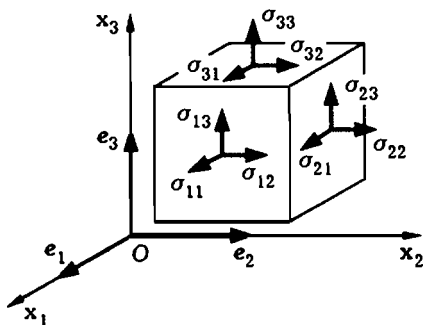


Рис. 2.12

Из (2.80) следует, что

$$p_i = (\sigma_{1i}e_1 + \sigma_{2i}e_2 + \sigma_{3i}e_3)n = p^{(i)}n, \quad i = 1, 2, 3, \quad (2.82)$$

т.е. p_i является проекцией вектора

$$p^{(i)} = \sigma_{1i}e_1 + \sigma_{2i}e_2 + \sigma_{3i}e_3 \quad (2.83)$$

на направление единичного вектора n нормали. Поскольку (2.82) верно и для точек на поверхности S , ограничивающей объем V , занятый средой, то (2.77) можно записать в проекциях на координатные оси в виде

$$\int_V \left(\rho \frac{dv_i}{dt} - b_i \right) dV = \int_S p_i dS = \int_S p^{(i)}n dS, \quad i = 1, 2, 3,$$

где v_i и b_i — проекции векторов v и b на координатные оси Ox_i , $i = 1, 2, 3$. Отсюда, используя формулу Остроградского — Гаусса, находим

$$\int_V \left(\rho \frac{dv_i}{dt} - b_i - \nabla p^{(i)} \right) dV = 0, \quad i = 1, 2, 3. \quad (2.84)$$

В прямоугольной системе координат с учетом правила суммирования по повторяющимся индексам

$$\nabla p^{(i)} = \frac{\partial \sigma_{1i}}{\partial x_1} + \frac{\partial \sigma_{2i}}{\partial x_2} + \frac{\partial \sigma_{3i}}{\partial x_3} = \frac{\partial \sigma_{ji}}{\partial x_j}, \quad j = 1, 2, 3.$$

Тогда в случае непрерывности подынтегральной функции в (2.84), учитывая произвольность рассматриваемого объема V , получаем локальную форму закона сохранения количества движения в проекциях на координатные оси

$$\rho \frac{dv_i}{dt} - b_i - \frac{\partial \sigma_{ji}}{\partial x_j} = 0, \quad i, j = 1, 2, 3, \quad (2.85)$$

называемую **уравнениями движения**. В векторной форме (2.84) и (2.85) с учетом правила суммирования по повторяющимся индексам $i, j = 1, 2, 3$ примут вид

$$\int_V \left(\rho \frac{dv}{dt} - b - \frac{\partial \sigma_{ji}}{\partial x_j} e_i \right) dV = 0, \quad \rho \frac{dv}{dt} - b - \frac{\partial \sigma_{ji}}{\partial x_j} e_i = 0. \quad (2.86)$$

Если среда неподвижна, то (2.85) переходят в **уравнения равновесия**

$$\frac{\partial \sigma_{ji}}{\partial x_j} + b_i = 0, \quad \text{или} \quad \frac{\partial \sigma_{ji}}{\partial x_j} e_i + b = 0, \quad i, j = 1, 2, 3. \quad (2.87)$$

При этом (2.84) принимает вид

$$\int_V (\nabla p^{(i)} + b_i) dV = 0, \quad i = 1, 2, 3. \quad (2.88)$$

Отметим, что (2.88) сохраняет силу и в случае, если множество точек в $V \in \mathbb{R}^3$, в которых векторная функция $p^{(i)}$ не имеет непрерывных производных по координатам радиус-вектора x или даже не является непрерывной, имеет меру Лебега, равную нулю. Если множество таких точек образует некоторую

поверхность разрыва S^* , делящую V на две подобласти V_1 и V_2 (см. рис. 2.1), то процедура, аналогичная использованной в 2.2, дает условие для скачка $[p^{(i)}n^*]$ проекции вектора $p^{(i)}$ на направление единичного вектора n^* нормали к S^* в виде

$$[p^{(i)}n^*] + p_i^* = 0, \quad i = 1, 2, 3, \quad (2.89)$$

где p_i^* — проекции на координатные оси Ox_i вектора p^* плотности внешних поверхностных сил, которые могут быть приложены на этой поверхности. Такие силы могут возникнуть, например, при взаимодействии среды с электромагнитным полем*. Умножая (2.89) на направляющие косинусы n_i , $i = 1, 2, 3$, вектора нормали n^* , суммируя по индексу i и учитывая (2.83), приходим к условию

$$[\sigma_{ji}n_i^*n_j^*] + p^*n^* = 0, \quad i, j = 1, 2, 3, \quad (2.90)$$

выраженному через компоненты тензора напряжений и записанному с учетом правила суммирования по повторяющимся индексам.

Чтобы получить условия на поверхности разрыва в случае движущейся среды, в уравнении неразрывности (2.2) перейдем к координатной записи дифференциальной операции дивергенции:

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho v_j}{\partial x_j} = 0, \quad i, j = 1, 2, 3,$$

а также учтем, что

$$\frac{dv_i}{dt} = \frac{\partial v_i}{\partial t} + v_j \frac{\partial v_i}{\partial x_j}, \quad i, j = 1, 2, 3.$$

Тогда вместо (2.85) будем иметь

$$\frac{\partial \rho v_i}{\partial t} + \frac{\partial}{\partial x_j} (\rho v_j v_i - \sigma_{ji}) - b_i = 0, \quad i, j = 1, 2, 3. \quad (2.91)$$

*См.: Седов Л.И.

В четырехмерном евклидовом арифметическом пространстве $\mathbb{R}^4$ с ортонормированным базисом $e = (e_t, e_1, e_2, e_3)$ введем векторы

$$\tilde{p}^{(i)} = \rho v_i v_0 e_t + (\rho v_j v_i - \sigma_{ji}) e_j, \quad i, j = 1, 2, 3,$$

где, как и в (2.25), v_0 — произвольная константа, имеющая размерность скорости. Используя векторный оператор (2.26), перейдем от (2.91) к *обобщенной дивергентной форме* закона сохранения количества движения в виде

$$\tilde{\nabla} \tilde{p}^{(i)} - b_i = 0, \quad i = 1, 2, 3. \quad (2.92)$$

Теперь процедура, аналогичная использованной в 2.2, приводит к следующим условиям на поверхности разрыва S^* :

$$[(\rho v_j v_i - \sigma_{ji}) n_j^*] = [\rho v_i] v^* n^* + p_i^*, \quad i = 1, 2, 3, \quad (2.93)$$

где v^* — вектор скорости перемещения точек поверхности S^* . Умножая (2.93) на n_i^* и суммируя не только по индексу j , но и по индексу i , получаем условие

$$[\rho (v n^*)^2 - \sigma_{ji} n_j^* n_i^*] = [\rho v n^*] v^* n^* + p^* n^*, \quad i, j = 1, 2, 3, \quad (2.94)$$

обобщающее (2.90) на случай движущейся среды. Если поверхность S^* неподвижна, т.е. $|v^*| = 0$, то первое слагаемое в правой части (2.94) исчезает. Тот же результат имеет место, если среда несжимаема, так как в этом случае в соответствии с (2.23) $[\rho v n^*] = 0$ на S^* .

Вектор момента K количества движения среды, находящейся в объеме V , можно представить в виде (1.5). Интегральная форма закона сохранения момента количества движения относительно некоторой фиксированной точки устанавливает равенство скорости изменения вектора K моменту относительно этой же точки всех действующих объемных и поверхностных сил. Можно показать*, что при отсутствии моментов, распре-

*См.: Седов Л.И.

деленных по объему V или ограничивающей его поверхности S , из этого закона следует симметрия тензора напряжений, т.е. $\sigma_{ij} = \sigma_{ji}$, $i, j = 1, 2, 3$. Однако в общем случае может возникнуть необходимость учитывать распределенные по объему и поверхности моменты (в частности, при движении среды в электромагнитном поле*).

Дополнение 2.1. Формулы векторного анализа в случае неоднородной среды

Законы сохранения физических субстанций в неоднородной сплошной среде при постановке задач математической физики приводят к дифференциальным уравнениям с переменными коэффициентами. Например, в уравнении (2.53) нестационарной теплопроводности, следующем из закона сохранения тепловой энергии (см. 2.3), теплопроводность $\lambda(M)$ среды может быть функцией координат точки $M \in \mathbb{R}^3$. Более того, если среда еще и анизотропна, т.е. ее свойства передавать тепловую энергию различны в различных направлениях, то вместо λ в (2.53) войдет **тензор $\tilde{\lambda}(M)$ теплопроводности**. Он является **тензором** второго ранга и в прямоугольной системе координат $Ox_1x_2x_3$ имеет вид квадратной матрицы третьего порядка:

$$\tilde{\lambda} = \begin{pmatrix} \lambda_{11} & \lambda_{12} & \lambda_{13} \\ \lambda_{21} & \lambda_{22} & \lambda_{23} \\ \lambda_{31} & \lambda_{32} & \lambda_{33} \end{pmatrix}, \quad (2.95)$$

причем $\lambda_{ij} = \lambda_{ji}$, $i, j = 1, 2, 3$. Каждая из **компонент** этого **тензора** может быть функцией координат x_i точки $M \in \mathbb{R}^3$. В этом случае (2.53) примет вид

$$\frac{\partial T}{\partial t} = \frac{1}{c} \sum_{i=1}^3 \sum_{j=1}^3 \frac{\partial}{\partial x_i} \left(\lambda_{ij} \frac{\partial T}{\partial x_j} \right) + \frac{I_V^{(q)}}{c}, \quad (2.96)$$

*См., например: *Можен Ж.*

где T — искомая функция температуры; $c > 0$ — теплоемкость единицы объема среды; $I_V^{(q)}$ — объемная мощность источников тепловой энергии.

Замечание 2.1. Напомним, что симметрическую матрицу вида (2.95) *ортгональным преобразованием* можно привести к диагональному виду. Координатные оси, задаваемые ортонормированным базисом, в котором матрица (2.95) является диагональной, называют **главными осями тензора**, которому соответствует эта матрица. Тогда при совпадении координатных осей Ox_i , $i = 1, 2, 3$, с главными осями тензора $\tilde{\lambda}(M)$ вместо (2.96) получим

$$\frac{\partial T}{\partial t} = \frac{1}{c} \sum_{i=1}^3 \frac{\partial}{\partial x_i} \left(\lambda_i \frac{\partial T}{\partial x_i} \right) + \frac{I_V^{(q)}}{c},$$

где λ_i — диагональные элементы матрицы (2.95) после ее приведения к диагональному виду, называемые в данном случае **главными коэффициентами теплопроводности** анизотропной среды. #

Уравнения с переменными коэффициентами могут быть результатом линеаризации нелинейных *математических моделей* физических процессов. При приближенном решении нелинейных задач *методом итераций* обычно приходится решать такие уравнения на каждой итерации.

Выражение в виде двойной суммы в правой части (2.96) характерно для многих уравнений с переменными коэффициентами, встречающихся в задачах математической физики. В более общем случае m -мерного евклидова арифметического пространства $\mathbb{R}^m$ рассмотрим интегралы

$$\int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) d\Omega \quad \text{и} \quad \int_{\Omega} v \sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) d\Omega$$

по конечной области $\Omega \subset \mathbb{R}^m$ с границей $\partial\Omega$. Здесь u и v — действительные функции, дважды непрерывно дифференцируемые по всем своим аргументам $x_1, x_2, \dots, x_m$ на замыкании $\bar{\Omega}$ области Ω , а переменные коэффициенты $a_{ij} = a_{ji}$, $i, j = \overline{1, m}$, непрерывно дифференцируемы по тем же аргументам на $\bar{\Omega}$.

Если в (1.25) принять

$$u_i = \sum_{j=1}^m a_{ij} \frac{\partial u}{\partial x_j}, \quad i = \overline{1, m},$$

то интеграл в левой части (1.25) совпадает с первым из рассматриваемых интегралов. Тогда получим обобщение *формулы Остроградского* на случай переменных коэффициентов:

$$\int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) d\Omega = \int_{\partial\Omega} \sum_{i=1}^m \sum_{j=1}^m a_{ij} \frac{\partial u}{\partial x_j} n_i d(\partial\Omega). \quad (2.97)$$

Используя (2.97) и равенство

$$\sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij} v \frac{\partial u}{\partial x_j} \right) = v \sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) + \sum_{i=1}^m \sum_{j=1}^m \frac{\partial v}{\partial x_i} a_{ij} \frac{\partial u}{\partial x_j},$$

для второго из рассматриваемых интегралов находим

$$\begin{aligned} & \int_{\Omega} v \sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) d\Omega = \\ &= \int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij} v \frac{\partial u}{\partial x_j} \right) d\Omega - \int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial v}{\partial x_i} a_{ij} \frac{\partial u}{\partial x_j} d\Omega = \\ &= \int_{\partial\Omega} v \sum_{i=1}^m \sum_{j=1}^m a_{ij} \frac{\partial u}{\partial x_j} n_i d(\partial\Omega) - \int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial v}{\partial x_i} a_{ij} \frac{\partial u}{\partial x_j} d\Omega. \quad (2.98) \end{aligned}$$

Меняя в (2.98) местами функции u и v и вычитая полученный результат из (2.98), с учетом равенств $a_{ij} = a_{ji}$, $i, j = \overline{1, m}$, получаем

$$\begin{aligned} \int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m \left(v \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) - u \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial v}{\partial x_j} \right) \right) d\Omega = \\ = \int_{\partial\Omega} \sum_{i=1}^m \sum_{j=1}^m a_{ij} \left(v \frac{\partial u}{\partial x_j} - u \frac{\partial v}{\partial x_j} \right) n_i d(\partial\Omega). \quad (2.99) \end{aligned}$$

Таким образом, (2.98) и (2.99) можно считать обобщением *первой и второй формул Грина* соответственно на случай m -мерного пространства и переменных коэффициентов.

Вопросы и задачи

2.1. Поле скорости несжимаемой среды задано функцией

$$v = \frac{Ax_2}{|x|^2} e_1 + v_2(x_1, x_2) e_2,$$

где $|x| = \sqrt{x_1^2 + x_2^2}$. Найти функцию $v_2(x_1, x_2)$, если $v_2(0, x_2) = 0$, и проверить, является ли поле скорости потенциальным.

2.2. В проходящих через фиксированную точку площадках с единичными векторами нормали n и n' действуют векторы напряжений p и p' соответственно. Доказать, что $pn' = p'n$.

2.3. Найти зависимость вектора плотности объемных сил от координат точки, если тензор напряжений в неподвижной среде задан матрицей

$$\begin{pmatrix} 3ax_1x_2 & 5ax_2^2 & 0 \\ 5ax_2^2 & 0 & 2ax_3 \\ 0 & 2ax_3 & 0 \end{pmatrix}, \quad a = \text{const.}$$

3. МАТЕМАТИЧЕСКИЕ МОДЕЛИ НЕКОТОРЫХ СРЕД

В общем случае количество неизвестных функций, входящих в уравнения, полученные из законов сохранения физических субстанций, превышает количество этих уравнений. Поэтому таких уравнений обычно недостаточно для построения замкнутой формулировки задач математической физики. Так, для получения из закона сохранения тепловой энергии дифференциальных уравнений (2.53) и (2.54) нестационарной теплопроводности, содержащих единственную неизвестную функцию температуры среды, пришлось использовать эмпирический закон (1.12) теплопроводности и зависимость (2.52) *объемной плотности* тепловой энергии от температуры. Эта зависимость является характерным примером *уравнения состояния* среды.

Чтобы использовать (2.85) для формулировки задач математической физики, полагая заданным поле вектора $\mathbf{b}$ плотности объемных сил, необходимо установить связь между вектором $\mathbf{v}$ скорости среды и *тензором напряжений*. Эта связь также может быть выражена соответствующим уравнением состояния среды, часто называемым *математической моделью* среды. Рассмотрим некоторые из таких моделей.

3.1. Модели идеальной жидкости (газа)

К жидкости (или газу) относят такую среду, *вектор напряжений* в любой точке которой в состоянии покоя всегда коллинеарен вектору нормали к любой площадке, проходящей через эту точку. Для твердых тел это имеет место лишь в условиях всестороннего сжатия или растяжения.

Идеальной (невязкой) жидкостью (или идеальным газом) считают среду, для которой указанное свойство сохраняется и при движении. При этом в любой точке такой среды **тензор напряжений** с компонентами σ_{ij} , $i, j = 1, 2, 3$, является **шаровым**, т.е. $\sigma_{ij} = \sigma^\circ$ при $i = j$ и $\sigma_{ij} = 0$ при $i \neq j$. В этом случае величину σ° , взятую с обратным знаком, называют давлением жидкости (или газа) и обозначают p , т.е. $p = -\sigma^\circ$. Тогда вместо уравнений движения (2.85) получим в прямоугольной (декартовой) системе координат $Ox_1x_2x_3$ соотношения

$$\rho \frac{dv_i}{dt} = b_i - \frac{\partial p}{\partial x_i} = 0, \quad i = 1, 2, 3, \quad (3.1)$$

называемые **уравнениями Эйлера**. Они установлены Л. Эйлером в 1755 году. Их можно представить в векторной форме:

$$\rho \frac{dv}{dt} = b - \nabla p. \quad (3.2)$$

Принимая во внимание, что [VII]

$$\frac{dv}{dt} = \frac{\partial v}{\partial t} + (v \nabla) v = \frac{\partial v}{\partial t} + \frac{1}{2} \nabla v^2 - v \times (\nabla \times v),$$

можно записать (3.2) также в виде

$$\frac{\partial v}{\partial t} + \frac{1}{2} \nabla v^2 + \frac{\nabla p}{\rho} = \frac{b}{\rho} + v \times (\nabla \times v). \quad (3.3)$$

При умеренных давлениях жидкость обычно считают несжимаемой. Если несжимаемая жидкость однородна, т.е. $\rho = \rho^\circ = \text{const}$, то при заданном значении ρ° (3.1) и уравнение неразрывности $\nabla v = 0$ (2.4) образуют замкнутую систему четырех уравнений относительно четырех неизвестных функций v_i , $i = 1, 2, 3$, и p . Несжимаемая жидкость может быть неоднородной, т.е. состоять из частиц, имеющих различную плотность, но сохраняющих ее в процессе движения неизменной (т.е.

$\frac{d\rho}{dt} = 0$). Тогда к указанным четырем следует добавить уравнение $\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0$ (2.2). Оно позволяет найти функцию $\rho = \rho(t, \mathbf{x})$, описывающую распределение плотности неоднородной несжимаемой жидкости в пространстве в различные моменты времени t , если в момент времени $t = 0$, принимаемый за начальный, задано начальное распределение $\rho_0(\mathbf{x}) = \rho(0, \mathbf{x})$.

Пусть при $\rho = \rho^0 = \text{const}$ векторные поля скорости однородной несжимаемой жидкости и объемных сил являются потенциальными, т.е. $\nabla \times \mathbf{v} = \mathbf{0}$, $\mathbf{v} = \nabla \varphi$ и $\mathbf{b} = \nabla B$, где φ и B — действительные функции, зависящие в общем случае от времени и пространственных координат. Тогда (3.3) принимает вид

$$\nabla \left(\frac{\partial \varphi}{\partial t} + \frac{1}{2} (\nabla \varphi)^2 + \frac{p - B}{\rho^0} \right) = \mathbf{0}. \quad (3.4)$$

Отсюда следует **интеграл Коши — Лагранжа**

$$\frac{\partial \varphi}{\partial t} + \frac{1}{2} (\nabla \varphi)^2 + \frac{p - B}{\rho^0} = g(t), \quad (3.5)$$

причем для нахождения функции $g(t)$ достаточно располагать зависимостью от t левой части (3.5) в какой-либо одной точке области движения жидкости. Из (2.4) получаем, что *потенциал скорости* φ удовлетворяет уравнению Лапласа $\nabla^2 \varphi = 0$, т.е. является *гармонической функцией*.

При установившемся движении $\frac{\partial \varphi}{\partial t} = 0$ и (3.5) переходит в известный **интеграл Бернулли***

$$\frac{v^2}{2} + \frac{p - B}{\rho^0} = \text{const}. \quad (3.6)$$

Это уравнение Д. Бернулли установил в 1738 году для течения жидкости в поле силы тяжести, т.е. при $B = -\rho^0 g_0 x_3$, где g_0 —

* Д. Бернулли (1700–1782) — швейцарский математик и механик

ускорение свободного падения, а x_3 — координата, отсчитываемая вертикально вверх.

Если плотность среды зависит только от давления, то процесс в такой среде называют баротропным. Газ называют совершенным, если он подчиняется уравнению состояния

$$p = \rho RT, \quad (3.7)$$

где R — удельная газовая постоянная для данного газа. Движение совершенного газа является баротропным процессом при известной температуре T газа или же при отсутствии теплообмена частиц газа между собой и с внешней средой, т.е. в случае адиабатического течения. Действительно, в этом случае для единицы массы газа

$$c_v dT + p d(1/\rho) = 0, \quad (3.8)$$

где c_v — удельная теплоемкость газа при постоянном объеме. В (3.8) первое слагаемое характеризует изменение внутренней энергии газа, а второе — работу давления при изменении объема $1/\rho$ единицы массы газа. Из (3.7) находим

$$dT = \frac{dp}{\rho R} - \frac{p d\rho}{\rho^2 R}$$

и, подставляя в (3.8), получаем дифференциальное уравнение с разделяющимися переменными

$$\frac{c_v dp}{\rho R} - \left(\frac{c_v}{R} + 1 \right) \frac{p d\rho}{\rho^2} = 0,$$

интегрирование которого приводит к зависимости

$$p = C \rho^\gamma, \quad C = \text{const}, \quad (3.9)$$

где $\gamma = 1 + \frac{R}{c_v} = \frac{c_p}{c_v}$ — показатель адиабаты; c_p — удельная теплоемкость газа при постоянном давлении.

Таким образом, в случае баротропного движения газа зависимость $p = f(\rho)$, уравнения Эйлера (3.1) и уравнение неразрывности (2.2) образуют замкнутую систему относительно пяти неизвестных функций v_1, v_2, v_3, ρ и p . Если векторное поле плотности $\mathbf{b}_\rho = \mathbf{b}/\rho$ объемных сил, приходящихся на единицу массы газа, является потенциальным, т.е. $\mathbf{b}_\rho = \nabla B_\rho$, где B_ρ — действительная функция, зависящая в общем случае от времени и пространственных координат, то при потенциальном движении газа, когда $\mathbf{v} = \nabla\varphi$, можно получить интеграл Коши — Лагранжа (3.5). Для этого введем функцию давления

$$P = \int_{p_0}^p \frac{dp}{\rho},$$

где p_0 — некоторое заданное значение давления, полагая, что плотность газа возрастает с ростом давления, т.е. $\frac{d\rho}{dp} = \frac{1}{f'(\rho)} > 0$. Тогда $\frac{1}{\rho}\nabla p = \nabla P$, так что из (3.3) при $\nabla \times \mathbf{v} = \mathbf{0}$ следует

$$\frac{\partial\varphi}{\partial t} + \frac{1}{2}(\nabla\varphi)^2 + P - B_\rho = h(t), \quad (3.10)$$

причем функцию $h(t)$ можно найти тем же путем, что и функцию $g(t)$ в (3.5).

Дифференциал функции давления можно представить в виде

$$dP = \frac{dp}{\rho} = \frac{a^2 d\rho}{\rho}, \quad a^2 = \frac{dp}{d\rho} = f'(\rho) > 0.$$

Отсюда следует, что

$$\frac{1}{\rho} \frac{d\rho}{dt} = \frac{1}{a^2} \frac{dP}{dt}.$$

Тогда, используя уравнение неразрывности в виде (2.3) и зависимость $\mathbf{v} = \nabla\varphi$, имеем

$$\frac{1}{a^2} \frac{dP}{dt} + \nabla^2\varphi = 0. \quad (3.11)$$

Система (3.10), (3.11) замкнута относительно неизвестных действительных функций P и φ , поскольку a^2 можно также представить как функцию P . Например, в случае адиабатического течения, учитывая (3.9), получаем

$$a^2 = \frac{dp}{d\rho} = \gamma C \rho^{\gamma-1} = \gamma \frac{p}{\rho}, \quad (3.12)$$

а для функции давления

$$P = \int_{\rho_0}^p \frac{dp}{\rho} = \gamma C \int_{\rho_0}^p \rho^{\gamma-2} d\rho = \frac{\gamma C}{\gamma-1} \rho^{\gamma-1} \Big|_{\rho_0}^p = \frac{a^2}{\gamma-1} - \frac{\gamma C}{\gamma-1} \rho_0^{\gamma-1},$$

где $\rho_0 = \left(\frac{p_0}{C}\right)^{1/\gamma}$. Отсюда следует линейная зависимость $a^2 = (\gamma-1)P + C\rho_0^{\gamma-1} = (\gamma-1)P + C^{1/\gamma} p_0^{1-1/\gamma}$.

В ряде прикладных задач аэродинамики и акустики течение газа можно рассматривать как возмущенное относительно известного движения или состояния покоя. Так, при малых возмущениях около состояния покоя систему (3.10), (3.11) можно линеаризовать и привести к виду

$$\frac{\partial \varphi}{\partial t} + P = 0, \quad \frac{1}{a_0^2} \frac{\partial P}{\partial t} + \nabla^2 \varphi = 0,$$

где $a_0^2 = \text{const}$ — значение $\frac{dp}{d\rho}$, вычисленное для невозмущенного состояния. Исключая отсюда P , получаем так называемое волновое уравнение

$$\frac{\partial^2 \varphi}{\partial t^2} = a_0^2 \nabla^2 \varphi, \quad (3.13)$$

которое является дифференциальным уравнением гиперболического типа. Значение a_0 имеет смысл скорости распространения волн в газе и носит название скорости звука. Если в невозмущенном состоянии газ имеет температуру T , то для

адиабатического течения из (3.7) и (3.12) следует известная формула $a_0 = \sqrt{\gamma RT}$. И. Ньютон предполагал, что при распространении звуковых волн течение газа изотермическое, и получил для скорости звука выражение $\sqrt{RT}$. Эта ошибка позднее была исправлена П.С. Лапласом.

При установившихся колебаниях с некоторой частотой ω функцию φ можно представить произведением зависящей только от времени t периодической функции вида $\sin \omega t$ или $\cos \omega t$ (или линейной комбинации этих функций) и функции u , зависящей лишь от пространственных координат и описывающей форму волны. Тогда (3.13) переходит в уравнение эллиптического типа

$$\nabla^2 u + \frac{\omega^2}{a_0^2} u = 0, \quad (3.14)$$

называемое *уравнением Гельмгольца*.

При взаимодействии идеальной жидкости (идеального газа) с непроницаемой границей области течения из-за того, что вектор напряжения коллинеарен вектору $\mathbf{n}$ нормали к границе, возможно проскальзывание частиц, т.е. отсутствует их прилипание к границе. Поэтому на таких границах совпадают проекции векторов скорости жидкости (газа) и заданной скорости $\mathbf{v}^0$ границы на направление нормали, т.е. $\mathbf{v}\mathbf{n} = \mathbf{v}^0\mathbf{n}$, или (в случае потенциального поля скорости) $(\nabla\varphi)\mathbf{n} = \mathbf{v}^0\mathbf{n}$. На свободной границе области должно быть задано давление, а на проницаемых границах — вектор скорости или давление.

3.2. Модели вязкой жидкости

В отличие от идеальной (невязкой) жидкости при движении вязкой жидкости компоненты *тензора напряжений* в выбранной прямоугольной системе координат $Ox_1x_2x_3$ в общем случае могут принимать произвольные значения. Рассмотрим наиболее простую модель изотропной вязкой жидкости, связывающую компоненты σ_{ij} тензора напряжений с компонентами

тензора скоростей деформаций

$$\xi_{ij} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right), \quad i, j = 1, 2, 3, \quad (3.15)$$

линейными соотношениями

$$\sigma_{ij} = -p\delta_{ij} + k(\nabla v)\delta_{ij} + 2\eta\xi_{ij}, \quad i, j = 1, 2, 3, \quad (3.16)$$

где коэффициенты k и η характеризуют вязкое сопротивление при движении среды, а $\delta_{ij} = 1$ при $i = j$ и $\delta_{ij} = 0$ при $i \neq j$ (**тензор второго ранга** с компонентами δ_{ij} называют **единичным**). В этой модели предполагают отсутствие проскальзывания частиц жидкости на границе с твердым телом, т.е. векторы скоростей жидкости и твердого тела в точках такой границы совпадают (в точках неподвижной границы скорость жидкости равна нулю).

Из (3.16) следует, что среднее нормальное напряжение равно

$$\sigma^\circ = \frac{\sigma_{11} + \sigma_{22} + \sigma_{33}}{3} = -p + \left(k + \frac{2}{3}\eta \right) \nabla v. \quad (3.17)$$

Так как ∇v характеризует скорость изменения единицы объема среды в окрестности точки с радиус-вектором x , то $k + \frac{2}{3}\eta$ в (3.17) называют коэффициентом объемной вязкости.

Тензор напряжений, для которого $\sigma^\circ = 0$, называют **девиатором напряжений**. Любой тензор напряжений можно представить суммой **шарового тензора** и девиатора с компонентами $s_{ij} = \sigma_{ij} - \sigma^\circ\delta_{ij}$. Из (3.16) следует, что

$$s_{ij} = 2\eta\tilde{\xi}_{ij}, \quad \tilde{\xi}_{ij} = \xi_{ij} - \frac{1}{3}(\nabla v)\delta_{ij}, \quad i, j = 1, 2, 3, \quad (3.18)$$

где $\tilde{\xi}_{ij}$ являются компонентами **девиатора скоростей деформаций**. Они характеризуют скорость изменения формы частицы среды постоянного объема путем сдвиговых деформаций. Поэтому η называют коэффициентом сдвиговой вязкости.

Если подставить (3.16) в уравнения движения (2.85), то с учетом суммирования по повторяющемуся индексу j получим

$$\rho \frac{dv_i}{dt} = b_i - \frac{\partial p}{\partial x_i} + \frac{\partial k \nabla v}{\partial x_i} + 2 \frac{\partial \eta \xi_{ij}}{\partial x_j}, \quad i, j = 1, 2, 3. \quad (3.19)$$

Отсюда при постоянных k и η с учетом (3.15) приходим к векторной форме **уравнений Навье — Стокса**

$$\rho \frac{d\mathbf{v}}{dt} = \mathbf{b} - \nabla p + (k + \eta) \nabla(\nabla \mathbf{v}) + \eta \nabla^2 \mathbf{v}. \quad (3.20)$$

Уравнения движения вязкой жидкости при упрощенных предположениях вывел в 1822 году французский инженер и механик А. Навье (1785–1836). Эти уравнения в виде, соответствующем (3.20), были получены в 1845 году английским физиком и математиком Дж.Г. Стоксом (1819–1903).

В отличие от идеальной жидкости при контакте вязкой жидкости с твердым телом отсутствует проскальзывание, возникает эффект прилипания. Поэтому при задании граничных условий должны совпадать векторы скорости частиц жидкости и точек такой границы (а не только их проекции на направление нормали к границе).

Для несжимаемой ($\nabla \mathbf{v} = 0$) неоднородной вязкой жидкости с постоянным значением η уравнения неразрывности (2.2), (2.4) и уравнение (3.20) в виде

$$\rho \frac{d\mathbf{v}}{dt} = \mathbf{b} - \nabla p + \eta \nabla^2 \mathbf{v} \quad (3.21)$$

образуют замкнутую систему относительно неизвестных функций $\rho(t, \mathbf{x})$, $\mathbf{v}(t, \mathbf{x})$ и $p(t, \mathbf{x})$. Неоднородность несжимаемой жидкости может быть вызвана зависимостью ее плотности от температуры или от концентрации растворенной в жидкости примеси. В поле тяготения такая неоднородность приводит к изменению векторного поля объемных сил. Если несжимаемая жидкость однородна, т.е. $\rho = \text{const}$, то в указанной системе

уравнения (2.2) и (2.4) равносильны и система (2.4), (3.21) будет замкнута относительно функций $v(t, \mathbf{x})$ и $p(t, \mathbf{x})$. Рассмотрим один из путей модификации *математической модели*, содержащей систему (2.4), (3.21), на сравнительно простом примере двумерного течения.

Пример 3.1. При достаточно медленном движении вязкой жидкости можно пренебречь в (3.21) инерционными силами, приравняв правую часть нулю. Тогда в частном случае течения однородной несжимаемой вязкой жидкости параллельно координатной плоскости $x_1 O x_2$ получим систему уравнений

$$\begin{cases} b_1 - \frac{\partial p}{\partial x_1} + \eta \left(\frac{\partial^2 v_1}{\partial x_1^2} + \frac{\partial^2 v_1}{\partial x_2^2} \right) = 0, \\ b_2 - \frac{\partial p}{\partial x_2} + \eta \left(\frac{\partial^2 v_2}{\partial x_1^2} + \frac{\partial^2 v_2}{\partial x_2^2} \right) = 0, \\ \frac{\partial v_1}{\partial x_1} + \frac{\partial v_2}{\partial x_2} = 0. \end{cases} \quad (3.22)$$

Ясно, что последнее уравнение в (3.22) превращается в тождество, если ввести функцию ψ , удовлетворяющую соотношениям

$$v_1 = \frac{\partial \psi}{\partial x_2}, \quad v_2 = -\frac{\partial \psi}{\partial x_1} \quad (3.23)$$

(ее называют *функцией тока*). Физический смысл этой функции состоит в том, что она постоянна вдоль каждой линии тока, которая при установившемся течении совпадает с траекторией частиц жидкости, а расход жидкости между двумя любыми линиями тока пропорционален разности значений ψ , соответствующих этим линиям.

Функцию $\zeta = \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2}$, характеризующую угловую скорость вращения частиц жидкости при движении в плоскости, называют *завихренностью*. Учитывая (3.23), получаем уравнение

$$\frac{\partial^2 \psi}{\partial x_1^2} + \frac{\partial^2 \psi}{\partial x_2^2} = -\zeta, \quad (3.24)$$

связывающее функцию тока и завихренность. Если продифференцировать первое уравнение в (3.22) по x_2 , а второе — по x_1 , и из второго результата вычесть первый, то получим уравнение Пуассона

$$\frac{\partial b_2}{\partial x_1} - \frac{\partial b_1}{\partial x_2} + \eta \left(\frac{\partial^2 \zeta}{\partial x_1^2} + \frac{\partial^2 \zeta}{\partial x_2^2} \right) = 0, \quad (3.25)$$

не содержащее давления p .

По найденной из (3.25) функции ζ затем из (3.24) можно найти функцию тока ψ и из (3.23) вычислить проекции v_i вектора скорости на координатные оси Ox_1, Ox_2 . Если необходимо найти поле давления, то дифференцированием первого уравнения (3.22) по x_1 , второго — по x_2 и сложением результатов с учетом третьего уравнения (3.22) приходим к уравнению Пуассона

$$\frac{\partial^2 p}{\partial x_1^2} + \frac{\partial^2 p}{\partial x_2^2} = \frac{\partial b_1}{\partial x_1} + \frac{\partial b_2}{\partial x_2}. \quad (3.26)$$

Обычно контур Γ , ограничивающий двумерную область F , в которой рассматривают течение жидкости, содержит участки, соответствующие твердым стенкам. На таких участках не удастся задать в корректной форме граничные условия для завихренности и давления и при решении прикладных задач эти условия обычно получают последовательными приближениями из (3.24) и первых двух уравнений (3.22) соответственно. Отмеченные трудности можно частично преодолеть, если из (3.24) и (3.25) исключить завихренность и получить уравнение четвертого порядка

$$\left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right) \left(\frac{\partial^2 \psi}{\partial x_1^2} + \frac{\partial^2 \psi}{\partial x_2^2} \right) = \frac{1}{\eta} \left(\frac{\partial b_2}{\partial x_1} - \frac{\partial b_1}{\partial x_2} \right) \quad (3.27)$$

относительно функции тока.

Для левой части (3.27) используют запись $\nabla^2(\nabla^2\psi)$ (иногда $\nabla^4\psi$ или $\Delta^2\psi$, если оператор Лапласа обозначают Δ), имея в виду, что в данном случае оператор Лапласа действует в плоскости. При этом оператор $\nabla^2\nabla^2$ называют **бигармоническим**.

Если векторное поле объемных сил является потенциальным, т.е. $\mathbf{b} = \nabla B$, то $\frac{\partial b_2}{\partial x_1} = \frac{\partial b_1}{\partial x_2}$ и (3.27) переходит в **бигармоническое уравнение** $\nabla^2(\nabla^2\psi) = 0$. **Функцию**, удовлетворяющую такому уравнению, называют **бигармонической**.

Несмотря на высокий порядок бигармонического уравнения, математическая модель течения вязкой жидкости на основе функции тока имеет то преимущество, что появляется возможность корректно сформулировать граничные условия. Согласно физическому смыслу функции тока, на непроницаемом для жидкости участке контура ее значение не изменяется, т.е. $\psi = \text{const}$. Если весь контур Γ , ограничивающий односвязную область F (рис. 3.1), является непроницаемым, то на нем можно принять $\psi = 0$. На участке Γ_* (рис. 3.2), через который жидкость вытекает из области F (или поступает в эту область) и на котором задана скорость течения $\mathbf{v}^o(P)$, $P \in \Gamma_*$, нетрудно вычислить изменение функции тока

$$\psi(P) = \psi(A) + \int_0^{s(P)} \mathbf{v}^o(P') \mathbf{n}(P') ds(P'), \quad P \in \Gamma_*,$$

где $\psi(A)$ — значение функции тока в точке $A \in \Gamma_*$, от которой отсчитывают длину $s(P')$ дуги до текущей точки $P' \in \Gamma_*$ с единичным вектором $\mathbf{n}(P')$ внешней нормали.

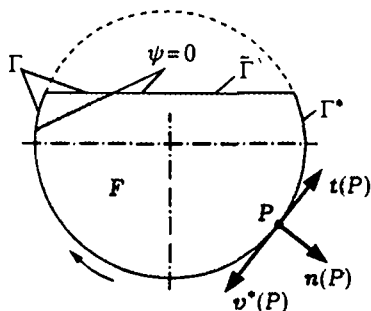


Рис. 3.1

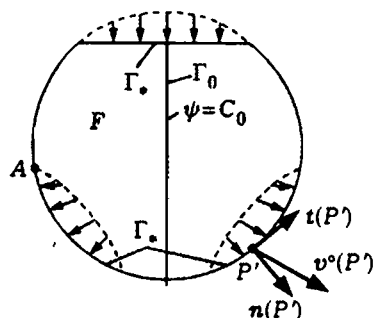


Рис. 3.2

Если в области F до решения задачи можно установить линию Γ_0 симметрии течения (см. рис. 3.2), то она будет совпадать с одной из линий тока, на которой $\psi = C_0 = \text{const}$, а в точках $P \in \Gamma_0$ частицы жидкости не будут вращаться, т.е. $\zeta(P) = 0$, и в соответствии с (3.24) $\nabla^2 \psi(P) = 0$. Совмещая в любой точке $P \in \Gamma_0$ оси координат с направлениями $t(P)$ касательной и $n(P)$ нормали к линии симметрии, в силу инвариантности оператора Лапласа относительно поворота ортогональной системы координат [VII] получаем

$$\frac{\partial^2 \psi(P)}{\partial t^2(P)} + \frac{\partial^2 \psi(P)}{\partial n^2(P)} = 0.$$

Но так как $\psi(P) = \text{const}$, $P \in \Gamma_0$, то

$$\frac{\partial^2 \psi(P)}{\partial t^2(P)} = 0$$

и, следовательно, в качестве граничного условия можно взять

$$\frac{\partial^2 \psi(P)}{\partial n^2(P)} = 0.$$

Аналогичные рассуждения можно провести применительно к свободной поверхности жидкости (участок $\tilde{\Gamma}$ на рис. 3.1), если пренебречь трением жидкости с воздухом (или иным газом) на этой поверхности. Однако на участке Γ^* , соответствующем твердой стенке, происходит прилипание частиц жидкости и в точках $P \in \Gamma^*$ вектор $v(P)$ скорости жидкости равен заданному вектору $v^*(P)$ скорости стенки (например, при вращении сосуда с жидкостью). Поэтому в соответствии с (3.23) имеем на такой стенке

$$\frac{\partial \psi(P)}{\partial n(P)} = v^*(P) t(P) = v_t^*(P), \quad P \in \Gamma^*, \quad (3.28)$$

где $t(P)$ — единичный вектор в направлении касательной к контуру в точке $P \in \Gamma^*$, повернутый относительно единичного вектора $n(P)$ внешней нормали против хода часовой стрелки. Ясно, что в случае неподвижной стенки $v_t^*(P) = 0$.

Таким образом, модель вязкой жидкости позволяет задать в каждой точке контура Γ по два граничных условия для функции тока ψ , необходимых для решения уравнения (3.27) с бигармоническим оператором. Эти граничные условия можно записать в достаточно общем виде следующим образом:

$$\begin{cases} \psi(P) = f_0(P), & \frac{\partial \psi(P)}{\partial n(P)} = f_1(P), & P \in \Gamma_1 \subset \Gamma; \\ \psi(P) = \psi^\circ = \text{const}, & \frac{\partial^2 \psi(P)}{\partial n^2(P)} = 0, & P \in \Gamma_2 = \Gamma \setminus \Gamma_1, \end{cases} \quad (3.29)$$

где $f_0(P)$ и $f_1(P)$ — заданные функции точки на контуре Γ , ограничивающем область F , а число ψ° можно принять равным нулю.

3.3. Упругое твердое тело

Переходя к уравнению состояния твердого тела, ограничимся математической моделью линейно упругой изотропной среды. Эта модель в выбранной прямоугольной системе координат $Ox_1x_2x_3$ связывает компоненты σ_{ij} тензора напряжений с компонентами **тензора деформаций**

$$\varepsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad i, j = 1, 2, 3, \quad (3.30)$$

где u_i — проекции вектора $\mathbf{u}$ перемещения точек тела, линейными соотношениями

$$\sigma_{ij} = \tilde{\lambda} \Theta \delta_{ij} + 2\tilde{\mu} \varepsilon_{ij}, \quad \Theta = \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33} = \varepsilon_{ii} = \nabla \mathbf{u}. \quad (3.31)$$

Здесь Θ — объемная деформация, $\tilde{\lambda}$ и $\tilde{\mu}$ — **константы Ламе***. Соотношения (3.31) выражают так называемый **обобщенный закон Гука****, а (3.30) называют **соотношениями Коши**.

*Г. Ламе (1795–1870) — французский математик и инженер.

**Р. Гук (1635–1703) — английский ученый

Из (3.31) для среднего нормального напряжения находим

$$\sigma^o = \left(\tilde{\lambda} + \frac{2}{3}\tilde{\mu} \right) \Theta = K\Theta, \quad (3.32)$$

а для компонентов девиатора напряжений получаем

$$s_{ij} = 2\tilde{\mu}\tilde{\varepsilon}_{ij}, \quad \tilde{\varepsilon}_{ij} = \varepsilon_{ij} - \frac{1}{3}\Theta\delta_{ij}, \quad i, j = 1, 2, 3. \quad (3.33)$$

Коэффициент $K = \tilde{\lambda} + \frac{2}{3}\tilde{\mu}$ называют модулем объемного сжатия, а $\tilde{\mu}$ — модулем сдвига. В инженерных задачах чаще используют коэффициенты

$$E = \tilde{\mu} \frac{3\tilde{\lambda} + 2\tilde{\mu}}{\tilde{\lambda} + \tilde{\mu}}, \quad \nu = \frac{\tilde{\lambda}}{2(\tilde{\lambda} + \tilde{\mu})}, \quad (3.34)$$

называемые модулем упругости, равным отношению $\sigma_{11}/\varepsilon_{11}$ нормального напряжения σ_{11} при растяжении образца материала тела вдоль оси Ox_1 к соответствующему удлинению ε_{11} , и коэффициентом Пуассона, равным взятому с обратным знаком отношению $\varepsilon_{22}/\varepsilon_{11}$ удлинений также при растяжении вдоль этой оси соответственно.

После подстановки (3.30) в (3.31) и затем в уравнения движения (2.85), полагая линейно упругое тело однородным, т.е. константы Ламе не зависящими от координат, получаем векторную форму **уравнений Ламе**

$$\rho \frac{dv}{dt} = \mathbf{b} + (\tilde{\lambda} + \tilde{\mu}) \nabla(\nabla \mathbf{u}) + \tilde{\mu} \nabla^2 \mathbf{u}, \quad (3.35)$$

причем $\nabla \mathbf{u} = \Theta$. В соотношениях Коши (3.30) компоненты тензора деформаций предполагают малыми по сравнению с единицей. Поэтому изменение плотности ρ твердого тела также мало и при использовании (3.35) **уравнение неразрывности** можно не рассматривать, положив $\rho \approx \rho^o = \text{const}$.

При малых деформациях перемещения, скорости и ускорения частиц тела могут быть, вообще говоря, значительными. Однако на практике часто встречаются задачи, в которых малы и эти величины. Тогда в (3.35) левую часть заменяют на

$\rho^0 \frac{\partial^2 \mathbf{u}}{\partial t^2}$ и получают уравнение относительно единственной неизвестной функции $\mathbf{u} = \mathbf{u}(t, \mathbf{x})$:

$$\frac{\partial^2 \mathbf{u}}{\partial t^2} = \frac{\mathbf{b}}{\rho^0} + \frac{\tilde{\lambda} + \tilde{\mu}}{\rho^0} \nabla(\nabla \mathbf{u}) + \frac{\tilde{\mu}}{\rho^0} \nabla^2 \mathbf{u}. \quad (3.36)$$

Учитывая, что $\nabla^2 \mathbf{u} = \nabla(\nabla \mathbf{u}) - \nabla \times (\nabla \times \mathbf{u})$ [VII], это уравнение можно записать в следующем виде:

$$\frac{\partial^2 \mathbf{u}}{\partial t^2} = \frac{\mathbf{b}}{\rho^0} + \frac{\tilde{\lambda} + 2\tilde{\mu}}{\rho^0} \nabla(\nabla \mathbf{u}) - \frac{\tilde{\mu}}{\rho^0} \nabla \times (\nabla \times \mathbf{u}). \quad (3.37)$$

Помимо (3.36) или (3.37) в формулировку задачи должны входить векторные поля начальных перемещения и скорости, а на границах тела должны быть заданы векторы перемещения или напряжения.

Любое векторное поле в односвязной области можно представить как сумму потенциального и соленоидального векторных полей [VII], т.е.

$$\mathbf{u} = \mathbf{u}^* + \mathbf{u}_* = \nabla \Phi + \nabla \times \mathbf{w}, \quad \frac{\mathbf{b}}{\rho^0} = \nabla \Psi + \nabla \times \mathbf{B}, \quad (3.38)$$

где Φ , Ψ и $\mathbf{w}$, $\mathbf{B}$ — скалярные и векторные потенциалы полей перемещений $\mathbf{u}$ и массовых сил $\mathbf{b}/\rho^0$ соответственно, $\mathbf{u}^* = \nabla \Phi$ и $\mathbf{u}_* = \nabla \times \mathbf{w}$. Подставляя (3.38) в (3.37), получаем

$$\begin{aligned} \frac{\partial^2 \nabla \Phi}{\partial t^2} + \frac{\partial^2 \nabla \times \mathbf{w}}{\partial t^2} &= \nabla \Psi + \nabla \times \mathbf{B} + \\ &+ \frac{\tilde{\lambda} + 2\tilde{\mu}}{\rho^0} \nabla(\nabla^2 \Phi) - \frac{\tilde{\mu}}{\rho^0} \nabla \times (\nabla \times (\nabla \times \mathbf{w})), \end{aligned}$$

или, считая $a_1^2 = (\tilde{\lambda} + 2\tilde{\mu})/\rho^0$ и $a_2^2 = \tilde{\mu}/\rho^0$ константами,

$$\begin{aligned} \nabla \left(\frac{\partial^2 \Phi}{\partial t^2} - a_1^2 \nabla^2 \Phi - \Psi \right) + \\ + \nabla \times \left(\frac{\partial^2 \mathbf{w}}{\partial t^2} - a_2^2 \nabla \times (\nabla \times \mathbf{w}) - \mathbf{B} \right) = \mathbf{0}, \end{aligned} \quad (3.39)$$

где $\mathbf{0}$ — нулевой вектор.

Если теперь к (3.39) применить операцию дивергенции и учесть, что $\nabla(\nabla \times \mathbf{f}) = 0$ для любой дважды непрерывно дифференцируемой векторной функции $\mathbf{f}$, то получим

$$\begin{aligned} \nabla \left(\nabla \left(\frac{\partial^2 \Phi}{\partial t^2} - a_1^2 \nabla^2 \Phi - \Psi \right) \right) = \\ = \nabla \left(\frac{\partial^2 \mathbf{u}^*}{\partial t^2} - a_1^2 \nabla^2 \mathbf{u}^* - \nabla \Psi \right) = \nabla C_1 = 0. \end{aligned}$$

Так как $\nabla \times \mathbf{u}^* = \nabla \times (\nabla \Phi) = 0$, то и $\nabla \times C_1 = 0$. Известно [VII], что если в некоторой области дивергенция и ротор некоторого вектора являются нулевыми, то этот вектор в данной области тождественно постоянен. Обозначая $C_1 = \nabla H$, где H — некоторая гармоническая функция, для которой $\nabla^2 H = 0$, приходим к неоднородному волновому уравнению

$$\frac{\partial^2 \mathbf{u}^*}{\partial t^2} = a_1^2 \nabla^2 \mathbf{u}^* + \nabla(\Psi + H),$$

описывающему распространение со скоростью a_1 волн, сопровождаемых растяжением и сжатием среды, но не вызывающих в силу $\nabla \times \mathbf{u}^* = \nabla \times (\nabla \Phi) = 0$ изменения формы ее элементарного объема.

Если же применить операцию ротора к (3.36) и учесть, что $(\tilde{\lambda} + \tilde{\mu})/\rho^0 = \text{const}$, а $\nabla \times (\nabla f) = 0$ для любой дважды непрерывно дифференцируемой действительной функции f , то будем иметь

$$\begin{aligned} \nabla \times \left(\frac{\partial^2 (\mathbf{u}^* + \mathbf{u}_*)}{\partial t^2} - a_2^2 \nabla^2 (\mathbf{u}^* + \mathbf{u}_*) - \nabla \Psi - \nabla \times \mathbf{B} \right) = \\ = \nabla \times \left(\frac{\partial^2 \mathbf{u}_*}{\partial t^2} - a_2^2 \nabla^2 \mathbf{u}_* - \nabla \times \mathbf{B} \right) = 0. \end{aligned}$$

Выполнение этого равенства можно обеспечить, если положить

$$\frac{\partial^2 \mathbf{u}_*}{\partial t^2} = a_2^2 \nabla^2 \mathbf{u}_* + \nabla \times \mathbf{B}.$$

Таким образом, пришли к неоднородному волновому уравнению, описывающему распространение со скоростью a_2 волн сдвига, не вызывающих в силу $\nabla \mathbf{u}_* = \nabla(\nabla \times \mathbf{w}) = 0$ изменения элементарного объема среды, но изменяющих его форму.

Отметим, что в установившемся состоянии, когда перемещения не зависят от времени, из (3.36) следует векторное **уравнение теории упругости в перемещениях**

$$\nabla^2 \mathbf{u} + \frac{1}{1-2\nu} \nabla \Theta + \frac{\mathbf{b}}{\tilde{\mu}} = 0. \quad (3.40)$$

В проекциях на оси Ox_i оно переходит в уравнения

$$\nabla^2 u_i + \frac{1}{1-2\nu} \frac{\partial \Theta}{\partial x_i} + \frac{b_i}{\tilde{\mu}} = 0, \quad i = 1, 2, 3, \quad (3.41)$$

полученные в 1821 году А. Навье для частного случая $\tilde{\lambda} = \tilde{\mu}$, что соответствует $\nu = 1/4$, а в 1828 году О. Коши для общего случая.

Применяя к (3.40) операцию дивергенции и учитывая, что $\Theta = \nabla \mathbf{u}$, получим

$$\nabla^2 \Theta + \frac{1-2\nu}{2\tilde{\mu}(1-\nu)} \nabla \mathbf{b} = 0.$$

Отсюда следует, что при отсутствии объемных сил объемная деформация и в соответствии с (3.32) среднее напряжение являются гармоническими функциями. В этом случае, используя (3.41), получаем трехмерное *бигармоническое уравнение* $\nabla^2(\nabla^2 u_i) = 0$, $i = 1, 2, 3$, т.е. проекции перемещений являются *бигармоническими функциями*. Так как в соответствии с соотношениями Коши (3.30) и обобщенным законом Гука (3.31) компоненты тензоров деформаций и напряжений являются линейными комбинациями производных перемещений, то эти компоненты при отсутствии объемных сил также будут бигармоническими функциями.

3.4. Уравнение переноса энергии в среде

Рассмотренные выше (см. 3.1–3.3) модели идеальной жидкости, вязкой жидкости (газа) и линейно упругого твердого тела позволяют формулировать задачи математической физики в тех случаях, когда поведение среды не связано существенно с теплофизическими и электромагнитными явлениями. Необходимость учитывать влияние теплофизических явлений при постановке задач математической физики вызвана прежде всего тем, что процессы в реальной среде часто сопровождается преобразование одного вида энергии в другой. Это приводит к изменению температуры, от которой могут существенно зависеть свойства среды. Изучение процессов преобразования энергии требует привлечения сведений из термодинамики. Здесь мы ограничимся лишь рассмотрением уравнения переноса энергии.

Из закона сохранения энергии следует, что скорость изменения полной энергии E любого тела равна подводимой к нему мощности от внешних источников энергии. Пусть тело имеет изменяющийся во времени t объем $V(t)$, ограниченный подвижной поверхностью S (см. рис. 1.1), точки которой перемещаются вместе с частицами тела. Поле скорости частиц тела описывает векторная функция $\mathbf{v} = \mathbf{v}(t, M)$ времени и координат $x_i, i = 1, 2, 3$, точки $M \in \mathbb{R}^3$ в прямоугольной системе координат $Ox_1x_2x_3$ с ортами $\mathbf{e}_i$, а распределение в V плотности среды — функция $\rho = \rho(t, M)$.

Объемную плотность ε энергии среды в $V(t)$ представим как сумму объемных плотностей $\rho \mathbf{v}^2/2$ и ρu кинетической и внутренней энергий соответственно, где $u = u(t, M)$ — внутренняя энергия единицы массы среды. Тогда полная энергия тела будет

$$E = \int_{V(t)} \varepsilon dV = \int_{V(t)} \rho \left(\frac{\mathbf{v}^2}{2} + u \right) dV. \quad (3.42)$$

Заменяя в (2.6) C_m на $v^2/2 + u$, имеем

$$\frac{dE}{dt} = \frac{d}{dt} \int_{V(t)} \rho \left(\frac{v^2}{2} + u \right) dV = \int_{V(t)} \rho \frac{d(v^2/2 + u)}{dt} dV. \quad (3.43)$$

Мощность, подводимая от внешних источников энергии, включает совершаемую в единицу времени работу

$$W = \int_{V(t)} \mathbf{b} \mathbf{v} dV + \int_S \mathbf{p} \mathbf{v} dS \quad (3.44)$$

объемных $\mathbf{b}$ и поверхностных $\mathbf{p}$ сил, поток энергии

$$Q_S = - \int_S \mathbf{q}^{(E)} \mathbf{n} dS, \quad (3.45)$$

передаваемый через поверхность S ($\mathbf{q}^{(E)}$ — вектор плотности потока энергии, $\mathbf{n}$ — единичный вектор внешней нормали к этой поверхности), и количество полной энергии

$$Q_V = \int_{V(t)} I_V^{(E)} dV, \quad (3.46)$$

выделенное в объеме V в единицу времени источниками с объемной мощностью $I_V^{(E)}$. Тогда в силу закона сохранения энергии имеем $\frac{dE}{dt} = W + Q_S + Q_V$, откуда, учитывая (3.43)–(3.46), получаем

$$\begin{aligned} \int_{V(t)} \rho \frac{d(v^2/2 + u)}{dt} dV &= \int_{V(t)} \rho \mathbf{v} \frac{d\mathbf{v}}{dt} dV + \int_{V(t)} \rho \frac{du}{dt} dV = \\ &= \int_{V(t)} \mathbf{b} \mathbf{v} dV + \int_S \mathbf{p} \mathbf{v} dS - \int_S \mathbf{q}^{(E)} \mathbf{n} dS + \int_V I_V^{(E)} dV. \end{aligned} \quad (3.47)$$

Приняв во внимание (2.81), запишем с учетом правила суммирования по повторяющимся индексам $i, j = 1, 2, 3$

$$\mathbf{p}\mathbf{v} = p_i \mathbf{e}_i \mathbf{v} = n_j \sigma_{ji} v_i = \sigma_{ji} v_i \mathbf{e}_j n,$$

где σ_{ij} — компоненты тензора напряжений. Тогда, применяя формулу Остроградского — Гаусса, получаем

$$\begin{aligned} \int_S \mathbf{p}\mathbf{v} dS &= \int_S \sigma_{ji} v_i \mathbf{e}_j n dS = \\ &= \int_{V(t)} \nabla(\sigma_{ji} v_i \mathbf{e}_j) dV = \int_{V(t)} \frac{\partial \sigma_{ji} v_i}{\partial x_j} dV. \end{aligned} \quad (3.48)$$

Используя ту же формулу, имеем

$$\int_S \mathbf{q}^{(E)} n dS = \int_{V(t)} \nabla \mathbf{q}^{(E)} dV. \quad (3.49)$$

Умножив второе уравнение (2.88) скалярно на вектор $\mathbf{v}$ и проинтегрировав результат по объему $V(t)$, запишем

$$\int_{V(t)} \rho \mathbf{v} \frac{d\mathbf{v}}{dt} dV = \int_{V(t)} \mathbf{b}\mathbf{v} dV + \int_{V(t)} \frac{\partial \sigma_{ji}}{\partial x_j} v_i dV. \quad (3.50)$$

Подставляя (3.48)–(3.50) в (3.47), находим

$$\int_{V(t)} \left(\rho \frac{du}{dt} + \nabla \mathbf{q}^{(E)} - I_V^{(E)} - \sigma_{ji} \frac{\partial v_i}{\partial x_j} \right) dV = 0. \quad (3.51)$$

Используя представление (3.15) компонентов ξ_{ij} тензора скоростей деформаций, получаем

$$\frac{\partial v_i}{\partial x_j} = \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) + \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} \right) = \xi_{ij} + \frac{1}{2} \left(\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} \right).$$

Отсюда с учетом правила суммирования по повторяющимся индексам и симметрии тензора напряжений ($\sigma_{ij} = \sigma_{ji}$) находим $\sigma_{ji} \frac{\partial v_i}{\partial x_j} = \sigma_{ij} \xi_{ij}$, поскольку

$$\sigma_{ji} \left(\frac{\partial v_i}{\partial x_j} - \frac{\partial v_j}{\partial x_i} \right) = \sigma_{ji} \frac{\partial v_i}{\partial x_j} - \sigma_{ij} \frac{\partial v_i}{\partial x_j} = 0.$$

Таким образом, вместо (3.51) приходим к интегральной форме закона сохранения энергии в объеме V :

$$\int_{V(t)} \left(\rho \frac{du}{dt} + \nabla \mathbf{q}^{(E)} - I_V^{(E)} - \sigma_{ij} \xi_{ij} \right) dV = 0. \quad (3.52)$$

В случае непрерывности подынтегральной функции в (3.52) рассуждения, аналогичные проведенным при получении (2.2) и (2.3), приводят к локальной форме этого закона, представляющей собой уравнение переноса энергии

$$\rho \frac{du}{dt} = \sigma_{ij} \xi_{ij} - \nabla \mathbf{q}^{(E)} + I_V^{(E)}, \quad i, j = 1, 2, 3. \quad (3.53)$$

Наличие в правой части (3.53) слагаемого $\sigma_{ij} \xi_{ij}$, характеризующего объемную мощность источников энергии, связанных с работой напряжений, заставляет в общем случае рассматривать это уравнение вместе с уравнением переноса количества движения (см. 2.4).

Если в неподвижной среде происходит перенос только тепловой энергии, т.е. $\xi_{ij} = 0$, $\mathbf{q}^{(E)} = \mathbf{q}$, $I_V^{(E)} = I_V^{(q)}$, и массовая плотность внутренней энергии зависит лишь от температуры T , то (3.53) переходит в дифференциальное уравнение (2.53) нестационарной теплопроводности. В более общем случае u может зависеть не только от температуры, $I_V^{(E)}$ может характеризовать объемную мощность выделения не только тепловой энергии, а $\mathbf{q}^{(E)}$ быть суммой векторов плотности потоков различных видов энергии (например, тепловой и энергии электромагнитного излучения). Однако при этом могут возникнуть

распределенные по объему и поверхности моменты (в частности, при взаимодействии среды с электромагнитным полем*), что может нарушить симметрию тензора напряжений (см. 2.4, 3.3) и потребовать уточнения уравнения (3.53).

Вернемся к закону сохранения энергии в виде $\frac{dE}{dt} = W + Q_S + Q_V$ и снова используем (3.44)–(3.46), (3.48) и (3.49). Тогда, заменяя в (2.5) C объемной плотностью ε энергии и учитывая (3.42), получаем

$$\begin{aligned} \frac{dE}{dt} &= \frac{d}{dt} \int_{V(t)} \varepsilon dV = \int_{V(t)} \left(\frac{\partial \varepsilon}{\partial t} + \nabla(\varepsilon \mathbf{v}) \right) dV = \\ &= \int_{V(t)} \mathbf{b} \mathbf{v} dV + \int_S \mathbf{p} \mathbf{v} dS - \int_S \mathbf{q}^{(E)} \mathbf{n} dS + \int_{V(t)} I_E dV = \\ &= \int_{V(t)} (\mathbf{b} \mathbf{v} + \nabla(\sigma_{ji} v_i \mathbf{e}_j) - \nabla \mathbf{q}^{(E)} + I_V^{(E)}) dV, \end{aligned}$$

или

$$\int_V \left(\frac{\partial \varepsilon}{\partial t} + \nabla(\varepsilon \mathbf{v} + \mathbf{q}^{(E)} - \sigma_{ji} v_i \mathbf{e}_j) - I_V^{(E)} + \mathbf{b} \mathbf{v} \right) dV = 0. \quad (3.54)$$

Отсюда следует, что

$$\frac{\partial \varepsilon}{\partial t} + \nabla(\varepsilon \mathbf{v} + \mathbf{q}^{(E)} - \sigma_{ji} v_i \mathbf{e}_j) = I_V^{(E)} - \mathbf{b} \mathbf{v}, \quad i, j = 1, 2, 3. \quad (3.55)$$

Интегральная форма (3.54) закона сохранения энергии остается в силе и в том случае, если множество точек в $V \in \mathbb{R}^3$, в которых подынтегральная функция не является непрерывной, имеет меру Лебега, равную нулю. Пусть это множество образует поверхность S^* , которая делит область V на две подобласти

*См., например: *Можен Ж.*

V_1 и V_2 . Для получения условий на такой *поверхности разрыва* введем в четырехмерном евклидовом арифметическом пространстве $\mathbb{R}^4$ с ортонормированным базисом $e = (e_t, e_1, e_2, e_3)$ вектор $\tilde{e} = \varepsilon v_0 e_t + \varepsilon v + q^{(E)} - \sigma_{j,i} v_i e_j$, где, как и в (2.25), v_0 — произвольная константа, имеющая размерность скорости. Используя векторный оператор (2.26), вместо (3.55) запишем

$$\tilde{\nabla} \tilde{e} = I_V^{(E)} - b v. \quad (3.56)$$

Теперь процедура, аналогичная использованной в 2.2, приведет к следующим условиям на поверхности разрыва S^* :

$$[(\varepsilon v + q^{(E)} - \sigma_{j,i} v_i e_j) n^*] = [\varepsilon] v^* n^* + I_S^{(E)} - p^* v, \quad (3.57)$$

где v^* — вектор скорости перемещения точек поверхности S^* ; n^* — единичный вектор нормали к S^* ; $I_S^{(E)}$ — мощность источников энергии, действующих на поверхности S^* ; p^* — плотность внешних поверхностных сил, которые могут быть приложены на этой поверхности (см. 2.4).

3.5. Уравнения Максвелла

Математические модели, описывающие электромагнитные процессы в *сплошной среде*, используют ряд известных физических понятий и законов.

Электромагнитное поле характеризуют векторными функциями $E = E(t, x)$ и $H = H(t, x)$ напряженности электрического и магнитного полей соответственно, где t — время, x — радиус-вектор точки в прямоугольной системе координат $Ox_1 x_2 x_3$. Для среды, обладающей диэлектрической ε и магнитной μ проницаемостями, вводят также векторы электрического смещения $D = \varepsilon \varepsilon_0 E$ и магнитной индукции $B = \mu \mu_0 H$ ($\varepsilon_0 \approx 8,8542 \cdot 10^{-12} \frac{\text{А} \cdot \text{с}}{\text{В} \cdot \text{м}}$ и $\mu_0 = 4\pi \cdot 10^{-7} \approx 1,2566 \cdot 10^{-6} \frac{\text{В} \cdot \text{с}}{\text{А} \cdot \text{м}}$

электрическая и магнитная постоянные, причем $\frac{1}{\sqrt{\epsilon_0 \mu_0}} = c \approx 2,9979 \cdot 10^8$ м/с — скорость света в вакууме). Электромагнитные процессы в такой среде описывают *уравнения Максвелла**

$$\begin{cases} \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} = \mathbf{0}, & \nabla \cdot \mathbf{B} = 0, \\ \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} = \mathbf{j}^{(e)}, & \nabla \cdot \mathbf{D} = \rho_e, \end{cases} \quad (3.58)$$

где $\mathbf{0}$ — нулевой вектор; $\mathbf{j}^{(e)}$ — вектор *плотности электрического тока*; ρ_e — *объемная плотность электрического заряда*. В (3.58) модули векторов $\mathbf{E}$ и $\mathbf{H}$ измеряют в В/м и А/м соответственно, что предписано системой единиц СИ. Связь между $\mathbf{j}^{(e)}$ и $\mathbf{E}$ для изотропной среды при отсутствии распределенных сторонних источников электродвижущей силы (ЭДС) устанавливает закон Ома в виде (1.14). В случае анизотропной среды, свойства которой зависят от направления, вместо электрической проводимости σ в (1.14) следует использовать *тензор второго ранга* коэффициентов электрической проводимости.

Запись (3.58) предполагает, что среда неподвижна относительно системы координат $Ox_1x_2x_3$, а эта система инерциальна, т.е. неподвижна или движется поступательно с постоянной скоростью. Отметим, что вакууму соответствуют значения $\epsilon = \mu = 1$. Кроме того, в вакууме $\mathbf{j}^{(e)} \equiv \mathbf{0}$, т.е. $\sigma = 0$, и обычно $\rho_e \equiv 0$. Можно показать [XII], что тогда при возникновении электромагнитных колебаний каждая из проекций векторов напряженности электрического и магнитного полей, соответствующая функции φ времени и пространственных координат, будет удовлетворять волновому уравнению вида

$$\frac{\partial^2 \varphi}{\partial t^2} = c^2 \nabla^2 \varphi. \quad (3.59)$$

При установившихся колебаниях с некоторой частотой ω форма волны для каждой из этих проекций, описываемая функцией

*Дж.К. Максвелл (1831–1879) — шотландский физик и математик.

и только пространственных координат, удовлетворяет уравнению Гельмгольца

$$\nabla^2 u + \frac{\omega^2}{c^2} u = 0. \quad (3.60)$$

В среде с электрической проводимостью $\sigma > 0$ колебания будут затухающими, а вместо (3.59) при $\varepsilon \neq 1$ и $\mu \neq 1$ получим так называемое **телеграфное уравнение**

$$\frac{\partial^2 \varphi}{\partial t^2} + \frac{\sigma}{\varepsilon \varepsilon_0} \frac{\partial \varphi}{\partial t} = \frac{c^2}{\varepsilon \mu} \nabla^2 \varphi. \quad (3.61)$$

В частном случае постоянного электрического поля в покоящейся среде, называемого электростатическим, из (3.58) следует система уравнений электростатики

$$\nabla \times \mathbf{E} = 0, \quad \nabla \mathbf{D} = \rho_e. \quad (3.62)$$

В случае постоянного магнитного поля в неподвижной среде из (3.58) имеем систему уравнений магнитостатики

$$\nabla \times \mathbf{H} = \mathbf{j}^{(e)}, \quad \nabla \mathbf{B} = 0. \quad (3.63)$$

Первое уравнение (3.62) есть условие потенциальности электростатического поля, для которого с помощью соотношения $\mathbf{E} = -\nabla U$ можно ввести потенциал U . Из второго уравнения (3.62) при $\mathbf{D} = \varepsilon \varepsilon_0 \mathbf{E}$ получаем

$$\nabla(\nabla \varepsilon U) + \frac{\rho_e}{\varepsilon_0} = 0. \quad (3.64)$$

Если диэлектрическая проницаемость среды $\varepsilon = \text{const}$, то из (3.64) следует уравнение Пуассона

$$\nabla^2 U + \frac{\rho_e}{\varepsilon \varepsilon_0} = 0. \quad (3.65)$$

Так как объемная плотность ρ_e электрических зарядов в электростатическом поле не изменяется во времени, то в среде отсутствует электрический ток. Поэтому в средах с электри-

ческой проводимостью $\sigma > 0$ (например, в металлах) в соответствии с (1.14) должно быть $\mathbf{E} = -\nabla U = \mathbf{0}$, т.е. $U = \text{const}$ и, согласно (3.65), $\rho_e \equiv 0$. Это означает, что если металлическое тело объемом V , ограниченное поверхностью S , имеет электрический заряд Q_e , то он будет сосредоточен на поверхности тела.

Пример 3.2. Пусть в полости заземленного металлического тела с внутренней поверхностью S_0 расположено тело из металла, ограниченное замкнутой поверхностью S_* (рис. 3.3). В области V между этими телами находится среда с диэлектрической проницаемостью $\epsilon(M)$ и объемной плотностью $\rho_e(M)$ электрических зарядов, а внутреннее тело заряжено до потенциала U_* относительно заземленного тела. В этом случае распределение потенциала $U(M)$ описывается уравнением (3.64) с граничными условиями

$$U(P') = U_0 = 0, \quad P' \in S_0; \quad U(P_*) = U_*, \quad P_* \in S_*. \quad (3.66)$$

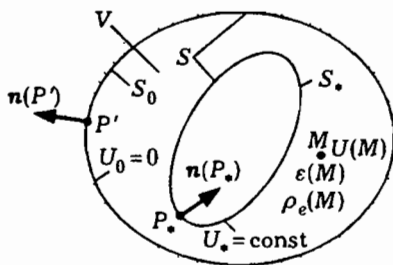


Рис. 3.3

Интегрируя второе уравнение (3.62) по области V , с учетом формулы Остроградского — Гаусса находим

$$\begin{aligned} \int_V \nabla \mathbf{D}(M) dV &= \int_{S_*} \mathbf{D}(P) \mathbf{n}(P) dS + \\ &+ \int_{S_0} \mathbf{D}(P) \mathbf{n}(P) dS = \int_V \rho_e(M) dV, \end{aligned}$$

где $\mathbf{n}(P)$ — единичный вектор внешней (относительно области V) нормали в точках P поверхностей S_* и S_0 . Это соотношение останется верным и в частном случае $\rho_e(M) \equiv 0$. В этом случае получаем равенство

$$\int_{S_*} \mathbf{D}(P) \mathbf{n}(P) dS = - \int_{S_0} \mathbf{D}(P) \mathbf{n}(P) dS. \quad (3.67)$$

Если ввести поверхностную плотность электрического заряда $\rho_S^{(e)}(P) = \mathbf{D}(P) \mathbf{n}(P)$, то (3.67) можно интерпретировать как равенство по абсолютной величине электрического заряда Q , сосредоточенного на поверхности S_* внутреннего тела, и наведенного электрического заряда на поверхности S_0 заземленного тела.

Таким образом, решение краевой задачи (3.64), (3.66) позволяет определить векторную функцию $\mathbf{D}(P) = \varepsilon(P) \varepsilon_0 \mathbf{E}(P) = -\varepsilon(P) \varepsilon_0 \nabla U(P)$, $P \in S = S_* \cup S_0$, электрического смещения и поверхностную плотность $\rho_S^{(e)} = \mathbf{D}(P) \mathbf{n}(P)$ электрического заряда, а затем, используя (3.67), вычислить заряд внутреннего тела

$$Q = \int_{S_*} \mathbf{D}(P) \mathbf{n}(P) dS = \int_{S_*} \rho_S^{(e)}(P) dS. \quad (3.68)$$

При $\rho_e(M) \equiv 0$ рассматриваемую систему характеризуют электрической емкостью $C = Q/U_*$. #

При изучении электромагнитных процессов в поляризующихся и (или) намагничивающихся средах уравнения (3.58) Максвелла рассматривают совместно с соотношениями

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} \quad \text{и} \quad \mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}), \quad (3.69)$$

где $\mathbf{P}$ и $\mathbf{M}$ — векторы электрической поляризованности и намагниченности среды соответственно. В некоторых случаях для сред, называемых диэлектриками, можно принять

$$\mathbf{P} = \varepsilon_0 \chi^{(e)} \mathbf{E}, \quad \mathbf{D} = \varepsilon \varepsilon_0 \mathbf{E}, \quad \varepsilon = 1 + \chi^{(e)}, \quad (3.70)$$

где $\chi^{(e)}$ — электрическая восприимчивость среды. Для сред, относящихся к диамагнетикам и парамагнетикам, можно также использовать линейные соотношения

$$\mathbf{M} = \chi^{(m)} \mathbf{H}, \quad \mathbf{B} = \mu \mu_0 \mathbf{H}, \quad \mu = 1 + \chi^{(m)}, \quad (3.71)$$

где $\chi^{(m)}$ — магнитная восприимчивость. Ясно, что для анизотропных сред $\chi^{(e)}$ и $\chi^{(m)}$, а значит, ϵ и μ будут тензорами второго ранга, а единицу в последних равенствах (3.70) и (3.71) заменит *единичный тензор второго ранга*.

Для некоторых сред зависимость $\mathbf{P}$ от $\mathbf{E}$ и $\mathbf{M}$ от $\mathbf{H}$ нелинейна и даже не является однозначной функцией. Так, для сегнетоэлектриков график зависимости $\mathbf{P}$ от $\mathbf{E}$ имеет вид петли гистерезиса (рис. 3.4). На рисунке стрелками отмечено направление движения точки по кривой при возрастании и убывании модуля $E = |\mathbf{E}|$ с учетом смены направления вектора $\mathbf{E}$ на противоположное, что условно соответствует значениям $E < 0$. При этом с некоторым запаздыванием происходит смена на противоположное направление и вектора $\mathbf{P}$, что также условно соответствует значениям $P < 0$ модуля $P = |\mathbf{P}|$. Аналогичный вид имеет график зависимости $\mathbf{M}$ от $\mathbf{H}$ для ферромагнетиков. На характер указанных зависимостей могут повлиять механические напряжения в среде (так называемые пьезоэлектрический и пьезомагнитный эффекты*, которые широко используют в технических устройствах). В этом случае возникают дополнительные составляющие электрической поляризованности и намагниченности, выражаемые векторами

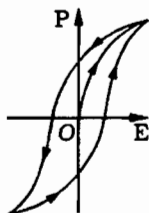


Рис. 3.4

$$\mathbf{P}^{(\sigma)} = \mathbf{d}^{(e)} \tilde{\sigma} \quad \text{и} \quad \mathbf{M}^{(\sigma)} = \mathbf{d}^{(m)} \tilde{\sigma},$$

где $\tilde{\sigma}$ — *тензор напряжений второго ранга*; $\mathbf{d}^{(e)}$ и $\mathbf{d}^{(m)}$ — тензоры третьего ранга пьезоэлектрических и пьезомагнитных коэффициентов.

*См.: Сиротин Ю.И., Шаскольская М.П.

Следует отметить существование и обратных пьезоэлектрического и пьезомагнитного эффектов, когда электрическое и магнитное поля вызывают деформацию среды. В некоторых средах возникает деформация, пропорциональная квадрату напряженности электрического или магнитного поля и неизменяемая при изменении направления вектора напряженности на противоположное. Эти явления называют соответственно электрострикцией и магнестрикцией. Они, в свою очередь, могут повлиять на поляризуемость и намагничивание среды.

Некоторые среды обладают так называемыми прямым и обратным магнитоэлектрическими эффектами: поляризуются под действием магнитного поля и намагничиваются под действием электрического поля. Дополнительные составляющие электрической поляризованности и намагниченности в этом случае можно представить в виде

$$P^{(m)} = GH \quad \text{и} \quad M^{(e)} = G^{-1} E,$$

где G — тензор второго ранга магнитоэлектрических коэффициентов.

Таким образом, математические модели взаимодействия среды с электрическим и магнитным полями достаточно сложны и многообразны. Дополнительное усложнение таких моделей связано с необходимостью учитывать влияние температурного состояния среды на ее свойства. Но при всем разнообразии этих моделей уравнения Максвелла остаются в силе, являются универсальными и сохраняют свое фундаментальное значение. Отметим, что если к третьему уравнению (3.58) применить операцию дивергенции и использовать четвертое уравнение (3.58), то придем к локальной форме закона сохранения электрического заряда (2.46) при условии, что в среде отсутствуют источники заряда ($I_e \equiv 0$).

Силы воздействия электромагнитного поля на среду называют пондеромоторными. На неподвижную относительно выбранной инерциальной системы координат среду при отсут-

ствии электрической поляризации и намагниченности ($\chi^{(e)} = \chi^{(m)} = 0$, т.е. $\varepsilon = \mu = 1$) действует так называемая сила Лоренца, вектор объемной плотности которой равен

$$\mathbf{b}^{(e)} = \rho_e \mathbf{E} + \mu_0 \mathbf{j}^{(e)} \times \mathbf{H}. \quad (3.72)$$

Для электрически поляризованной и намагниченной неподвижной среды с учетом правила суммирования по повторяющемуся индексу имеем*

$$\mathbf{b}^{(e)} = \rho_e \mathbf{E} + \mathbf{j}^{(e)} \times \mathbf{B} + \\ + \frac{1}{2} (P_i \nabla E_i - E_i \nabla P_i + \mu_0 M_i \nabla H_i - \mu_0 H_i \nabla M_i), \quad (3.73)$$

где индекс $i = 1, 2, 3$ указывает номер координатной оси Ox_i , на которую спроектирован соответствующий вектор. Если для среды справедливы (3.70) и (3.71), причем электрическая и магнитная восприимчивости являются скалярами и не зависят от координат, то выражение в круглых скобках в (3.73) равно нулю. Это выражение может быть отлично от нуля для анизотропной среды или в случае, когда электрическая и магнитная восприимчивости зависят от координат.

Ясно, что при взаимодействии среды с электромагнитным полем ponderomotorные силы должны быть учтены в *уравнениях движения* или *равновесия* (см. 2.4). Более того, в общем случае при таком взаимодействии возникают распределенные в объеме среды моменты (пары сил), что приводит к утрате свойства симметрии тензором напряжений**.

Известно, что объемная плотность энергии электромагнитного поля в вакууме равна $w = (\varepsilon_0 E^2 + \mu_0 H^2)/2$. Направление и интенсивность переноса этой энергии характеризуют вектором $\mathbf{S} = \mathbf{E} \times \mathbf{H}$ Умова — Пойнтинга, введенном в работах русского физика Н.А. Умова (1846–1915) и английского физика

*См.: Седов Л.И.

**См. Можен Ж

Дж. Г. Пойнтинга (1852–1914). Если первое и третье уравнения (3.58) скалярно умножить на $\mathbf{H}$ и $\mathbf{E}$ соответственно и затем, приняв для вакуума $\varepsilon = \mu = 1$ и $\mathbf{j}^{(e)} = \mathbf{0}$, из первого результата вычесть второй, то получим

$$\mathbf{H}(\nabla \times \mathbf{E}) - \mathbf{E}(\nabla \times \mathbf{H}) = -\mu_0 \mathbf{H} \frac{\partial \mathbf{H}}{\partial t} - \varepsilon_0 \mathbf{E} \frac{\partial \mathbf{E}}{\partial t}.$$

Отсюда в силу $\mathbf{H}(\nabla \times \mathbf{E}) - \mathbf{E}(\nabla \times \mathbf{H}) = \nabla(\mathbf{E} \times \mathbf{H})$ [VII] приходим к уравнению Умова — Пойнтинга

$$\frac{\partial w}{\partial t} + \nabla S = 0, \quad (3.74)$$

которое является локальной формой закона сохранения энергии электромагнитного поля в вакууме.

Аналогично из (3.58) следует уравнение Умова — Пойнтинга

$$\frac{\partial \tilde{w}}{\partial t} + \nabla S + \mathbf{j}^{(e)} \mathbf{E} = 0, \quad \tilde{w} = \mathbf{E} \frac{\partial \mathbf{D}}{\partial t} + \mathbf{H} \frac{\partial \mathbf{B}}{\partial t}, \quad (3.75)$$

для электрически поляризуемой и намагничиваемой среды. Слагаемое $\mathbf{j}^{(e)} \mathbf{E}$ характеризует для неподвижной среды так называемое джоулево тепло, т.е. объемную мощность $I_q^{(e)}$ источников тепловой энергии, передаваемой среде при ее взаимодействии с электромагнитным полем. Эти источники должны быть учтены в уравнении (3.53) переноса энергии.

В среду также поступает энергия, затрачиваемая на электрическую поляризацию и намагничивание. Объемная мощность источников этой энергии равна

$$\frac{\partial(\tilde{w} - w)}{\partial t} = \mathbf{E} \left(\frac{\partial \mathbf{D}}{\partial t} - \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) + \mathbf{H} \left(\frac{\partial \mathbf{B}}{\partial t} - \mu_0 \frac{\partial \mathbf{H}}{\partial t} \right) = I_P + I_M,$$

где с учетом (3.69)

$$I_P = \mathbf{E} \frac{\partial \mathbf{P}}{\partial t} \quad \text{и} \quad I_M = \mu_0 \mathbf{H} \frac{\partial \mathbf{M}}{\partial t}. \quad (3.76)$$

Если зависимость P от E (или M от H) имеет вид петли гистерезиса (см. рис. 3.4), то часть затрачиваемой на процессы поляризации (или намагничивания) энергии электромагнитного поля может переходить в теплоту, что также следует учитывать в уравнении (3.53) переноса энергии.

3.6. Электромагнитные процессы в медленно движущейся среде

Выше (см. 3.5) приведены уравнения (3.58) Максвелла для электрически поляризованной и намагниченной среды, неподвижной относительно некоторой инерциальной системы координат. Их запись будет неизменна (инвариантна), если при переходе к другой инерциальной системе координат, движущейся относительно прежней поступательно с постоянной скоростью, применить преобразование, предложенное в 1904 году нидерландским физиком и математиком Х.А. Лоренцом (1853–1928).

Пусть оси системы координат $O'x'_1x'_2x'_3$ параллельны соответствующим осям системы координат $Ox_1x_2x_3$ и точка O' движется относительно точки O вдоль оси Ox_1 с постоянной скоростью v в положительном направлении этой оси (рис. 3.5). Тогда при преобразовании Лоренца время и координаты в этих системах будут связаны соотношениями

$$t' = \frac{t - vx_1/c^2}{\sqrt{1 - v^2/c^2}}, \quad x'_1 = \frac{x_1 - vt}{\sqrt{1 - v^2/c^2}}, \quad x'_2 = x_2, \quad x'_3 = x_3,$$

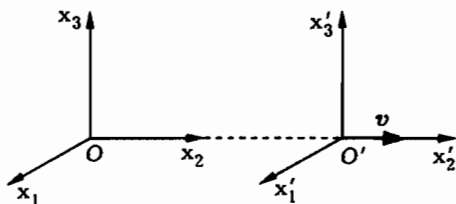


Рис. 3.5

где c — скорость распространения электромагнитных волн в вакууме. Эти соотношения отличаются от формул преобразования Галилея*

$$t' = t, \quad x'_1 = x - vt, \quad x'_2 = x_2, \quad x'_3 = x_3,$$

относительно которого инвариантны уравнения классической (нерелятивистской) механики.

При сравнительно медленном движении среды можно пренебречь значением v^2/c^2 по сравнению с единицей и при преобразовании системы координат использовать приближенные формулы. В этом случае (3.58) будут инвариантны относительно преобразований Галилея, если в новой инерциальной системе координат, движущейся относительно прежней со скоростью V , принять**

$$\begin{cases} E' = E + V \times B, & B' = B - \frac{V \times E}{c^2}, & j^{(e)'} = j^{(e)} - V \rho_e, \\ D' = D + \varepsilon_0 V \times B, & H' = H - V \times D, & \rho_e' = \rho_e, \end{cases} \quad (3.77)$$

где величины со штрихом относятся к новой системе координат, а соответствующие им величины без штриха — к прежней системе.

Г. Минковский*** сформулировал принцип, позволяющий использовать (3.77) при произвольном поле скорости среды, задаваемом в прежней системе координат векторной функцией $v = v(t, x)$, если в (3.77) V заменить на v . Тогда в рассматриваемом приближении уравнения (3.58) будут справедливы для элементарного объема движущейся среды в сопутствующей системе координат, движущейся вместе с этим объемом.

В частном случае, когда $|H| \ll |v||D|$ и $|B| \ll |v||E|$, в (3.77) можно принять $E' = E$ и $D' = D$, а (3.58) привести к виду

$$\nabla \times E = 0, \quad \nabla D = \rho_e, \quad \nabla \times H = \frac{\partial D}{\partial t} + j^{(e)}, \quad \nabla B = 0.$$

*Г. Галилэй (1564–1642) — итальянский физик, механик и математик.

**См. Можен Ж.

***Г. Минковский (1864–1909) — немецкий математик и физик.

В этом приближении изменение электрического поля порождает магнитное поле, но обратным эффектом (электромагнитной индукцией) пренебрегают. С учетом (3.77) в прежней системе координат для движущейся среды имеем

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}, \quad \mathbf{B} = \mu_0 (\mathbf{H} + 4\pi \mathbf{M}), \quad \mathbf{j}^{(e)} = \sigma \mathbf{E} + \rho_e \mathbf{v},$$

где σ — электрическая проводимость среды.

В случае, когда $|\mathbf{E}| \ll |\mathbf{v}| |\mathbf{B}|$ и $|\mathbf{v}| |\mathbf{D}| \ll |\mathbf{H}|$, в (3.77) допустимо принять $\mathbf{H}' = \mathbf{H}$ и $\mathbf{B}' = \mathbf{B}$, а (3.58) записать в виде

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \mathbf{D} = \rho_e, \quad \nabla \times \mathbf{H} = \mathbf{j}^{(e)}, \quad \nabla \mathbf{B} = 0.$$

Здесь учитывается эффект электромагнитной индукции, порождающий электрическое поле при изменении магнитного поля, но пренебрегается обратным (магнитоэлектрическим) эффектом. Принимая во внимание (3.77), получаем

$$\begin{cases} \mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P}, \\ \mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}), \\ \mathbf{j}^{(e)} = \sigma (\mathbf{E} + \mathbf{v} \times \mathbf{B}) + \rho_e \mathbf{v}. \end{cases} \quad (3.78)$$

Пример 3.3. В математической модели магнитной гидродинамики обычно принимают*, что в движущейся проводящей среде отсутствуют поляризация и намагниченность, т.е. $\mathbf{D} = \varepsilon_0 \mathbf{E}$ и $\mathbf{B} = \mu_0 \mathbf{H}$, но может протекать электрический ток ($\mathbf{j}^{(e)} \neq 0$), причем $|\mathbf{v}| |\mathbf{D}| \ll |\mathbf{H}|$, так что приближенно $\mathbf{H}' = \mathbf{H}$. Будем считать, что среда идеальная (невязкая) и имеет плотность ρ . Тогда уравнение неразрывности имеет вид (2.2)

$$\frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{v}) = 0, \quad (3.79)$$

*См.. Седов Л.И.

а вместо векторной формы (3.2) уравнений Эйлера получим

$$\rho \frac{dv}{dt} + \nabla p = \mathbf{b} + \mathbf{b}^{(e)}, \quad (3.80)$$

где p — давление, и в соответствии с (3.72) и (3.77) (с заменой $\mathbf{V}$ на $\mathbf{v}$) в инерциальной системе координат вектор объемной плотности силы Лоренца

$$\mathbf{b}^{(e)} = \rho_e (\mathbf{E} + \mathbf{v} \times \mathbf{B}) + \mathbf{j}^{(e)} \times \mathbf{B},$$

причем вектор $\mathbf{j}^{(e)}$ плотности электрического тока задан последним равенством в (3.78). В уравнении (3.53) переноса энергии мощность объемных источников энергии, обусловленных выделением джоулева тепла (см. 3.5), равна

$$I_E = \mathbf{j}^{(e)} (\mathbf{E} + \mathbf{v} \times \mathbf{B}).$$

К указанным соотношениям необходимо добавить уравнения Максвелла (3.58), записанные с учетом приближенного равенства $\mathbf{H}' = \mathbf{H}$ в сопутствующей системе координат, движущейся с элементарным объемом среды:

$$\begin{cases} \nabla \times \mathbf{E}' + \mu_0 \frac{\partial \mathbf{H}}{\partial t} = \mathbf{0}, & \nabla \mathbf{H} = 0, \\ \nabla \times \mathbf{H} - \varepsilon_0 \frac{\partial \mathbf{E}'}{\partial t} = \mathbf{j}^{(e)}, & \nabla \mathbf{E}' = \rho_e. \end{cases} \quad (3.81)$$

В частном случае бесконечно большой электрической проводимости ($\sigma \rightarrow \infty$) среды математическую модель магнитной гидродинамики удастся упростить. Так как плотность электрического тока должна быть конечной, из последнего равенства (3.78) следует, что $\mathbf{E} = -\mathbf{v} \times \mathbf{B} = -\mu_0 \mathbf{v} \times \mathbf{H}$, т.е. напряженность электрического поля в идеальном проводнике в сопутствующей системе координат равна нулю ($\mathbf{E}' \equiv 0$). Тогда из третьего и четвертого уравнений (3.81) имеем

$$\mathbf{j}^{(e)} = \nabla \times \mathbf{H}, \quad \rho_e = 0, \quad (3.82)$$

а первое и второе уравнения (3.81) дают

$$\frac{\partial \mathbf{H}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{H}), \quad \nabla \mathbf{H} = 0. \quad (3.83)$$

Ясно, что для идеального проводника $I_E = 0$, а в (3.80) $\mathbf{b}^{(e)} = (\nabla \times \mathbf{H}) \times \mathbf{B} = \mu_0(\nabla \times \mathbf{H}) \times \mathbf{H}$.

Используя формулы векторного анализа [VII], с учетом второго уравнения (3.83) находим

$$\begin{aligned} \nabla \times (\mathbf{v} \times \mathbf{H}) &= (\mathbf{H} \nabla) \mathbf{v} - (\mathbf{v} \nabla) \mathbf{H} + \mathbf{v}(\nabla \mathbf{H}) - \mathbf{H}(\nabla \mathbf{v}) = \\ &= (\mathbf{H} \nabla) \mathbf{v} - (\mathbf{v} \nabla) \mathbf{H} - \mathbf{H}(\nabla \mathbf{v}), \\ (\nabla \times \mathbf{H}) \times \mathbf{H} &= (\mathbf{H} \nabla) \mathbf{H} - \frac{1}{2} \nabla H^2. \end{aligned}$$

Тогда из (3.80) и первого уравнения (3.83) получим два векторных уравнения

$$\begin{cases} \rho \frac{d\mathbf{v}}{dt} + \nabla p = \mathbf{b} + \mu_0(\mathbf{H} \nabla) \mathbf{H} - \frac{\mu_0}{2} \nabla H^2, \\ \frac{\partial \mathbf{H}}{\partial t} = (\mathbf{H} \nabla) \mathbf{v} - (\mathbf{v} \nabla) \mathbf{H} - \mathbf{H}(\nabla \mathbf{v}) \end{cases} \quad (3.84)$$

относительно двух векторных функций $\mathbf{v}$ и $\mathbf{H}$ и двух действительных функций ρ и p . Вместе с (3.79) и уравнением состояния $p = p(\rho)$ уравнения (3.84) составляют замкнутую систему.

Предположим, что все неизвестные функции зависят лишь от одной пространственной координаты x_1 , причем вектор $\mathbf{v}$ скорости среды имеет лишь одну ненулевую проекцию v на координатную ось Ox_1 , а $\mathbf{b} = 0$. Применяя к первому уравнению (3.83) операцию дивергенции и учитывая, что $\nabla(\nabla \times \mathbf{f}) = 0$ для любой дважды дифференцируемой векторной функции $\mathbf{f}$, имеем $\frac{\partial(\nabla \mathbf{H})}{\partial t} = 0$. Из второго уравнения (3.83) следует, что $\frac{\partial H_1}{\partial x_1} = 0$, где H_1 — проекция вектора $\mathbf{H}$ на ось Ox_1 .

Следовательно, $\frac{\partial^2 H_1}{\partial t \partial x_1} = 0$, т.е. $H_1 = \text{const}$. Примем $H_1 = 0$, а

проекцию вектора $\mathbf{H}$ на направление, перпендикулярное оси Ox_1 , обозначим H . Тогда $(\mathbf{H}\nabla)\mathbf{H} = \mathbf{0}$ и $(\mathbf{H}\nabla)v = 0$, а вместо (3.79) и (3.84) с учетом уравнения состояния $p = p(\rho)$ получим систему трех уравнений

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial x_1} = 0, \quad \rho \frac{dv}{dt} + \frac{\partial}{\partial x_1} \left(p(\rho) + \frac{\mu_0}{2} H^2 \right) = 0, \quad \frac{\partial H}{\partial t} + \frac{\partial v H}{\partial x_1} = 0$$

относительно трех искомых величин ρ , v и H .

Дополнение 3.1. Поверхности разрыва в электромагнитном поле

Используя формулу Остроградского - Гаусса и теорему Стокса, уравнения Максвелла в некоторой инерциальной системе координат можно представить в интегральной форме:

$$\left\{ \begin{array}{l} \int_{S_0} D_n dS = \int_{V_0} \rho_e dV, \\ \int_{V_0} \nabla B dV = \int_{S_0} B_n dS = 0, \\ \oint_{L_0} \mathbf{E} d\mathbf{x} = - \int_{S_L} \frac{\partial B_n}{\partial t} dS, \\ \oint_{L_0} \mathbf{H} d\mathbf{x} = \int_{S_L} \frac{\partial D_n}{\partial t} dS + \int_{S_L} j_n^{(e)} dS, \end{array} \right. \quad (3.85)$$

где L_0 и S_L — замкнутый контур и натянутая на него поверхность; V_0 и S_0 — объем и ограничивающая его поверхность (все они неподвижны относительно выбранной системы координат); $\mathbf{x}$ — радиус-вектор точки $M \in L_0$ в этой системе координат; $\mathbf{E}$ и $\mathbf{H}$ — векторы напряженности электрического и магнитного полей соответственно; ρ_e — объемная плотность электрического заряда. Нижний индекс n обозначает проекцию на направление единичного вектора $\mathbf{n}$ внешней нормали к S или нормали

к S_L векторов $\mathbf{D}$, $\mathbf{B}$ и $\mathbf{j}^{(e)}$ электрического смещения, магнитной индукции и плотности тока соответственно (нормаль к S_L выбрана так, что наблюдаемый с ее стороны обход контура L осуществляется против хода часовой стрелки). К уравнениям (3.85) следует добавить закон сохранения заряда в интегральной форме (см. 2.3)

$$\int_{S_0} \mathbf{j}_n^{(e)} dS = \int_{V_0} \left(I_V^{(e)} - \frac{\partial \rho^{(e)}}{\partial t} \right) dV, \quad (3.86)$$

где $I_V^{(e)}$ — интенсивность объемных источников электрического заряда.

Так как (3.58) справедливы для среды, неподвижной относительно выбранной системы координат, то применение (3.85) также ограничено этим случаем. Пусть $M^* \in S^* \subset V_0$ — некоторая точка на поверхности S^* сильного разрыва относительно функций, входящих в (3.85)–(3.86). Условие (2.48) на поверхности разрыва, которое следует из закона сохранения заряда в движущейся среде, установлено выше (см. 2.3). В данном случае для среды, неподвижной относительно сопутствующей системы координат, с учетом равенства $\mathbf{j}^{(e)} = \sigma \mathbf{E}$ вместо (2.48) получаем

$$[\mathbf{j}^{(e)} \mathbf{n}^*] + \frac{\partial \rho_S^{(e)}}{\partial t} = 0 \text{ на } S^*, \quad (3.87)$$

где $\rho_S^{(e)}$ — поверхностная плотность зарядов, расположенных на S^* в окрестности точки $M^* \in S^*$, а символ $[\cdot]$ обозначает скачок значений функции при переходе в точке $M^* \in S^*$ через поверхность S^* в направлении, противоположном единичному вектору $\mathbf{n}^*(M^*)$ к этой поверхности. Аналогично, используя рассмотренную выше (см. 2.2) процедуру, из (3.85) в сопутствующей для точки $M^* \in S^*$ системе координат находим

$$[\mathbf{D} \mathbf{n}^*] = \rho_S^{(e)}, \quad [\mathbf{B} \mathbf{n}^*] = 0. \quad (3.88)$$

Проведем к поверхности S^* касательную в точке $M^* \in S^*$ в направлении единичного вектора t^* и выберем в качестве поверхности S_L участок плоскости, содержащий векторы n^* и t^* и ограниченной прямоугольным контуром L_0 протяженностью $2h$ в направлении единичного вектора n^* (рис. 3.6). Обход этого контура против хода часовой стрелки будет соответствовать условию, что векторы n^* , t^* и единичный вектор n нормали к S_L образуют *правую тройку некопланарных векторов*, т.е. $n = n^* \times t^*$.

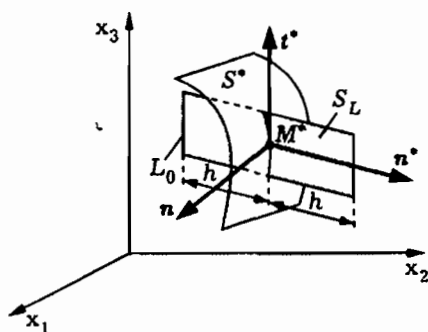


Рис. 3.6

При $h \rightarrow 0$ интегралы по S_L в (3.85) также стремятся к нулю, и при последующем стягивании контура L_0 в точку M^* для произвольного направления касательного вектора t^* , перпендикулярного вектору n^* в точке $M^* \in S^*$, при отсутствии поверхностных токов получаем

$$[Et^*] = 0, \quad [Ht^*] = 0. \quad (3.89)$$

Еще раз подчеркнем, что условия (3.87)–(3.89) справедливы в случае неподвижной относительно сопутствующей для точки $M^* \in S^*$ системы координат. При сравнительно медленном относительно этой системы координат движении среды векторы, входящие в (3.87)–(3.89), можно преобразовать в соответствии с (3.77). Если векторы скорости среды на обеих сторонах поверхности S^* перпендикулярны S^* в точке $M^* \in S^*$, то из (3.77)

следует, что условия (3.88) не изменятся, а если эти векторы касательны к S^* в этой точке, то не изменятся условия (3.87) и (3.89).

Установленные условия на поверхности разрыва дают полезную информацию для корректной формулировки граничных условий при решении задач математической физики в областях непрерывного изменения искомых функций и их производных. В частности, из (3.87)–(3.89) следует, что при решении уравнений Максвелла в дифференциальной форме (3.58) в сочетании с локальной формой (2.46) закона сохранения электрического заряда можно задать на границе области проекции векторов $\mathbf{D}$, $\mathbf{B}$, $\mathbf{j}^{(e)}$ на нормаль к этой границе и проекции векторов $\mathbf{E}$ и $\mathbf{H}$ на направление касательной к ней.

Дополнение 3.2. Примеры задач, описываемых интегральными уравнениями

Математические модели ряда физических процессов могут содержать не только дифференциальные уравнения с частными производными, но и **интегральные** и **интегро-дифференциальные уравнения**, в которых искомые функции (а иногда и их производные) входят и под знак интеграла. Характерным примером интегрального уравнения является **интегральная формула Грина**, содержащая **объемный потенциал** и **потенциалы простого и двойного слоя**, которой удовлетворяет решение задачи для дифференциального уравнения Пуассона [XII]. В некоторых случаях математическая формулировка задачи в виде интегрального уравнения оказывается более простой и может быстрее привести к цели, нежели использование соответствующих дифференциальных уравнений.

Пример 3.4. Вернемся к системе металлических тел, рассмотренной в примере 3.2. Возможен иной подход к построению *математической модели* такой системы.

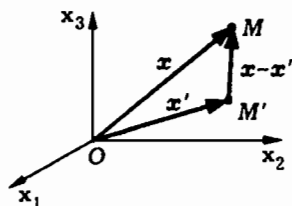


Рис. 3.7

Из физики известно, что, согласно закону Кулона*, неподвижный точечный электрический заряд q , находящийся в точке $M' \in \mathbb{R}^3$ с радиус-вектором x' , определяющим положение этой точки относительно начала прямоугольной системы координат $Ox_1x_2x_3$ (рис. 3.7), создает в

точке $M \in \mathbb{R}^3$ с радиус-вектором x электростатическое поле с вектором напряженности

$$E(x) = \frac{q}{4\pi\epsilon\epsilon_0} \frac{x - x'}{|x - x'|^3}$$

и потенциалом

$$U(x) = \frac{q}{4\pi\epsilon\epsilon_0} \frac{1}{|x - x'|},$$

где ϵ — диэлектрическая проницаемость среды; ϵ_0 — электрическая постоянная (см. 3.5). Заменяем q зарядом $\rho_e(x') dV(x')$ в элементарном объеме $dV(x')$ в окрестности точки с радиус-вектором x' , находящейся в некоторой области V , ограниченной замкнутой поверхностью S , или зарядом $\rho_S^{(e)}(x') dS(x')$ на элементарной площадке $dS(x')$, где ρ_e и $\rho_S^{(e)}$ — объемная и поверхностная плотности электрического заряда. Тогда, суммируя действие таких зарядов и полагая $\epsilon = \text{const}$, получаем

$$U(x) = \frac{1}{4\pi\epsilon\epsilon_0} \int_V \frac{\rho_e(x') dV(x')}{|x - x'|} + \frac{1}{4\pi\epsilon\epsilon_0} \int_S \frac{\rho_S^{(e)}(x') dS(x')}{|x - x'|}. \quad (3.90)$$

В соответствии с граничными условиями (3.66) значения U в точках $P_* \in S_*$ и $P \in S_0$ поверхности $S = S_* \cup S_0$, ограничивающей область V , заключенную между металлическими телами (см. рис. 3.3), заданы. Приравнивая этим значениям левую часть (3.90), приходим к интегральному уравнению

*Ш.О.Кулон (1736–1806) — французский инженер и физик.

относительно неизвестной поверхностной плотности $\rho_S^{(e)}$ электрического заряда в точках поверхности S . После решения этого уравнения из (3.90) можно найти потенциал $U(\mathbf{x})$ электростатического поля в любой точке $M \in V$ с радиус-вектором $\mathbf{x}$, а при помощи (3.68) вычислить электрический заряд Q внутреннего металлического тела. #

Пример 3.4 иллюстрирует возможность свести математическую формулировку задачи, содержащую одно или несколько дифференциальных уравнений, к интегральному уравнению. Приведем примеры задач, интегральная формулировка которых следует непосредственно из их физической постановки.

Пример 3.5. Рассмотрим замкнутую тонкостенную оболочку с поверхностью S (рис. 3.8), в полости которой находится диатермичная среда, т.е. среда, полностью прозрачная для теплового излучения. Извне оболочка поглощает тепловой поток с заданной плотностью $\tilde{q}(M)$, $M \in S$, а внутри оболочки происходит лучистый теплообмен между ее отдельными участками. Построим математическую модель такого процесса переноса тепловой энергии.

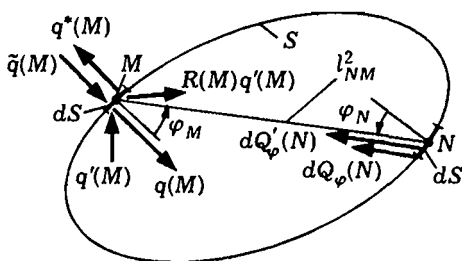


Рис. 3.8

Выделим в окрестности точки $M \in S$ участок оболочки с площадью $dS(M)$ внутренней поверхности и составим для этого участка уравнение баланса тепловой энергии. Примем, что по толщине оболочки температура $T(M)$ однородна. Согласно за-

кону Стефана — Больцмана* теплового излучения, внутренней и внешней поверхности рассматриваемого участка излучают тепловые потоки плотностью $q(M) = \varepsilon(M)\sigma_0 T^4(M)$ и $q^*(M) = \varepsilon^*(M)\sigma_0 T^4(M)$ соответственно, где $\varepsilon(M)$ и $\varepsilon^*(M)$ — коэффициенты излучения этих поверхностей, а $\sigma_0 = 5,67 \cdot 10^{-8} \frac{\text{Вт}}{\text{м}^2 \cdot \text{К}^4}$ — постоянная Стефана — Больцмана. Плотность теплового потока, проходящего через диатермичную среду и падающего на этот участок с внутренней поверхности оболочки, обозначим $q'(M)$. Этот поток частично поглощается, а частично отражается, т.е. $q'(M) = A(M)q'(M) + R(M)q'(M)$, где $A(M)$ и $R(M)$ — коэффициенты поглощения и отражения внутренней поверхности данного участка.

В итоге для рассматриваемого участка (если не учитывать перенос тепловой энергии между соседними участками тонкостенной оболочки путем теплопроводности) получаем $\tilde{q}(M) + A(M)q'(M) = q^*(M) + q(M)$, или

$$\tilde{q}(M) + A(M)q'(M) = (\varepsilon(M) + \varepsilon^*(M))\sigma_0 T^4(M). \quad (3.91)$$

В (3.91) входят две неизвестные функции положения точки $M \in S$: температура $T(M)$ этого участка и плотность $q'(M)$ падающего на него теплового потока. Это означает, что постановка задачи пока не замкнута и необходимо использовать дополнительные соотношения.

Примем, что распределение по направлениям излучения, отраженного от внутренней поверхности оболочки, является рассеянным (диффузным) и подчиняется закону Ламберта**

$$dQ'_\varphi(M) = \frac{R(M)q'(M)}{\pi} \cos \varphi_M dS(M) d\Omega, \quad M \in S, \quad (3.92)$$

где $dQ'_\varphi(M)$ — проходящий через телесный угол $d\Omega$ поток излучения, отраженного от площадки $dS(M)$ под углом φ_M к направлению внутренней нормали к оболочке в точке $M \in S$ (рис. 3.9).

*Й. Стéфан (1835–1893) и Л. Бóльцман (1844–1906) — австрийские физики.

**И.Г. Лáмберт (1728–1777) — немецкий математик и физик.

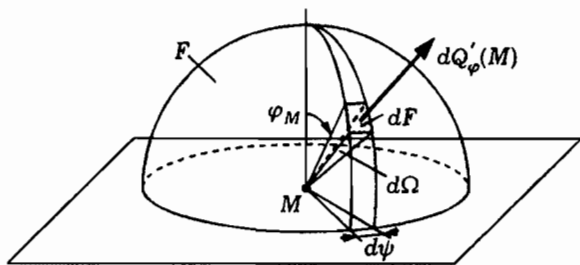


Рис. 3.9

Закон Ламберта строго справедлив для поверхности, обладающей свойствами абсолютно черного тела, но его допустимо применять и в случае достаточно шероховатой поверхности (в противоположность гладкой поверхности, отражение от которой является зеркальным).

Будем считать, что распределение по направлениям собственного излучения также подчиняется закону Ламберта

$$dQ_{\varphi}(M) = \frac{q(M)}{\pi} \cos \varphi_M dS(M) d\Omega, \quad M \in S. \quad (3.93)$$

Нетрудно убедиться, что интегрирование (3.93) по поверхности F полусферы радиуса r с центром в точке $M \in S$ приведет к значению $q(M) dS(M)$, соответствующему потоку собственного излучения площадкой $dS(M)$. Действительно, для телесного угла имеем $d\Omega = dF/r^2$, где $dF = r^2 \sin \varphi_M d\varphi_M d\psi$ — площадь элементарного участка полусферы, на который опирается этот телесный угол (см. рис. 3.9). Тогда получим

$$\begin{aligned} \int_F dQ_{\varphi}(M) dF &= \frac{q(M) dS(M)}{\pi} \int_0^{2\pi} d\psi \int_0^{\pi/2} \sin \varphi_M \cos \varphi_M d\varphi_M = \\ &= 2q(M) dS(M) \int_0^{\pi/2} \sin \varphi_M d(\sin \varphi_M) = q(M) dS(M). \end{aligned}$$

Аналогичный результат можно получить и для отраженного излучения.

В окрестности точки $N \in S$ выделим участок оболочки с площадью внутренней поверхности $dS(N)$ (см. рис. 3.8). С этого участка на площадку $dS(M)$, согласно закону Ламберта, направлен поток $dQ_\varphi(N) = \frac{q(N)}{\pi} \cos \varphi_N dS(N) d\Omega_M$ собственно-го излучения и поток $dQ\varphi'(N) = R(N) \frac{q'(N)}{\pi} \cos \varphi_N dS(N) d\Omega_M$ излучения, отраженного от этого участка. В данном случае $d\Omega_N = dS(M) \frac{\cos \varphi_M}{l_{NM}^2}$, где l_{NM}^2 — расстояние между точками N и M , а φ_N и φ_M — углы прямой NM с внутренними нормальными к оболочке в точках N и M соответственно.

В итоге на единичную площадку внутренней поверхности оболочки в окрестности точки $M \in S$ падает посылаемый площадкой $dS(N)$ поток плотностью

$$dq'_N(M) = \frac{q(N) + R(N)q'(N)}{\pi l_{NM}^2} \cos \varphi_M \cos \varphi_N dS(N).$$

Если проинтегрировать правую часть этого равенства по всей поверхности S , то найдем плотность суммарного теплового потока, падающего на площадку в окрестности точки $M \in S$:

$$q'(M) = \int_S (q(N) + R(N)q'(N)) \omega(M, N) dS(N). \quad (3.94)$$

Здесь

$$\omega(M, N) = \frac{\cos \varphi_M \cos \varphi_N}{\pi l_{NM}^2}.$$

Отметим, что при вычислении интеграла в (3.94) точка $M \in S$ фиксирована.

Обозначим $q^\circ(M) = \sigma_0 T^4(M)$. Тогда $q(M) = \varepsilon(M)q^\circ(M)$, $q^*(M) = \varepsilon^*(M)q^\circ(M)$ и вместо (3.91) можно записать

$$\tilde{q}(M) + A(M)q'(M) = (\varepsilon(M) + \varepsilon^*(M))q^\circ(M). \quad (3.95)$$

Выражая q' из (3.95) и подставляя в (3.94), приходим к интегральному уравнению относительно функции $q^o(M)$, $M \in S$:

$$\begin{aligned} & \frac{\varepsilon(M) + \varepsilon^*(M)}{A(M)} q^o(M) - \\ & - \int_S \frac{\varepsilon(N) + R(N)\varepsilon^*(N)}{A(N)} q^o(N) \omega(M, N) dS(N) = \\ & = \frac{\tilde{q}(M)}{A(M)} - \int_S \frac{R(N)}{A(N)} \tilde{q}(N) \omega(M, N) dS(N). \quad (3.96) \end{aligned}$$

Решение этого уравнения позволяет найти установившееся распределение температуры $T(M) = (q^o(M)/\sigma_0)^{1/4}$ по поверхности S рассматриваемой оболочки.

Для сферической оболочки $\varphi_M = \varphi_N$ и $\omega(M, N) = 1/(4\pi r_0^2) = \text{const}$ (рис. 3.10), где r_0 — радиус оболочки. В этом случае из (3.94) следует, что

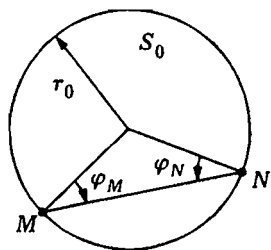


Рис. 3.10

$$q'(M) = \frac{1}{4\pi r_0^2} \int_{S_0} (\varepsilon(N) q^o(N) + R(N) q'(N)) dS(N) = q' = \text{const},$$

т.е. в окрестности любой точки $M \in S_0$ внутренней поверхности S_0 сферической оболочки плотность q' падающего потока одинакова. Подставляя в это равенство $q^o(N)$ из (3.95), находим

$$q' = \left(\int_{S_0} \frac{\varepsilon^*(N) A(N) dS(N)}{\varepsilon(N) + \varepsilon^*(N)} \right)^{-1} \int_{S_0} \frac{\varepsilon(N) \tilde{q}(N) dS(N)}{\varepsilon(N) + \varepsilon^*(N)} \quad (3.97)$$

и затем при помощи (3.95) вычисляем

$$T(M) = \sqrt[4]{\frac{q^o(M)}{\sigma_0}} = \sqrt[4]{\frac{\tilde{q}(M) + A(M) q'}{(\varepsilon(M) + \varepsilon^*(M)) \sigma_0}}. \quad (3.98)$$

Пусть свойства внутренней и внешней поверхностей оболочки соответствуют серому телу, т.е. $\varepsilon(M) = A(M) = \varepsilon = \text{const}$ и $\varepsilon^*(M) = \varepsilon^* = \text{const}$, $M \in S_0$. Тогда, согласно (3.97), в (3.98) имеем

$$q' = \frac{1}{4\pi r_0^2 \varepsilon^*} \int_{S_0} \tilde{q}(N) dS(N). \quad (3.99)$$

Примем, что $\tilde{q}(M)$ — плотность поглощенного внешней поверхностью сферической оболочки потока солнечного излучения на околоземной орбите. Если направить ось Oz отсчета угла α на Солнце (рис. 3.11), то получим при $0 < \beta \leq 2\pi$

$$\tilde{q}(\alpha) = \begin{cases} \varepsilon^* q_S \cos \alpha, & 0 \leq \alpha \leq \frac{\pi}{2}; \\ 0, & \frac{\pi}{2} < \alpha \leq \pi, \end{cases} \quad (3.100)$$

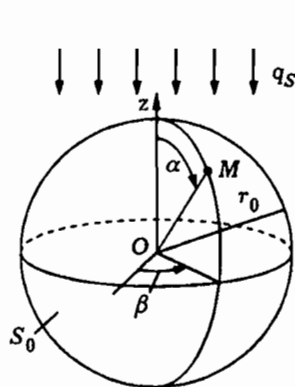


Рис. 3.11

где q_S — солнечная постоянная, характеризующая плотность потока солнечного излучения на среднем расстоянии Земли от Солнца (примем, что $(q_S/\sigma_0)^{1/4} = 390$ К). Полагая для сферы $dS = r_0^2 \sin \alpha d\alpha d\beta$, из (3.98)–(3.100) при $\varepsilon^* = \varepsilon$ находим, что установившееся распределение температуры, симметричное относительно оси Oz , на освещенной стороне оболочки будет

$$T(\alpha) = 328 \sqrt[4]{\frac{1}{4} + \cos \alpha} \text{ К}.$$

Наибольшая температура $T_{\max} = 347$ К соответствует значению $\alpha = 0$, а вся неосвещенная сторона оболочки имеет температуру $T_{\min} = 232$ К.

Пример 3.6. Пусть рассмотренная в примере 3.5 сферическая оболочка S_0 радиуса r_0 имеет толщину h , а материал оболочки обладает теплопроводностью λ . Составим уравнение

переноса тепловой энергии излучением с учетом передачи теплоты путем теплопроводности, считая по-прежнему температуру по толщине оболочки однородной. Теперь баланс тепловой энергии участка оболочки в окрестности точки $M \in S_0$ вместо (3.91) примет вид

$$\lambda h \nabla^2 T(M) + \tilde{q}(M) + A(M)q'(M) = (\varepsilon(M) + \varepsilon^*(M))\sigma_0 T^4(M), \quad (3.101)$$

причем в данном случае оператор Лапласа от искомой функции $T(M)$ температуры следует вычислять по поверхности сферы [VII]:

$$\nabla^2 T = \frac{1}{r_0^2 \sin \alpha} \frac{\partial T}{\partial \alpha} \left(\sin \alpha \frac{\partial T}{\partial \alpha} \right) + \frac{1}{r_0^2 \sin^2 \alpha} \frac{\partial^2 T}{\partial \beta^2},$$

где α, β — угловые координаты точки $M \in S_0$ (см. рис. 3.11).

Отметим, что $\nabla^2 T = \frac{2}{r_0^2} \frac{\partial^2 T}{\partial \alpha^2}$ при $\alpha = 0$ и $\alpha = \pi$.

Используя (3.101), вместо (3.97) теперь получаем

$$q' = \frac{\int_{S_0} \varepsilon(N) \frac{\tilde{q}(N) + \lambda h \nabla^2 T(M)}{\varepsilon(N) + \varepsilon^*(N)} dS(N)}{\int_{S_0} \frac{\varepsilon^*(N) A(N)}{\varepsilon(N) + \varepsilon^*(N)} dS(N)}$$

и, подставляя в (3.101), приходим к интегро-дифференциальному уравнению

$$\begin{aligned} \lambda h \nabla^2 T(M) = & (\varepsilon(M) + \varepsilon^*(M))\sigma_0 T^4(M) - \tilde{q}(M) - \\ & - A(M) \frac{\int_{S_0} \varepsilon(N) \frac{\tilde{q}(N) + \lambda h \nabla^2 T(M)}{\varepsilon(N) + \varepsilon^*(N)} dS(N)}{\int_{S_0} \frac{\varepsilon^*(N) A(N)}{\varepsilon(N) + \varepsilon^*(N)} dS(N)} \end{aligned}$$

относительно искомой функции $T(M)$, $M \in S_0$.

Вопросы и задачи

3.1. Доказать, что для идеальной жидкости (газа) мощность напряжений можно представить в виде $\frac{p}{\rho} \frac{d\rho}{dt}$.

3.2. Показать, что равновесие покоящейся идеальной жидкости (или газа) возможно лишь при условии $\mathbf{b}(\nabla \times \mathbf{b}) = 0$.

3.3. Доказать, что разложение тензора напряжений на шаровой тензор и девиатор единственно.

3.4. Значения компонент (в МПа) тензора напряжений в некоторой точке среды заданы матрицей

$$\begin{pmatrix} 70 & 0 & -20 \\ 0 & 50 & 0 \\ -20 & 0 & 40 \end{pmatrix}.$$

Разложить его на шаровой тензор и девиатор. Найти вектор напряжения, действующий в проходящей через эту точку площадке с единичным вектором нормали $\mathbf{n} = 2\mathbf{e}_1/3 - 2\mathbf{e}_2/3 + \mathbf{e}_3/3$, где $\mathbf{e}_i$, $i = 1, 2, 3$, — орты прямоугольной системы координат $Ox_1x_2x_3$. Вычислить модуль этого вектора, его проекцию на направление вектора нормали и угол между этими векторами.

3.5. Вывести (3.20) из (3.19).

3.6. Используя (3.31), получить зависимости компонент тензора деформаций от компонент тензора напряжений и вывести (3.34).

3.7. В некоторой точке среды известен тензор скоростей деформаций. Записать выражения для объемной мощности источников энергии, связанной с работой напряжений в идеальной и вязкой сжимаемой и несжимаемой средах.

3.8. Вывести волновое (3.59) и телеграфное (3.61) уравнения из уравнений (3.58) Максвелла.

ЧАСТЬ II

Элементы функционального анализа и приближенные аналитические методы

4. НОРМИРОВАННЫЕ ПРОСТРАНСТВА И ОПЕРАТОРЫ

Постановки задач математической физики, рассмотренные в части I этой книги, включают *функциональные уравнения*, связывающие искомые и заданные функции, принадлежащие некоторым множествам функций. Поэтому при изучении приближенных методов математической физики возникает необходимость использовать некоторые понятия и методы функционального анализа. Одними из основных объектов изучения в функциональном анализе являются бесконечномерные нормированные *пространства*, элементами которых во многих случаях являются функции. В этой главе использованы некоторые положения функционального анализа, изложенные в [IX].

4.1. Нормированные пространства

Множество, между элементами которого установлены определенные соотношения, часто называют *пространством*. Если такое множество состоит из функций, то говорят о *функциональном пространстве*. Так, если множество функций удовлетворяет аксиомам линейного пространства [IV], то его часто называют линейным функциональным пространством. Его элементами могут быть, например, скалярные или векторные (действительные или комплексные) функции. Векторные функции будем обозначать так же, как и векторы — полужирным курсивом (например, $\mathbf{f}$, $\mathbf{u}$, $\mathbf{v}$), а скалярные — светлым курсивом (например, f , u , v).

Пример 4.1. Множество всех определенных на числовой прямой $\mathbb{R}$ многочленов $P_k(x) = a_0x^k + a_1x^{k-1} + \dots + a_{k-1}x + a_k$ степени не выше некоторого натурального числа $k \in \mathbb{N}$ с произвольными действительными или комплексными коэффициента-

ми (с обычными операциями сложения функций и умножения на числа) является линейным функциональным пространством. Множества скалярных действительных функций или вектор-функций, непрерывных или непрерывно дифференцируемых на отрезке $[a, b]$, удовлетворяют аксиомам линейного пространства и поэтому являются линейными функциональными пространствами. #

В дальнейшем, если специально не оговорено, будем рассматривать линейные функциональные пространства, элементами которых являются скалярные действительные функции с операцией умножения только на действительные числа. Если система $\{u_n\}_N$ элементов линейного функционального пространства $\mathcal{U}$ является в $\mathcal{U}$ линейно зависимой (независимой) [IV], то для краткости эту систему будем называть линейно зависимыми (независимыми) функциями в $\mathcal{U}$. В основном будем изучать *бесконечномерные линейные функциональные пространства*, т.е. такие, в которых можно указать сколь угодно большое число линейно независимых функций.

Под *линейной оболочкой* системы функций $u_n \in \mathcal{U}$, $n \in \mathbb{N}$, образующих (бесконечную) последовательность $\{u_n\}$, будем понимать множество всевозможных (конечных) линейных комбинаций функций этой системы. Систему $\{u_n\}$ функций из $\mathcal{U}$ называют линейно независимой, если любая ее конечная подсистема линейно независима. Линейная оболочка бесконечной линейно независимой системы $\{u_n\}$ функций является бесконечномерным линейным функциональным пространством.

Пример 4.2. Множество всех действительных функций одного действительного переменного, непрерывных на отрезке $[a, b]$, является линейным пространством (см. пример 4.1), причем бесконечномерным, поскольку для любого $N \in \mathbb{N}$ существует N линейно независимых элементов этого линейного пространства. Например, многочлены x^n , $n = \overline{1, N}$, линейно независимы. Линейное пространство многочленов степени не выше некоторого натурального числа N конечномерно [IV]. #

Говорят, что в линейном пространстве $\mathcal{U}$ задана *норма*, если каждому элементу $u \in \mathcal{U}$ поставлено в соответствие действительное число $\|u\|$, причем верны три аксиомы нормы:

- 1) $\|u\| \geq 0$ и $\|u\| = 0$ тогда и только тогда, когда $u = 0$;
- 2) $\|\alpha u\| = |\alpha| \|u\|$, $\alpha \in \mathbb{R}$;
- 3) $\|u + v\| \leq \|u\| + \|v\|$, $u, v \in \mathcal{U}$ (неравенство треугольника).

Линейное пространство, в котором задана норма, называют *нормированным пространством*. В случае *линейного функционального пространства*, в котором определена норма, говорят о *функциональном нормированном пространстве*, а норму элемента (функции) в таком функциональном пространстве называют *нормой функции*.

Чтобы подчеркнуть, что речь идет о норме функции u в пространстве $\mathcal{U}$, будем писать $\|u\|_{\mathcal{U}}$, а правило, которое устанавливает соответствие между функцией $u \in \mathcal{U}$ и ее нормой в $\mathcal{U}$, будем обозначать $\|\cdot\|_{\mathcal{U}}$.

Отметим, что в нормированном пространстве $\mathcal{U}$ справедливо неравенство

$$0 \leq \left| \|u\| - \|v\| \right| \leq \|u \pm v\| \leq \|u\| + \|v\|, \quad u, v \in \mathcal{U}. \quad (4.1)$$

Действительно, если в неравенстве треугольника (аксиома 3) заменить сначала u на $u - v$, а затем v на $v - u$, то получим неравенства $\|u\| - \|v\| \leq \|u - v\|$ и $\|v\| - \|u\| \leq \|v - u\|$. Учитывая равенства $\| -v \| = \|v\|$ и $\|u - v\| = \|v - u\|$ (аксиома 2), приходим к (4.1).

Нормированное пространство $\mathcal{U}$ является *метрическим пространством с метрикой*, индуцированной нормой, т.е. $\rho(u, v) = \|u - v\|$, $u, v \in \mathcal{U}$. Под сходимостью последовательности $\{u_n\}$ по норме $\|\cdot\|_{\mathcal{U}}$ понимают сходимость этой последовательности по метрике ρ , индуцированной нормой $\|\cdot\|_{\mathcal{U}}$. Элемент $u \in \mathcal{U}$ называют *пределом последовательности* $\{u_n\} \subset \mathcal{U}$, *сходящейся по норме*, если

$$\lim_{n \rightarrow \infty} \|u_n - u\| = 0. \quad (4.2)$$

Понятия окрестности точки, внутренней, изолированной, граничной и предельной точки множества, фундаментальной последовательности, ограниченности, замкнутости и компактности множества, введенные для метрического пространства [1], применимы и для нормированного пространства.

Определение 4.1. Последовательность $\{u_n\} \subset U$ называют **фундаментальной** в нормированном пространстве U , если для любого числа $\varepsilon > 0$ можно указать такое число $N \in \mathbb{N}$, что при всех $m > N$ и $n > N$ верно неравенство $\|u_m - u_n\| < \varepsilon$.

В полном нормированном (банаховом*) пространстве любая фундаментальная последовательность сходится по норме. Банахово пространство будем обозначать символом B .

Пусть U — нормированное пространство. Напомним [IX], что если для множества $X \subset U$ его замыкание $\overline{X}$, т.е. объединение X со всеми предельными точками X из U , совпадает с U , то множество X называют *всюду плотным* в U . С учетом определения предельной точки множества в метрическом пространстве [1] можно дать следующее эквивалентное определение.

Определение 4.2. Множество $X \subset U$ называют *всюду плотным* в нормированном пространстве U , если для любого $u \in U$ и любого $\varepsilon > 0$ найдется такой элемент $u' \in X$, что $\|u - u'\| < \varepsilon$.

Нормированное пространство U называют *сепарабельным*, если в этом пространстве существует счетное всюду плотное подмножество.

Пусть $U_1 \subset U$ — линейная оболочка некоторой системы элементов u_n , $n \in \mathbb{N}$, нормированного пространства U . Если эта оболочка — замкнутое множество, то она является *подпространством* нормированного пространства U . Напомним, что систему элементов называют *замкнутой* в банаховом пространстве, если ее линейная оболочка является всюду плотной в

*С. Ба́нах (1892–1945) — польский математик.

этом пространстве. Замкнутость системы $\{u_n\}$ означает, что любой элемент $u \in \mathcal{U}$ можно сколь угодно точно (по норме этого пространства) представить конечными линейными комбинациями элементов данной системы, т.е. для любых $u \in \mathcal{U}$ и $\varepsilon > 0$ можно подобрать такой номер $N \in \mathbb{N}$ и такие коэффициенты $a_n \in \mathbb{R}$, $n = \overline{1, N}$, что для элемента

$$\tilde{u}_N = \sum_{n=1}^N a_n u_n$$

будет выполнено неравенство $\|u - \tilde{u}_N\| < \varepsilon$.

Пример 4.3. Функциональное пространство непрерывных на отрезке $[a, b]$ действительных функций $f(x)$ одного действительного переменного x является линейным (см. пример 4.1). Оно будет банаховым $[IX]$, если ввести норму

$$\|f\| = \max_{x \in [a, b]} |f(x)|. \quad (4.3)$$

Это банахово пространство обозначают $C[a, b]$.

В линейном функциональном пространстве $C[a, b]$ можно ввести и другую норму. Из свойств интеграла Римана и неравенства треугольника для действительных чисел следует, что соотношение

$$\|f\| = \int_a^b |f(x)| dx, \quad f \in C[a, b], \quad (4.4)$$

задает норму в $C[a, b]$. Но полученное нормированное пространство с нормой (4.4) не является полным. Покажем это для случая $[a, b] = [0, 1]$. Например, нетрудно убедиться, что последовательность непрерывных на отрезке $[0, 1]$ функций (рис. 4.1)

$$\varphi_n(x) = \begin{cases} 0, & 0 \leq x < 1/2 - 1/n; \\ nx + 1 - n/2, & 1/2 - 1/n \leq x < 1/2; \\ 1, & 1/2 \leq x \leq 1, \end{cases} \quad (4.5)$$

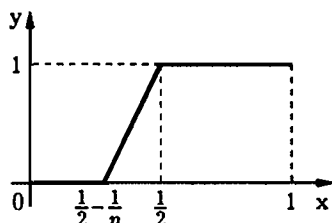


Рис. 4.1

в нормированном пространстве $C[a, b]$ с нормой (4.4) является фундаментальной. Действительно, для любого $\varepsilon > 0$ при любых $n > \frac{1}{2\varepsilon}$ и $m > n$ имеем

$$\begin{aligned} \|\varphi_n - \varphi_m\| &= \int_0^1 |\varphi_n(x) - \varphi_m(x)| dx = \\ &= \int_{1/2 - 1/n}^{1/2} |\varphi_n(x) - \varphi(x)_m| dx \leqslant \\ &\leqslant \int_{1/2 - 1/n}^{1/2} \left(nx + 1 - \frac{n}{2} \right) dx = \frac{1}{2n} < \varepsilon. \end{aligned}$$

Однако эта фундаментальная последовательность не имеет предела в $C[a, b]$. Предположим противное: пусть существует непрерывная на отрезке $[0, 1]$ функция φ_0 , для которой

$$\lim_{n \rightarrow \infty} \|\varphi_n - \varphi_0\| = \lim_{n \rightarrow \infty} \int_0^1 |\varphi_n(x) - \varphi_0(x)| dx = 0.$$

Заметим, что для разрывной в точке $x = 1/2$ функции

$$\varphi(x) = \begin{cases} 0, & 0 \leqslant x < 1/2; \\ 1, & 1/2 \leqslant x \leqslant 1, \end{cases} \quad (4.6)$$

справедливы равенства

$$\begin{aligned}\lim_{n \rightarrow \infty} \int_0^1 |\varphi_n(x) - \varphi(x)| dx &= \\ &= \lim_{n \rightarrow \infty} \int_{1/2-1/n}^{1/2} \left| nx + 1 - \frac{n}{2} \right| dx = \lim_{n \rightarrow \infty} \frac{1}{2n} = 0.\end{aligned}$$

Тогда

$$\begin{aligned}\int_0^1 |\varphi_0(x) - \varphi(x)| dx &= \lim_{n \rightarrow \infty} \int_0^1 |\varphi_0(x) - \varphi(x)| dx \leq \\ &\leq \lim_{n \rightarrow \infty} \int_0^1 |\varphi_0(x) - \varphi_n(x)| dx + \lim_{n \rightarrow \infty} \int_0^1 |\varphi_n(x) - \varphi(x)| dx = 0.\end{aligned}$$

Поскольку $|\varphi_0(x) - \varphi(x)| \geq 0$ для $x \in [0, 1]$ и

$$\int_0^1 |\varphi_0(x) - \varphi(x)| dx = \int_0^{1/2} |\varphi_0(x) - \varphi(x)| dx + \int_{1/2}^1 |\varphi_0(x) - \varphi(x)| dx,$$

то каждый из двух последних интегралов равен нулю. Отсюда в силу неотрицательности и непрерывности подынтегральной функции $|\varphi_0(x) - \varphi(x)|$ на каждом из промежутков $[0, 1/2]$ и $[1/2, 1]$ получаем $\varphi_0(x) = \varphi(x)$ для всех $x \in [0, 1/2] \cup [1/2, 1]$. Тогда $x = 1/2$ является для функции $\varphi_0(x)$ точкой разрыва первого рода, что противоречит предположению непрерывности φ_0 на отрезке $[0, 1]$.

Таким образом, последовательность $\{\varphi_n\}$, фундаментальная в линейном нормированном пространстве $C[a, b]$ с нормой (4.4), не является сходящейся в $\mathcal{U}$ по этой норме, т.е. $C[a, b]$ с нормой (4.4) не будет полным нормированным пространством.

Можно показать*, что банаховым является пространство суммируемых (или интегрируемых по Лебегу**) на отрезке $[a, b]$ функций $f(x)$, обозначаемое $L_1[a, b]$ и имеющее норму $\|f\|$

$$\|f\| = \int_a^b |f(x)| dx, \quad (4.7)$$

которая определена при помощи интеграла Лебега. Элементами пространства $L_1[a, b]$ являются классы функций, равных почти всюду на отрезке $[0, 1]$. В частности, этому пространству при $[a, b] = [0, 1]$ принадлежит не интегрируемая по Риману на отрезке $[0, 1]$, но интегрируемая по Лебегу на этом отрезке функция Дирихле

$$\chi(x) = \begin{cases} 1, & x \in \mathbb{Q} \cap [0, 1]; \\ 0, & x \in [0, 1] \setminus \mathbb{Q}, \end{cases}$$

причем $\|\chi(x)\| = 0$, поскольку эта функция почти всюду равна нулю.

Пример 4.4. Функциональное пространство $\mathcal{U}$ действительных функций одного действительного переменного, k раз непрерывно дифференцируемых на отрезке $[a, b]$, является линейным (см. пример 4.1). В этом функциональном пространстве можно вводить следующие нормы:

$$\begin{aligned} \|f\|_0 &= \max_{x \in [a, b]} |f(x)|, \\ \|f\|_1 &= \max_{0 \leq m \leq k} \max_{x \in [a, b]} |f^{(m)}(x)|, \\ \|f\|_2 &= \sum_{m=0}^k \max_{x \in [a, b]} |f^{(m)}(x)|, \end{aligned}$$

*См.: Колмогоров А.Н., Фомин С.В.

**А. Лебег (1875–1941) — французский математик.

где $f^{(m)}(x)$, $m \geq 1$, — производная порядка m функции $f(x)$, а $f^{(0)}(x) = f(x)$.

Функциональное пространство $\mathcal{U}$ с нормой $\|\cdot\|_0$ является линейным многообразием банахова пространства $C[a, b]$ (см. пример 4.3). Поскольку это линейное многообразие не является замкнутым в $C[a, b]$ (существуют последовательности непрерывно дифференцируемых функций, сходящиеся в $C[a, b]$, т.е. по норме $\|\cdot\|_0$, к непрерывной функции, которая ни в одной точке отрезка $[a, b]$ не имеет конечной производной*), то функциональное пространство $\mathcal{U}$ с нормой $\|\cdot\|_0$ не является банаховым.

Нетрудно показать, что нормы $\|\cdot\|_1$ и $\|\cdot\|_2$, действующие в функциональном пространстве $\mathcal{U}$, являются эквивалентными, т.е. найдутся такие положительные числа α и $\beta > 0$, что

$$\alpha\|f\|_2 \leq \|f\|_1 \leq \beta\|f\|_2, \quad f \in \mathcal{U}.$$

Отметим, что последовательность $\{f_n\} \subset \mathcal{U}$ сходится по нормам $\|\cdot\|_1$ или $\|\cdot\|_2$ тогда и только тогда, когда функциональные последовательности $\{f_n^{(m)}(x)\}$ производных сходятся равномерно на отрезке $[a, b]$ при любом $m = \overline{0, k}$ [IX]. Отсюда, в частности, ясно, что из сходимости последовательности $\{f_n\} \subset \mathcal{U}$ по нормам $\|\cdot\|_1$ или $\|\cdot\|_2$ следует сходимость этой последовательности по норме $\|\cdot\|_0$. Обратное утверждение неверно. Так, например, последовательность функций $f_n(x) = \frac{1}{n} \sin n^2 x$, $n \in \mathbb{N}$, равномерно сходится на отрезке $[0, 1]$ к нулевой функции, поскольку $\max_{x \in [0, 1]} |f_n(x)| \leq \frac{1}{n} \rightarrow 0$ при $n \rightarrow \infty$ и, следовательно, сходится к нулю по норме $\|\cdot\|_0$. Функциональная последовательность $\{f'_n(x)\} = \{n \cos n^2 x\}$ первых производных не является равномерно сходящейся на отрезке $[0, 1]$. Более того, она расходится при $x = 0$, так как $f'_n(0) = n$.

Функциональное пространство $\mathcal{U}$ с нормой $\|\cdot\|_1$ или нормой $\|\cdot\|_2$ является банаховым и его обозначают $C^k[a, b]$.

*См.: Фигтенгольц Г.М.

4.2. Операторы в нормированных пространствах

Пусть $\mathcal{U}$ и $\mathcal{W}$ — некоторые множества. Если множества $\mathcal{U}$ и $\mathcal{W}$ наделены некоторой структурой, например, являются *функциональными пространствами* (линейными, нормированными и т.п.), то обычно отображение $A: \mathcal{U} \rightarrow \mathcal{W}$ называют **оператором**. При действиях с операторами используют терминологию, связанную с отображениями множеств. В частности, $\mathcal{U}$ называют **областью определения оператора** A , обычно обозначаемую $D(A)$, а подмножество

$$R(A) = \{w \in \mathcal{W}: \exists u \in \mathcal{U} (w = A(u))\} \quad —$$

областью значений оператора A .

Если заданы операторы $A: \mathcal{U} \rightarrow \mathcal{W}$ и $B: \mathcal{W}_1 \rightarrow \mathcal{V}$, где $\mathcal{U}$, $\mathcal{W}$, $\mathcal{W}_1$, $\mathcal{V}$ — некоторые множества, а $R(A) \subset D(B) = \mathcal{W}_1$, то говорят о **композиции операторов** $B \circ A$ с областью определения $D(B \circ A) = \mathcal{U}$. Если включение $R(A) \subset D(B)$ не имеет места, но $R(A) \cap D(B) \neq \emptyset$, то композицию операторов $B \circ A$ можно определить на более узком множестве $D' \subset \mathcal{U}$, таком, что для любого $u \in D'$ элемент $A(u) \in D(B)$.

Если заданы два оператора $A_1: \mathcal{U}_1 \rightarrow \mathcal{W}$ и $A_2: \mathcal{U}_2 \rightarrow \mathcal{W}$, причем $\mathcal{U}_1 \subset \mathcal{U}_2$ и $A_1(u) = A_2(u)$ для любых $u \in \mathcal{U}_1$, то оператор A_1 называют **сужением оператора** A_2 , а оператор A_2 — **расширением** A_1 .

Определение 4.3. Оператор $I_{\mathcal{U}}: \mathcal{U} \rightarrow \mathcal{U}$, который переводит любой элемент $u \in \mathcal{U}$ в себя, т.е. $I_{\mathcal{U}}(u) = u$, $u \in \mathcal{U}$, называют **тождественным**. Оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, $R(A) = \mathcal{W}$, называют **взаимно однозначным**, если каждому элементу $w \in R(A)$ отвечает единственный прообраз $u \in D(A)$, т.е. для каждого $w \in \mathcal{W}$ существует и притом единственное решение **уравнения**

$$A(u) = w, \quad (4.8)$$

называемого **операторным**. Если оператор $A: \mathcal{U} \rightarrow \mathcal{W}$ является взаимно однозначным, то **оператор** $A^{-1}: \mathcal{W} \rightarrow \mathcal{U}$, удо-

влетворяющий условию $A^{-1}(w) = u$ тогда и только тогда, когда выполнено (4.8), называют **обратным** к A . При этом $D(A^{-1}) = R(A)$ и $R(A^{-1}) = D(A)$. Композиции $A^{-1} \circ A = I_U$ и $A \circ A^{-1} = I_W$ являются тождественными операторами, преобразующими любой элемент множества U или W в себя.

Примером взаимно однозначного оператора является *линейное преобразование* линейного (векторного) пространства $\mathbb{R}^n$ в себя, имеющее в некотором базисе невырожденную матрицу. В этом случае обратному оператору в том же базисе будет соответствовать обратная матрица. Произведение невырожденной матрицы на обратную к ней равно, как известно, единичной матрице, которая соответствует тождественному оператору.

Пример 4.5. Пусть оператор $A: C[a, b] \rightarrow C[a, b]$ каждой непрерывной на отрезке $[a, b]$ функции f ставит в соответствие непрерывную функцию $\varphi \in C[a, b]$ по правилу

$$\varphi(x) = \int_a^x f(t) dt, \quad x \in [a, b].$$

Тогда $R(A) = U \subset C[a, b]$ — множество всех непрерывно дифференцируемых на отрезке $[a, b]$ функций с условием $\varphi(a) = 0$. Если рассматривать оператор A как действующий из $C[a, b]$ в U , то он имеет обратный оператор $A^{-1}: U \rightarrow C[a, b]$, причем равенство $A^{-1}(\varphi) = f$, $\varphi \in U$, эквивалентно тому, что $f(x) = \varphi'(x)$, $x \in [a, b]$.

Пример 4.6. Покажем, что оператор $A: C[0, 1] \rightarrow C[0, 1]$, который каждой непрерывной на отрезке $[0, 1]$ функции f ставит в соответствие непрерывную функцию $\varphi = A(f)$ по правилу

$$\varphi(x) = \int_0^x f(t) dt + f(x), \quad x \in [0, 1], \quad (4.9)$$

имеет обратный оператор $A^{-1}: C[0, 1] \rightarrow C[0, 1]$.

Рассмотрим сужение оператора A на множество $\mathcal{U}$ непрерывно дифференцируемых на отрезке $[0, 1]$ функций. Если функция $f \in \mathcal{U}$, то функция $\varphi = A(f)$ также принадлежит множеству $\mathcal{U}$. В соответствии с (4.9) имеем

$$\varphi'(x) = f(x) + f'(x), \quad x \in [0, 1], \quad (4.10)$$

причем $\varphi(0) = f(0)$. Разрешим (4.10) относительно функции f . Оно является линейным неоднородным обыкновенным дифференциальным уравнением (ОДУ) относительно неизвестной функции f , удовлетворяющей начальному условию $f(0) = \varphi(0)$. Общее решение соответствующего однородного ОДУ имеет вид $f(x) = Ce^{-x}$. Решая (4.10) методом Лагранжа [VIII], получаем

$$f(x) = e^{-x} \int_0^x e^t \varphi'(t) dt + Ce^{-x}.$$

Интегрируя по частям, получаем

$$\begin{aligned} f(x) &= e^{t-x} \varphi(t) \Big|_0^x - e^{-x} \int_0^x e^t \varphi(t) dt + Ce^{-x} = \\ &= \varphi(x) - \varphi(0)e^{-x} - e^{-x} \int_0^x e^t \varphi(t) dt + Ce^{-x}. \end{aligned}$$

Отсюда, используя начальное условие $f(0) = \varphi(0)$, находим $C = \varphi(0)$. Значит, решение уравнения (4.10) имеет вид

$$f(x) = \varphi(x) - \int_0^x e^{t-x} \varphi(t) dt, \quad x \in [0, 1]. \quad (4.11)$$

Таким образом, для любой функции $\varphi \in \mathcal{U}$ существует единственная функция $f \in \mathcal{U}$, удовлетворяющая уравнению $A(f) = \varphi$, причем функция f определена равенством (4.11).

Докажем, что оператор $B: C[0, 1] \rightarrow C[0, 1]$, ставящий в соответствие всякой функции $\varphi \in C[0, 1]$ функцию $f = B(\varphi)$ по правилу (4.11), является обратным к оператору A . Найдем образ функции $\varphi \in C[0, 1]$ при действии оператора $A \circ B$. Если $\varphi_1 = A \circ B(\varphi) = A(B(\varphi))$, то

$$\begin{aligned}\varphi_1(x) &= \int_0^x \left(\varphi(\tau) - \int_0^\tau e^{t-\tau} \varphi(t) dt \right) d\tau + \varphi(x) - \int_0^x e^{t-x} \varphi(t) dt = \\ &= \int_0^x \varphi(\tau) d\tau - \int_0^x e^{-\tau} d\tau \int_0^\tau e^t \varphi(t) dt + \varphi(x) - e^{-x} \int_0^x e^t \varphi(t) dt.\end{aligned}$$

Так как функция $g(\tau, t) = e^{t-\tau} \varphi(t)$ непрерывна в замкнутой области $D = \{(\tau, t) \in \mathbb{R}^2: 0 \leq \tau \leq x, 0 \leq t \leq \tau\}$, то можно поменять порядок интегрирования в повторном интеграле [VII]

$$\begin{aligned}\int_0^x e^{-\tau} d\tau \int_0^\tau e^t \varphi(t) dt &= \int_0^x e^t \varphi(t) dt \int_t^x e^{-\tau} d\tau = \\ &= - \int_0^x e^t (e^{-x} - e^{-t}) \varphi(t) dt = -e^{-x} \int_0^x e^t \varphi(t) dt + \int_0^x \varphi(t) dt.\end{aligned}$$

Итак, возвращаясь к вычислению $\varphi_1(x)$, получаем $\varphi_1(x) = \varphi(x)$, $x \in [0, 1]$, т.е. $A \circ B(\varphi) = \varphi$, $\varphi \in [0, 1]$.

Аналогично можно показать, что $B \circ A(f) = f$, $f \in C[0, 1]$. Таким образом, оператор B является обратным к оператору A , т.е. $B = A^{-1}$.

Определение 4.4. Операторы $A_l^{-1}: \mathcal{W} \rightarrow \mathcal{U}$ и $A_r^{-1}: \mathcal{W} \rightarrow \mathcal{U}$ называют соответственно *левым* и *правым обратным* к оператору $A: \mathcal{U} \rightarrow \mathcal{W}$, если

$$A_l^{-1} \circ A = I_{\mathcal{U}} \quad \text{и} \quad A \circ A_r^{-1} = I_{\mathcal{W}}, \quad (4.12)$$

где $I_{\mathcal{U}}$ и $I_{\mathcal{W}}$ — тождественные операторы в $\mathcal{U}$ и $\mathcal{W}$.

Отметим, что если существует оператор A^{-1} , обратный к оператору A , то, согласно определению 4.3, оба эти оператора являются взаимно однозначными и уравнение $Au = f$ имеет единственное решение $u = A^{-1}f \in \mathcal{U}$. Существование оператора A_l^{-1} не означает, что оператор A сюръективный, а существование оператора A_r^{-1} не означает, что оператор A инъективный.

Теорема 4.1. Если существует правый обратный оператор A_r^{-1} к оператору $A: \mathcal{U} \rightarrow \mathcal{W}$, то операторное уравнение

$$A(u) = f \quad (4.13)$$

для любого $f \in \mathcal{W}$ имеет решение $u = A_r^{-1}(f) \in \mathcal{U}$, а если существует левый обратный оператор A_l^{-1} , то уравнение (4.13) имеет не более одного решения. Если существуют и правый и левый обратные операторы, то они единственны и совпадают с обратным оператором A^{-1} , а (4.13) имеет единственное решение.

◀ Пусть существует A_r^{-1} . Тогда элемент $u = A_r^{-1}(f) \in \mathcal{U}$ обращает (4.13) в тождество, так как

$$A(A_r^{-1}(f)) = (A \circ A_r^{-1})(f) = I_{\mathcal{W}}(f) = f \in \mathcal{W}.$$

Допустим, что существует A_l^{-1} . Если $u \in \mathcal{U}$ является решением уравнения $A(u) = f$, то $A_l^{-1} \circ A(u) = A_l^{-1}(f)$, или $u = A_l^{-1}(f)$, т.е. решение u является образом элемента $f \in \mathcal{W}$ при действии оператора A_l^{-1} и определено однозначно.

Из существования и правого, и левого обратных операторов следует, что (4.13) имеет и притом единственное решение, т.е. оператор A взаимно однозначный и имеет, согласно определению 4.3, обратный оператор A^{-1} . Пусть наряду с A_l^{-1} существует еще один левый обратный оператор B , т.е. $A_l^{-1} \circ A = B \circ A = I_{\mathcal{U}}$. Тогда $A_l^{-1} \circ A \circ A^{-1} = B \circ A \circ A^{-1}$, или $A_l^{-1} \circ I_{\mathcal{W}} = B \circ I_{\mathcal{W}}$. Следовательно, $A_l^{-1} = B$, что доказывает единственность A_l^{-1} . Аналогично можно доказать единственность A_r^{-1} .

Так как A_l^{-1} и A_r^{-1} единственны, а A^{-1} — оператор, одновременно и левый, и правый обратный к A , то $A^{-1} = A_l^{-1} = A_r^{-1}$. ►

Пусть теперь $\mathcal{U}$ и $\mathcal{W}$ — нормированные функциональные пространства с нормами $\|\cdot\|_{\mathcal{U}}$ и $\|\cdot\|_{\mathcal{W}}$ соответственно, а область определения $D(A)$ оператора $A: \mathcal{U} \rightarrow \mathcal{W}$ содержит проколотую окрестность $\overset{\circ}{U}$ точки u_0 (в этой точке оператор A может быть и не определен). Элемент $w_0 \in \mathcal{W}$ называют пределом оператора A в точке u_0 , если для любого $\varepsilon > 0$ найдется такое $\delta = \delta(\varepsilon)$, что для любого $u \in \overset{\circ}{U}$ при условии $\|u - u_0\|_{\mathcal{U}} < \delta$ будет выполнено $\|A(u) - w_0\|_{\mathcal{W}} < \varepsilon$. При этом пишут $A(u) \xrightarrow{\mathcal{W}} w_0$ при $u \xrightarrow{\mathcal{U}} u_0$.

Это определение соответствует понятию предела отображения метрических пространств $\mathcal{U}$ и $\mathcal{W}$ с метриками, индуцированными нормами этих пространств.

Определение 4.5. Оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, где $\mathcal{U}$ и $\mathcal{W}$ — нормированные пространства, называют **ограниченным**, если он преобразует всякое ограниченное множество из своей области определения $D(A) = \mathcal{U}$ во множество элементов, ограниченное в $\mathcal{W}$.

Определение 4.6. Оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, где $\mathcal{U}$ и $\mathcal{W}$ — нормированные пространства и $D(A) = \mathcal{U}$, называют **непрерывным в точке $u_0 \in D(A)$** , если для любой окрестности $W \subset \mathcal{W}$ точки $w_0 = A(u_0)$ существует такая окрестность $U \subset \mathcal{U}$ точки u_0 , что $A(u) \in W$ для любой точки $u \in U$ (кратко пишут $A(u) \xrightarrow{\mathcal{W}} A(u_0)$ при $u \xrightarrow{\mathcal{U}} u_0$). Оператор, непрерывный в каждой точке $u_0 \in D(A)$, называют **непрерывным**.

Эти определения соответствуют определениям непрерывности отображения в метрических пространствах [1].

Пример 4.7. Покажем, что оператор $A: C[0, 1] \rightarrow C[0, 1]$, заданный правилом

$$\varphi = A(f) \iff \varphi(x) = f^2(x), \quad x \in [0, 1],$$

является ограниченным и непрерывным. Действительно, если $\mathcal{U}$ — ограниченное множество в $C[0, 1]$, то найдется такое число $M > 0$, что для любых функций $f \in \mathcal{U}$ справедливо неравенство $\|f\| = \max_{x \in [0, 1]} |f(x)| \leq M$. Тогда, если $\varphi = A(f)$, $f \in \mathcal{U}$, то $\|\varphi\| = \max_{x \in [0, 1]} f^2(x) \leq M^2$, и множество $A(\mathcal{U})$, состоящее из образов всех элементов $f \in \mathcal{U}$, ограничено.

Покажем, что оператор A является непрерывным. В самом деле, для любых функций $f, f_0 \in C[0, 1]$ имеем

$$\begin{aligned} \|A(f) - A(f_0)\| &= \max_{x \in [0, 1]} |f^2(x) - f_0^2(x)| = \\ &= \max_{x \in [0, 1]} |f(x) - f_0(x)| |f(x) + f_0(x)| \leq \\ &\leq \max_{x \in [0, 1]} |f(x) - f_0(x)|^2 + 2\|f_0\| \max_{x \in [0, 1]} |f(x) - f_0(x)| = \\ &= \|f - f_0\|^2 + 2\|f_0\| \|f - f_0\|. \end{aligned}$$

Если для произвольного числа $\varepsilon > 0$ положить

$$\delta = \min \left\{ \frac{\varepsilon}{1 + 2\|f_0\|}, 1 \right\},$$

то при условии $\|f - f_0\| < \delta$ будет справедливо неравенство $\|A(f) - A(f_0)\| < \varepsilon$. Согласно определению 4.6, оператор A является непрерывным. #

Прежде чем ввести следующий важный класс операторов в банаховых пространствах, дадим некоторые определения, связанные с компактными множествами в метрических (нормированных) пространствах.

Определение 4.7. Множество $X \subset M$, где M — метрическое пространство, называют компактным в M , если из всякого бесконечного подмножества этого множества можно выделить последовательность, сходящуюся к некоторому элементу из X .

Отметим, что это определение эквивалентно определению компактного множества в метрическом пространстве, введенному при помощи понятия *открытого покрытия множества* [1]. Множество X в метрическом пространстве M называют *относительно компактным* (или *предкомпактным*) в M , если его замыкание $\overline{X}$ компактно в M .

Определение 4.8. Оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, где $\mathcal{U}, \mathcal{W}$ — нормированные пространства, называют **вполне непрерывным**, если он непрерывен в $\mathcal{U}$ и преобразует любое ограниченное множество из $\mathcal{U}$ в множество, относительно компактное в $\mathcal{W}$.

Иными словами, из любой ограниченной в $\mathcal{U}$ последовательности можно выбрать *подпоследовательность*, которую вполне непрерывный оператор преобразует в сходящуюся к некоторому элементу из $\mathcal{W}$.

Не всякий непрерывный оператор является вполне непрерывным. Так, тождественный оператор $I_{\mathcal{U}}$ вполне непрерывен в нормированном пространстве $\mathcal{U}$ лишь при условии, что пространство $\mathcal{U}$ конечномерно. В таком пространстве любое ограниченное множество относительно компактно [IX]. Поэтому оператор $I_{\mathcal{U}}$ является вполне непрерывным. В бесконечномерном нормированном пространстве $\mathcal{U}$ этот оператор не будет вполне непрерывным, поскольку он преобразует единичный шар (ограниченное множество) в себя, а единичный шар не является относительно компактным в бесконечномерном пространстве. Это утверждение вытекает из следующей теоремы.

Теорема 4.2. Пусть последовательность $\{u_n\}$ образует линейно независимую систему элементов нормированного пространства $\mathcal{U}$, а $\mathcal{U}_n$ — подпространство, которое совпадает с линейной оболочкой системы элементов u_i , $i = \overline{1, n}$. Тогда существует последовательность $\{v_n\}$ элементов $v_n \in \mathcal{U}$, удовлетворяющая условиям: 1) $\|v_n\| = 1$; 2) $v_n \in \mathcal{U}_n$; 3) $\rho(v_n, \mathcal{U}_{n-1}) = \inf_{u \in \mathcal{U}_{n-1}} \|v_n - u\| > 1/2$.

◀ Так как $\{u_n\}$ — линейно независимая система в $\mathcal{U}$, то $u_n \notin \mathcal{U}_{n-1}$, $n \in \mathbb{N}$, и, следовательно [IX],

$$\rho(u_n, \mathcal{U}_{n-1}) = \inf_{u \in \mathcal{U}_{n-1}} \|u_n - u\| = \alpha > 0.$$

В силу свойств точной нижней грани существует такой элемент $u_n^* \in \mathcal{U}_{n-1}$, что

$$\alpha \leq \|u_n - u_n^*\| < 2\alpha. \quad (4.14)$$

Поскольку для любого элемента $u \in \mathcal{U}_{n-1}$ элемент $u_n^* + u \in \mathcal{U}_{n-1}$ ($\mathcal{U}_{n-1}$ — подпространство в $\mathcal{U}$) и $u_n - u_n^* \neq 0$ ($u_n \notin \mathcal{U}_{n-1}$, $u_n^* \in \mathcal{U}_{n-1}$), то

$$\rho(u_n, \mathcal{U}_{n-1}) = \inf_{u \in \mathcal{U}_{n-1}} \|u_n - u_n^* - u\| = \rho(u_n - u_n^*, \mathcal{U}_{n-1}).$$

Следовательно, $\rho(u_n - u_n^*, \mathcal{U}_{n-1}) = \alpha$. Очевидно, что элементы

$$v_1 = \frac{u_1}{\|u_1\|}, \quad v_n = \frac{u_n - u_n^*}{\|u_n - u_n^*\|} \in \mathcal{U}_n, \quad n = 2, 3, \dots,$$

удовлетворяют условиям 1 и 2 теоремы. Учитывая (4.14), проверим выполнение условия 3:

$$\begin{aligned} \rho(v_n, \mathcal{U}_{n-1}) &= \rho\left(\frac{u_n - u_n^*}{\|u_n - u_n^*\|}, \mathcal{U}_{n-1}\right) = \\ &= \frac{\rho(u_n - u_n^*, \mathcal{U}_{n-1})}{\|u_n - u_n^*\|} = \frac{\alpha}{\|u_n - u_n^*\|} > \frac{\alpha}{2\alpha} = \frac{1}{2}, \end{aligned}$$

что завершает доказательство теоремы. ►

Таким образом, согласно теореме 4.2, в единичном шаре бесконечномерного нормированного пространства можно выбрать последовательность $\{v_n\}$, для которой $\rho(v_n, \mathcal{U}_{n-1}) > 1/2$, $n = 2, 3, \dots$, и, значит, никакая ее подпоследовательность не может быть сходящейся (не является фундаментальной). Следовательно, единичный шар в нормированном пространстве

является относительно компактным множеством тогда и только тогда, когда это пространство конечномерно.

Прежде чем привести примеры вполне непрерывных операторов, сформулируем без доказательства* критерии относительной компактности множеств в некоторых нормированных пространствах.

Теорема 4.3 (теорема Арцела).** Для того чтобы множество X в $C[a, b]$ было относительно компактным, необходимо и достаточно, чтобы были выполнены следующие условия:

1) множество X должно быть **равномерно ограниченным множеством функций**, т.е. существует такое число $M > 0$, что для каждой функции $\varphi \in X$ верно соотношение $|\varphi(x)| \leq M$, $x \in [a, b]$;

2) множество X должно быть **равностепенно непрерывным множеством функций**, т.е. для всякого $\varepsilon > 0$ существует такое число $\delta = \delta(\varepsilon) > 0$, что для любой функции $\varphi \in X$ и любых точек $x_1, x_2 \in [a, b]$, таких, что $|x_1 - x_2| < \delta$, верно неравенство $|\varphi(x_1) - \varphi(x_2)| < \varepsilon$.

Для нормированного пространства $L_p[a, b]$ функций f , для которых конечен интеграл Лебега $\int_a^b |f(t)|^p dt$, $p \geq 1$, и норма определена равенством $[IX]$

$$\|f\|_p = \left(\int_a^b |f(t)|^p dt \right)^{1/p},$$

справедлива следующая теорема, доказанная М. Риссом***.

Теорема 4.4. Для относительной компактности в $L_p[a, b]$ множества $X \subset L_p[a, b]$ функций необходимо и достаточно, чтобы были выполнены следующие условия:

*Доказательство см.: Люстерник Л. А., Соболев В. И.

**Ч. Арцелá (1847–1912) — итальянский математик.

***М. Рисс (1886–1969) — шведский математик, венгр по национальности

1) множество X ограничено в $L_p[a, b]$, т.е. существует такое число $M > 0$, что $\|f\| \leq M$, $f \in X$.

2) для всякого $\varepsilon > 0$ найдется такое число $\delta = \delta(\varepsilon) > 0$, что

$$\int_a^b |f(t+h) - f(t)|^p dt < \varepsilon$$

для всех чисел h , таких, что $0 < h < \delta$, и для всех $f \in X$ (считаем, что $f(t) = 0$ для всех $t \in \mathbb{R} \setminus [a, b]$).

Пример 4.8. Покажем, что в $C[a, b]$ вполне непрерывным является оператор $A: C[a, b] \rightarrow C[a, b]$, действующий по правилу

$$\varphi = A(f) \iff \varphi(t) = \int_a^b K(s, t) f(s) ds, \quad t \in [a, b],$$

где $K(s, t)$ — действительная функция двух переменных, непрерывная на замкнутом квадрате $D = \{(s, t) \in \mathbb{R}^2: s, t \in [a, b]\}$. Из свойств интегралов с параметрами [VI] следует непрерывность оператора A . Докажем, что оператор A любое ограниченное множество в $C[a, b]$ переводит в относительно компактное.

Так как функция $K(s, t)$ непрерывна на ограниченном замкнутом множестве $D \subset \mathbb{R}^2$, то она ограничена на этом множестве [V], т.е. $|K(s, t)| \leq C$ для всех точек $(s, t) \in D$, где C — некоторое число. Пусть $X \subset C[a, b]$ — произвольное ограниченное множество непрерывных функций. Тогда для некоторого числа $R > 0$ для всех $f \in X$ верно неравенство $\|f\| = \max_{s \in [a, b]} |f(s)| \leq R$.

Если $\varphi = A(f)$, $f \in X$, то для всех $t \in [a, b]$ имеем

$$|\varphi(t)| \leq \int_a^b |K(s, t)| \max_{s \in [a, b]} |f(s)| ds \leq CR(b-a) = M.$$

Следовательно, условие 1 теоремы 4.3 выполнено и множество $A(X)$, состоящее из образов элементов множества X ,

равномерно ограничено. Докажем равностепенную непрерывность этого множества. Так как ограниченное и замкнутое множество $D \subset \mathbb{R}^2$ является компактным в $\mathbb{R}^2$, а функция $K(s, t)$ непрерывна на D , то она равномерно непрерывна на множестве D [1]. Следовательно, для любого числа $\varepsilon > 0$ найдется такое число $\delta = \delta(\varepsilon) > 0$, что $|K(s, t_1) - K(s, t_2)| < \frac{\varepsilon}{R(b-a)}$ для всех $t_1, t_2 \in [a, b]$, для которых $|t_1 - t_2| < \delta$, и для всех $s \in [a, b]$. Тогда для $\varphi = A(f)$, $f \in X$, имеем

$$\begin{aligned} |\varphi(t_1) - \varphi(t_2)| &= \left| \int_a^b K(s, t_1) f(s) ds - \int_a^b K(s, t_2) f(s) ds \right| \leq \\ &\leq \int_a^b |K(s, t_1) - K(s, t_2)| \max_{s \in [a, b]} |f(s)| ds \leq \int_a^b \frac{\varepsilon}{R(b-a)} R ds = \varepsilon \end{aligned}$$

для всех $t_1, t_2 \in [a, b]$, таких, что $|t_1 - t_2| < \delta$.

Таким образом, множество $A(X)$ равностепенно непрерывно. Условия 1 и 2 теоремы 4.3 для множества $A(X)$ выполнены и, следовательно, $A(X)$ — относительно компактное множество, а оператор A — вполне непрерывный.

Пример 4.9. Покажем, что оператор $A: L_2[a, b] \rightarrow L_2[a, b]$, действующий по правилу

$$\varphi = A(f) \iff \varphi(t) = \int_a^t f(x) dx, \quad t \in [a, b],$$

вполне непрерывен. При любых функциях $f_1, f_2 \in L_2[a, b]$ для $\varphi_1 = A(f_1)$ и $\varphi_2 = A(f_2)$ имеем

$$\begin{aligned} \|\varphi_1 - \varphi_2\|^2 &= \int_a^b \left(\int_a^t f_1(x) dx - \int_a^t f_2(x) dx \right)^2 dt \leq \\ &\leq \int_a^b \left(\int_a^t |f_1(x) - f_2(x)| dx \right)^2 dt. \end{aligned}$$

Применяя далее к внутреннему интегралу неравенство Коши — Буняковского для $L_2[a, b]$ [IX], получаем

$$\begin{aligned}\|\varphi_1 - \varphi_2\|^2 &= \int_a^b \left((t-a) \int_a^t (f_1(x) - f_2(x))^2 dx \right) dt \leq \\ &\leq (b-a)^2 \int_a^b (f_1(x) - f_2(x))^2 dx = (b-a)^2 \|f_1 - f_2\|^2,\end{aligned}$$

что и доказывает непрерывность оператора A . Докажем, что любое ограниченное в $L_2[a, b]$ множество он переводит в относительно компактное.

Пусть X — произвольное ограниченное множество в $L_2[a, b]$, т.е. для некоторого числа $C > 0$ для всех $f \in X$ справедливо неравенство

$$\|f\| = \left(\int_a^b |f(x)|^2 dx \right)^{1/2} \leq C.$$

Если $\varphi = A(f)$, $f \in X$, то имеем

$$\begin{aligned}\|\varphi\| &= \left(\int_a^b \left| \int_a^t f(x) dx \right|^2 dt \right)^{1/2} \leq \\ &\leq \left(\int_a^b \left(\int_a^b |f(x)| dx \right)^2 dt \right)^{1/2} = \sqrt{b-a} \int_a^b |f(x)| dx.\end{aligned}$$

Применяя к последнему интегралу неравенство Коши — Буняковского, получаем

$$\|\varphi\| = (b-a) \left(\int_a^b |f(x)|^2 dx \right)^{1/2} \leq C(b-a).$$

Следовательно, множество $A(X)$ образов элементов множества X ограничено в $L_2[a, b]$, т.е. выполнено условие 1 теоремы 4.4.

Пусть $0 < h < \delta$ и $\varphi = A(f)$, $f \in X$. Тогда, вновь используя неравенство Коши — Буняковского, получаем

$$\begin{aligned} \int_a^b |\varphi(t+h) - \varphi(t)|^2 dt &= \int_a^b \left| \int_a^{t+h} f(x) dx - \int_a^t f(x) dx \right|^2 dt = \\ &= \int_a^b \left| \int_t^{t+h} f(x) dx \right|^2 dt \leq \int_a^b \left(\int_t^{t+h} |f(x)| dx \right)^2 dt \leq \\ &\leq \int_a^b h \left(\int_t^{t+h} |f(x)|^2 dx \right) dt \leq \int_a^b h \left(\int_a^b |f(x)|^2 dx \right) dt \leq (b-a)C^2h. \end{aligned}$$

Если для произвольного $\varepsilon > 0$ выбирать $\delta = \frac{\varepsilon}{C^2(b-a)}$, то условие 2 теоремы 4.4 будет выполнено. Согласно этой теореме, множество $A(X)$ является относительно компактным в $L_2[a, b]$ и, следовательно, оператор A вполне непрерывен. #

Пусть оператор P отображает нормированное пространство $\mathcal{U}$ в себя. Запись $P^m(u)$, $m \in \mathbb{N}$, означает, что оператор P действует последовательно m раз: сначала на элемент $u \in \mathcal{U}$, затем на элемент $P(u)$, потом на $P(P(u)) = P^2(u)$ и, наконец, на $P^{m-1}(u)$, т.е. $P^m = P^{m-1} \circ P$. Таким образом, определен новый оператор $P^m: \mathcal{U} \rightarrow \mathcal{U}$, который называют m -й **степенью оператора**. В случае $m = 0$ полагают, что $P^0 = I_{\mathcal{U}}$.

Рассмотрим оператор $P: \mathcal{U} \rightarrow \mathcal{U}$, удовлетворяющий для некоторого числа $M > 0$ условию

$$\|P(u) - P(v)\|_{\mathcal{U}} \leq M\|u - v\|_{\mathcal{U}}, \quad u, v \in \mathcal{U}.$$

Если это условие верно для некоторого числа $M < 1$, то **оператор P называют сжимающим в $\mathcal{U}$** , а число M — **коэффициентом сжатия**. Для сжимающего оператора P выполняется неравенство $\|P(u) - P(v)\|_{\mathcal{U}} < \|u - v\|_{\mathcal{U}}$, $u, v \in \mathcal{U}$, т.е. расстояние между образами любых элементов (точек) из $\mathcal{U}$ меньше

расстояния между этими элементами. Элемент $u^* \in \mathcal{U}$, удовлетворяющий условию $P(u^*) = u^*$, называют **неподвижной точкой оператора** P в $\mathcal{U}$, т.е. при отображении P образ неподвижной точки совпадает с этой точкой. Любой оператор имеет некоторое множество неподвижных точек, причем это множество может быть пустым. В полном нормированном пространстве сжимающий оператор имеет единственную неподвижную точку [1].

Пример 4.10. Докажем, что уравнение

$$f(x) + \int_0^1 \frac{f(y) dy}{x + \sqrt{1+y^2}} = x^4 \quad (4.15)$$

имеет единственное решение в банаховом пространстве $C[0, 1]$. Для этого покажем, что оператор $P: C[0, 1] \rightarrow C[0, 1]$, действующий по правилу

$$\varphi = P(f) \iff \varphi(x) = x^4 - \int_0^1 \frac{f(y) dy}{x + \sqrt{1+y^2}}, \quad x \in [0, 1],$$

является сжимающим в банаховом пространстве $C[0, 1]$.

Для любых функций $f, g \in C[0, 1]$ имеем

$$\begin{aligned} \|P(f) - P(g)\| &= \max_{x \in [0, 1]} \left| x^4 - \int_0^1 \frac{f(y) dy}{x + \sqrt{1+y^2}} - x^4 + \int_0^1 \frac{g(y) dy}{x + \sqrt{1+y^2}} \right| \leq \\ &\leq \max_{x \in [0, 1]} \int_0^1 \frac{|f(y) - g(y)| dy}{x + \sqrt{1+y^2}} \leq \\ &\leq \max_{y \in [0, 1]} |f(y) - g(y)| \max_{x \in [0, 1]} \int_0^1 \frac{dy}{x + \sqrt{1+y^2}} \leq \\ &\leq \|f - g\| \int_0^1 \frac{dy}{\sqrt{1+y^2}} = \|f - g\| \ln(1 + \sqrt{2}). \end{aligned}$$

Так как $M = \ln(1 + \sqrt{2}) < 1$, то оператор P является сжимающим в банаховом пространстве $C[0, 1]$ и, следовательно, имеет одну неподвижную точку, т.е. (4.15) имеет единственное решение.

4.3. Линейные операторы

Перейдем к изучению линейных операторов в нормированных пространствах. Пусть $\mathcal{U}$ и $\mathcal{W}$ — нормированные пространства. Поскольку нормированное пространство является линейным, то для него остается в силе определение линейного оператора, действующего в линейном пространстве. Напомним, что если оператор $A: \mathcal{U} \rightarrow \mathcal{W}$ линейный, то по определению

$$A(\lambda u + \mu v) = \lambda Au + \mu Av, \quad u, v \in \mathcal{U}, \quad \lambda, \mu \in \mathbb{R}, \quad (4.16)$$

и вместо (4.8) обычно пишут $Au = w$, $w \in \mathcal{W}$. Область значений $R(A) \subset \mathcal{W}$ любого линейного оператора $A: \mathcal{U} \rightarrow \mathcal{W}$ является *линейным многообразием*. Действительно, выберем произвольные $y_1, y_2 \in R(A)$ и $\lambda, \mu \in \mathbb{R}$. Пусть x_1 — прообраз y_1 и x_2 — прообраз y_2 , т.е. $Ax_1 = y_1$ и $Ax_2 = y_2$. Используя (4.16), получаем

$$y = \lambda y_1 + \mu y_2 = \lambda Ax_1 + \mu Ax_2 = A(\lambda x_1 + \mu x_2).$$

Поэтому $y = \lambda y_1 + \mu y_2 \in R(A)$, и в силу произвольного выбора y_1, y_2 и чисел λ, μ заключаем, что $R(A)$ — линейное многообразие.

Линейный оператор $A: \mathcal{U} \rightarrow \mathcal{W}$ отображает нулевой элемент $0 \in \mathcal{U}$ в нулевой элемент $\tilde{0} \in \mathcal{W}$, т.е. $A0 = \tilde{0}$. Оператор $O: \mathcal{U} \rightarrow \mathcal{W}$, отображающий любой элемент $x \in \mathcal{U}$ в элемент $\tilde{0} \in \mathcal{W}$, называют *нулевым*.

Теорема 4.5. Линейный оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, где $\mathcal{U}$ и $\mathcal{W}$ — нормированные пространства и $D(A) = \mathcal{U}$, непрерывен в любой точке $u_0 \in D(A)$, если он непрерывен в точке $0 \in \mathcal{U}$.

◀ Пусть $u \rightarrow u_0$ в нормированном пространстве $\mathcal{U}$. Тогда $v = u - u_0 \rightarrow 0 \in \mathcal{U}$. Поскольку оператор $A: \mathcal{U} \rightarrow \mathcal{W}$ непрерывен в точке 0 , то, согласно определению 4.6, $Av \rightarrow \tilde{0} \in \mathcal{W}$ при $v \rightarrow 0$, и, следовательно, $A(u - u_0) \rightarrow \tilde{0}$ при $u \rightarrow u_0$. В силу линейности оператора A имеем $A(u - u_0) = Au - Au_0$, т.е. $Au - Au_0 \rightarrow \tilde{0}$ при $u \rightarrow u_0$, или $Au \rightarrow Au_0$ при $u \rightarrow u_0$, что, согласно определению 4.6, доказывает утверждение теоремы. ▶

Рассмотрим свойства *ограниченного* линейного оператора.

Теорема 4.6. Линейный оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, где $\mathcal{U}$ и $\mathcal{W}$ — нормированные пространства, является ограниченным тогда и только тогда, когда он ограничен на единичном *замкнутом* шаре

$$\bar{U}(0, 1) = \{u \in \mathcal{U}: \|u\|_{\mathcal{U}} \leq 1\} \quad (4.17)$$

с центром в точке $0 \in \mathcal{U}$, т.е. если существует такое число $M > 0$, что

$$\|Au\|_{\mathcal{W}} \leq M, \quad u \in \bar{U}(0, 1). \quad (4.18)$$

◀ Если линейный оператор A является ограниченным, то, очевидно, он ограничен на шаре $\bar{U}(0, 1)$ (см. определение 4.5). Если же выполнено (4.18), а $X \subset \mathcal{U}$ — произвольное ограниченное множество, то $X \subset \bar{U}(0, R) = \{u \in \mathcal{U}: \|u\|_{\mathcal{U}} \leq R\}$ для некоторого числа $R > 0$. Для любого элемента $u \in X$ имеем $\frac{1}{R}u \in \bar{U}(0, 1)$, а значит,

$$\|Au\|_{\mathcal{W}} = \left\| R A \left(\frac{1}{R} u \right) \right\|_{\mathcal{W}} = R \left\| A \left(\frac{1}{R} u \right) \right\|_{\mathcal{W}} \leq RM.$$

Таким образом, оператор A является ограниченным. ▶

Отметим, что если оператор A не является линейным, то (4.18) нельзя использовать в качестве критерия ограниченности этого оператора.

Теорема 4.7. Линейный оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, где $\mathcal{U}$, $\mathcal{W}$ — нормированные пространства, является ограниченным тогда и

только тогда, когда

$$\|Au\|_{\mathcal{W}} \leq M\|u\|_{\mathcal{U}}, \quad u \in \mathcal{U}, \quad (4.19)$$

где $M > 0$ — некоторое число.

◀ Согласно теореме 4.6, оператор A ограничен тогда и только тогда, когда он ограничен на единичном шаре (4.17). Поэтому докажем эквивалентность утверждений (4.18) и (4.19).

При $u = 0 \in \mathcal{U}$ имеем $Au = \tilde{0} \in \mathcal{W}$, так что выполнение (4.18) и (4.19) очевидно. Пусть $u \neq 0$ и выполнено (4.18). Положим $v = u/\|u\|_{\mathcal{U}}$. Тогда $\|v\|_{\mathcal{U}} = 1$, и поэтому в соответствии с (4.18) $\|Av\|_{\mathcal{W}} \leq M$, или, учитывая линейность оператора и однородность нормы (см. 4.1, аксиома 2 нормы), имеем

$$\|Av\|_{\mathcal{W}} = \left\| A\left(\frac{u}{\|u\|_{\mathcal{U}}}\right) \right\|_{\mathcal{W}} = \left\| \frac{Au}{\|u\|_{\mathcal{U}}} \right\|_{\mathcal{W}} = \frac{\|Au\|_{\mathcal{W}}}{\|u\|_{\mathcal{U}}} \leq M.$$

Отсюда следует (4.19). Обратно, если верно (4.19), то при $\|u\|_{\mathcal{U}} \leq 1$ из (4.19) следует (4.18). ▶

Теорема 4.8. Линейный оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, где $\mathcal{U}, \mathcal{W}$ — нормированные пространства, непрерывен тогда и только тогда, когда он ограничен.

◀ Пусть A непрерывен. Допустим, что он неограничен. Тогда множество элементов $Au \in \mathcal{W}$, где $u \in \bar{U}(0, 1)$, $\bar{U}(0, 1) \subset \mathcal{U}$ — единичный замкнутый шар с центром в точке 0 , неограничено. Поэтому в силу отрицания условия (4.18) для любого $n \in \mathbb{N}$ существует такой элемент $u_n \in \mathcal{U}$, $\|u_n\|_{\mathcal{U}} \leq 1$, что $\|Au_n\|_{\mathcal{W}} \geq n$. Отсюда для элемента $v_n = u_n/n$ имеем как

$$\|Av_n\|_{\mathcal{W}} = \left\| \frac{1}{n} Au_n \right\|_{\mathcal{W}} = \frac{1}{n} \|Au_n\|_{\mathcal{W}} \geq 1,$$

так и $\|v_n\|_{\mathcal{U}} = \frac{1}{n} \|u_n\|_{\mathcal{U}} \leq \frac{1}{n}$. Тогда $\|v_n\| \rightarrow 0$ при $n \rightarrow \infty$, и в силу непрерывности оператора A имеем $\|Av_n\|_{\mathcal{W}} \rightarrow 0$ при $n \rightarrow \infty$, что

противоречит неравенству $\|Av_n\|_W \geq 1$, $n \in \mathbb{N}$. Это доказывает, что непрерывный оператор ограничен.

Пусть теперь A ограничен. Тогда, согласно теореме 4.7, справедливо неравенство (4.19). Следовательно, если $\|u\|_U \rightarrow 0$, то и $\|Au\|_W \rightarrow 0$, а это равносильно тому, что $Au \xrightarrow{W} \tilde{0}$ при $u \xrightarrow{U} 0$, т.е., согласно определению 4.6, оператор A непрерывен в точке $0 \in U$ и поэтому в силу теоремы 4.5 непрерывен. ►

4.4. Линейные ограниченные функционалы

Рассмотрим важный частный случай линейного оператора.

Определение 4.9. Линейный оператор $F: U \rightarrow \mathbb{R}$, определенный на нормированном пространстве U , называют **линейным функционалом**.

Линейный функционал $F: U \rightarrow \mathbb{R}$ является **ограниченным**, если образом любого ограниченного в U множества будет ограниченное числовое множество, или, согласно теореме 4.6, если функционал F переводит единичный замкнутый шар (4.17) в ограниченное числовое множество. В силу общих свойств линейных операторов линейный функционал $F: U \rightarrow \mathbb{R}$ непрерывен в U , если он непрерывен в какой-либо одной точке из U (например, в точке $0 \in U$). Для непрерывности линейного функционала необходимо и достаточно его ограниченности на единичном шаре.

Точную верхнюю грань значений $|Fu| \in \mathbb{R}$ на единичном замкнутом шаре называют **нормой линейного функционала** F , т.е.

$$\|F\| = \sup_{\|u\| \leq 1} |Fu|. \quad (4.20)$$

Так как для линейного функционала $\frac{|Fu|}{\|u\|} = \left| F\left(\frac{u}{\|u\|}\right) \right|$, то

$$\|F\| = \sup_{u \neq 0} \frac{|Fu|}{\|u\|}. \quad (4.21)$$

Действительно, учитывая (4.20), с одной стороны,

$$\sup_{u \neq 0} \frac{|Fu|}{\|u\|} = \sup_{u \neq 0} \left| F\left(\frac{u}{\|u\|}\right) \right| \leq \sup_{\|v\|=1} |Fv| \leq \|F\|.$$

Так как $|Fu| \leq \frac{|Fu|}{\|u\|}$ при $\|u\| \leq 1$, $u \neq 0$, то, с другой стороны,

$$\|F\| = \sup_{\|u\| \leq 1} |Fu| \leq \sup_{\substack{\|u\| \leq 1 \\ u \neq 0}} \frac{|Fu|}{\|u\|} \leq \sup_{u \neq 0} \frac{|Fu|}{\|u\|}.$$

Из (4.21) следует, что $\|F\| \geq \frac{|Fu|}{\|u\|}$, $u \neq 0$, т.е. для любого $u \in \mathcal{U}$ справедливо неравенство

$$|Fu| \leq \|F\| \|u\|. \quad (4.22)$$

Пример 4.11. Покажем, что интеграл

$$I(x) = \int_a^b x(t) dt$$

определяет ограниченный линейный функционал в нормированном пространстве $C[a, b]$.

Поскольку

$$|I(x)| = \left| \int_a^b x(t) dt \right| \leq (b-a) \max_{t \in [a, b]} |x(t)| = (b-a) \|x\|, \quad (4.23)$$

то этот функционал ограничен и, согласно теореме 4.8, непрерывен. С учетом (4.21) и (4.23) имеем, с одной стороны,

$$\|I\| = \sup_{x \neq 0} \frac{|I(x)|}{\|x\|} \leq b-a,$$

а с другой стороны, $|I(x)| = b - a$ при $x(t) \equiv 1$ и $\|I\| \geq b - a$. Следовательно, норма этого функционала $\|I\| = b - a$.

Пример 4.12. Докажем, что интеграл

$$I(x) = \int_{-1}^1 tx(t) dt$$

задает ограниченный линейный функционал в нормированном пространстве $C[-1, 1]$. Пусть $x \in C[-1, 1]$ — произвольная функция, принадлежащая единичному шару, т.е. $\|x\| \leq 1$. Тогда

$$|I(x)| \leq \int_{-1}^1 |tx(t)| dt \leq \|x\| \int_{-1}^1 |t| dt = \|x\| \leq 1.$$

Следовательно, функционал $I(x)$ ограничен, причем в силу (4.20) имеем неравенство $\|I\| \leq 1$.

Заметим, что значение интеграла $I(x)$ от функции $x_0 = x_0(t) = \operatorname{sgn} t$, $t \in [-1, 1]$, равно единице:

$$I(x_0) = \int_{-1}^1 t \operatorname{sgn} t dt = \int_{-1}^1 |t| dt = 1.$$

Однако функция x_0 не принадлежит $C[-1, 1]$. Рассмотрим последовательность непрерывных на отрезке функций, определенных следующим образом (рис. 4.2):

$$x_n(t) = \begin{cases} -1, & t \in [-1, -1/n]; \\ nt, & t \in [-1/n, 1/n]; \\ 1, & t \in [1/n, 1]. \end{cases}$$

Предельной функцией (в смысле поточечной сходимости) на отрезке $[-1, 1]$ для функциональной последовательности $\{x_n(t)\}$

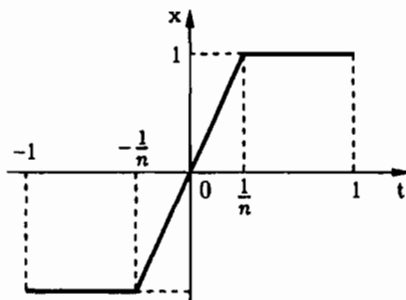


Рис. 4.2

как раз и является функция $x_0(t)$. Поскольку $tx_n(t)$ — четная функция, для любого $n \in \mathbb{N}$ имеем

$$\begin{aligned} I(x_n) &= \int_{-1}^1 tx_n(t) dt = 2 \int_0^1 tx_n(t) dt = 2 \int_0^{1/n} nt^2 dt + 2 \int_{1/n}^1 t dt = \\ &= \frac{2}{3} nt^3 \Big|_0^{1/n} + t^2 \Big|_{1/n}^1 = 1 - \frac{1}{3n^2}. \end{aligned}$$

Так как $|I(x_n)| \rightarrow 1$ при $n \rightarrow \infty$ и $\|x_n\| = 1$, $n \in \mathbb{N}$, то

$$\|I\| = \sup_{\|x\| \leq 1} |I(x)| \geq \sup_{n \in \mathbb{N}} |I(x_n)| = 1.$$

Учитывая полученное выше неравенство $\|I\| \leq 1$, приходим к выводу, что $\|I\| = 1$.

Пример 4.13. Вычислим норму линейного функционала $F: C^1[-1, 1] \rightarrow \mathbb{R}$, заданного следующим образом:

$$F(x) = \int_0^1 x(t) dt - x(-1), \quad x \in C^1[-1, 1].$$

В нормированном пространстве $C^1[-1, 1]$ норма определена равенством (см. пример 4.4).

$$\|x\| = \max \left\{ \max_{t \in [-1, 1]} |x(t)|, \max_{t \in [-1, 1]} |x'(t)| \right\}.$$

Пусть $\|x\| \leq 1$, т.е.

$$\max_{t \in [-1, 1]} |x(t)| \leq 1, \quad \max_{t \in [-1, 1]} |x'(t)| \leq 1.$$

Так как функция $x \in C^1[-1, 1]$ удовлетворяет на отрезке $[-1, t]$, $t \in (-1, 1]$, условиям теоремы Лагранжа [II], то $|x(t) - x(-1)| = |x'(c)| |t + 1|$, где $c \in (-1, t)$. Тогда

$$\begin{aligned} |F(x)| &= \left| \int_0^1 (x(t) - x(-1)) dt \right| \leq \int_0^1 |x'(c)| (t + 1) dt \leq \\ &\leq \int_0^1 \max_{c \in [-1, 1]} |x'(c)| (t + 1) dt \leq \int_0^1 (t + 1) dt = \frac{3}{2}. \end{aligned}$$

Следовательно, в соответствии с (4.20) получаем $\|F\| \leq \frac{3}{2}$.

Для функции $x_0(t) = t$ с нормой $\|x_0\| = 1$ справедливо равенство

$$|F(x_0)| = \int_0^1 t dt + 1 = \frac{3}{2}.$$

Отсюда, согласно (4.20), имеем $\|F\| \geq \frac{3}{2}$. В итоге $\|F\| = \frac{3}{2}$.

Пример 4.14. Интеграл

$$I(x) = \int_{-1}^1 t^2 x(t) dt$$

определяет линейный функционал в нормированном пространстве $L_1[-1, 1]$ суммируемых (интегрируемых по Лебегу) на отрезке $[-1, 1]$ функций с нормой

$$\|x\| = \int_{-1}^1 |x(t)| dt.$$

Найдем норму этого функционала. Если $\|x\| \leq 1$, то

$$|I(x)| \leq \int_{-1}^1 t^2 |x(t)| dt \leq \int_{-1}^1 |x(t)| dt = \|x\| \leq 1.$$

Поэтому в соответствии с (4.20) $\|I\| \leq 1$.

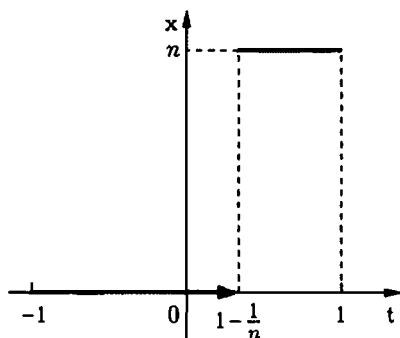


Рис. 4.3

Рассмотрим последовательность $\{x_n\}$ функций вида

$$x_n(t) = \begin{cases} 0, & t \in [-1, 1 - 1/n]; \\ n, & t \in [1 - 1/n, 1], \end{cases}$$

(рис. 4.3). Все эти функции принадлежат $L_1[-1, 1]$. Для каждого $n \in \mathbb{N}$ вычислим норму функции $x_n(t)$:

$$\|x_n\| = \int_{-1}^1 |x_n(t)| dt = \int_{1-1/n}^1 n dt = 1.$$

Так как

$$\begin{aligned}
 |I(x)| &= \left| \int_{-1}^1 t^2 x_n(t) dt \right| = \int_{1-1/n}^1 nt^2 dt = \\
 &= \frac{n}{3} \left(1 - \left(1 - \frac{1}{n} \right)^3 \right) = \frac{n}{3} \cdot \frac{1}{n} \left(3 - \frac{3}{n} + \frac{1}{n^2} \right) \rightarrow 1
 \end{aligned}$$

при $n \rightarrow \infty$, то

$$\|I\| = \sup_{\|x\| \leq 1} |I(x)| \geq \sup_{n \in \mathbb{N}} |I(x_n)| = 1.$$

Учитывая ранее установленное неравенство $\|I\| \leq 1$, получаем $\|I\| = 1$. #

Норма линейного функционала имеет простую геометрическую интерпретацию. Множество элементов нормированного пространства $\mathcal{U}$, удовлетворяющих уравнению $Fu = 1$, назовем **гиперплоскостью**. Расстояние до этой гиперплоскости от нулевого элемента $0 \in \mathcal{U}$ обратно пропорционально норме функционала F , т.е.

$$d = \inf_{Fu=1} \|u\| = \frac{1}{\|F\|}.$$

Действительно, согласно (4.21), имеем

$$\begin{aligned}
 \|F\| &= \sup_{u \neq 0} \frac{|Fu|}{\|u\|} = \sup_{\substack{u \neq 0 \\ Fu \neq 0}} \frac{1}{\left\| \frac{u}{F(u)} \right\|} = \sup_{v = \frac{u}{F(u)}} \frac{1}{\|v\|} = \\
 &= \sup_{Fv=1} \frac{1}{\|v\|} = \frac{1}{\inf_{Fv=1} \|v\|} = \frac{1}{d}.
 \end{aligned}$$

Пусть $\mathcal{U}_0$ — подпространство нормированного пространства $\mathcal{U}$ и F_0 — ограниченный линейный функционал в $\mathcal{U}_0$. Тогда **продолжением** этого **функционала** на все пространство $\mathcal{U}$ называют линейный функционал F , определенный на $\mathcal{U}$

и такой, что $Fu = F_0u$, $u \in \mathcal{U}_0$. Одной из центральных теорем о функционалах в нормированных пространствах является следующая.

Теорема 4.9 (теорема Хана* — Банаха). Любой линейный ограниченный функционал F_0 , определенный на подпространстве $\mathcal{U}_0$ нормированного пространства $\mathcal{U}$, имеет такое продолжение F на все нормированное пространство, что линейный функционал F ограничен в $\mathcal{U}$ и $\|F_0\|_{\mathcal{U}} = \|F\|_{\mathcal{U}}$. #

Доказательство этой теоремы можно найти, например, в учебниках А.Н. Колмогорова и С.В. Фомина или Л.А. Люстерника и В.И. Соболева. Ее геометрический смысл состоит в том, что всякий линейный ограниченный функционал F_0 , определенный на подпространстве $\mathcal{U}_0 \subset \mathcal{U}$, может быть продолжен до линейного ограниченного функционала F , определенного на $\mathcal{U}$ таким образом, что гиперплоскость в $\mathcal{U}$, заданная уравнением $Fu = 1$, будет находиться от точки $0 \in \mathcal{U}$ на том же расстоянии $1/\|F_0\|$, что и гиперплоскость в $\mathcal{U}_0$, заданная уравнением $F_0u = 1$.

4.5. Нормированное пространство линейных операторов

Пусть $A, B, C, \dots$ — линейные ограниченные операторы, областью определения которых является нормированное пространство $\mathcal{U}$, а множество значений лежит в нормированном пространстве $\mathcal{W}$. Множество этих операторов обозначим $\mathcal{L}(\mathcal{U}, \mathcal{W})$ и положим

$$(A + B)u = Au + Bu, \quad (\lambda A)u = \lambda Au, \quad u \in \mathcal{U}, \quad \lambda \in \mathbb{R}. \quad (4.24)$$

Несложно проверить, что операторы $A + B$ и λA являются линейными ограниченными и также принадлежат множеству $\mathcal{L}(\mathcal{U}, \mathcal{W})$. Таким образом, на множестве $\mathcal{L}(\mathcal{U}, \mathcal{W})$ заданы операции сложения элементов и умножения элемента на число, и

*Г. Хан (1879–1934) — австрийский математик

легко убедиться в том, что для этих операций выполнены все аксиомы линейного пространства. Значит, множество операторов $\mathcal{L}(\mathcal{U}, \mathcal{W})$ является линейным пространством [IV]. Отметим, что нулевым элементом этого линейного пространства является нулевой оператор.

Норму в $\mathcal{L}(\mathcal{U}, \mathcal{W})$ введем аналогично тому, как вводили норму для линейных функционалов (см. 4.4). **Нормой линейного оператора** $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ назовем число

$$\|A\| = \sup_{\|u\|_{\mathcal{U}} \leq 1} \|Au\|_{\mathcal{W}}, \quad (4.25)$$

т.е. точную верхнюю грань значений $\|Au\|_{\mathcal{W}}$ на единичном замкнутом шаре $\bar{U}(0, 1)$ (4.17) с центром в точке $0 \in \mathcal{U}$. Как и для линейных функционалов, нетрудно доказать, что

$$\|A\| = \sup_{u \neq 0} \frac{\|Au\|_{\mathcal{W}}}{\|u\|_{\mathcal{U}}}$$

и что нормой оператора A является точная нижняя грань множества всех чисел M , удовлетворяющих (4.19). Отсюда следует справедливость неравенства

$$\|Au\|_{\mathcal{W}} \leq \|A\| \|u\|_{\mathcal{U}}, \quad u \in \mathcal{U}. \quad (4.26)$$

Проверим выполнение аксиом нормы (см. 4.1) в линейном пространстве $\mathcal{L}(\mathcal{U}, \mathcal{W})$. Пусть $\|A\| = 0$. Тогда из (4.26) следует, что $Au = 0$ для всех $u \in \mathcal{U}$, т.е. $A = O$. Верно и обратное: если оператор A нулевой, то для всех $u \in \mathcal{U}$ имеем $\|Au\|_{\mathcal{W}} \equiv 0$ и из (4.25) находим $\|A\| = 0$. Согласно (4.25), имеем

$$\|\alpha A\| = \sup_{\|u\| \leq 1} \|\alpha Au\|_{\mathcal{W}} = |\alpha| \sup_{\|u\| \leq 1} \|Au\|_{\mathcal{W}} = |\alpha| \|A\|.$$

Наконец, используя (4.25), получим **неравенство треугольника** для норм операторов $A, B \in \mathcal{L}(\mathcal{U}, \mathcal{W})$:

$$\begin{aligned} \|A + B\| &= \sup_{\|u\| \leq 1} \|(A + B)u\|_{\mathcal{W}} \leq \\ &\leq \sup_{\|u\| \leq 1} \|Au\|_{\mathcal{W}} + \sup_{\|u\| \leq 1} \|Bu\|_{\mathcal{W}} = \|A\| + \|B\|. \end{aligned}$$

Таким образом, пространство $\mathcal{L}(\mathcal{U}, \mathcal{W})$ линейных ограниченных операторов является нормированным с нормой, определяемой (4.25).

Пример 4.15. Оператор $A: C[a, b] \rightarrow C[a, b]$, рассмотренный в примере 4.8, является линейным ограниченным. Линейность его очевидна, а ограниченность следует из непрерывности (см. теорему 4.8). Найдем норму оператора A , если $K(s, t) = e^{s+t}$, $s, t \in [a, b]$. Тогда оператор A определяется правилом

$$\varphi = Af \iff \varphi(t) = \int_a^b e^{t+s} f(s) ds, \quad t \in [a, b].$$

Пусть f — произвольная непрерывная функция, такая, что $\|f\| = \max_{s \in [a, b]} |f(s)| \leq 1$. Тогда получим

$$\begin{aligned} \|Af\| &= \max_{t \in [a, b]} \left| \int_a^b e^{t+s} f(s) ds \right| = \max_{t \in [a, b]} e^t \left| \int_a^b e^s f(s) ds \right| \leq \\ &\leq e^b \int_a^b e^s \|f\| ds \leq e^b \int_a^b e^s ds = e^b (e^b - e^a). \end{aligned}$$

Следовательно, $\|A\| \leq e^b (e^b - e^a)$. Если $f_0(s) = 1$, $s \in [a, b]$, то

$$\|Af_0\| = \max_{t \in [a, b]} \int_a^b e^{t+s} ds = e^b (e^b - e^a).$$

Так как $\|f_0\| = 1$, то $\|A\| \geq e^b (e^b - e^a)$. В итоге получаем $\|A\| = e^b (e^b - e^a)$.

Теорема 4.10. Если пространство $\mathcal{W}$ банахово, то и пространство $\mathcal{L}(\mathcal{U}, \mathcal{W})$ линейных ограниченных операторов банахово.

◀ Пусть $\{A_n\}$ — фундаментальная последовательность элементов $A_n \in \mathcal{L}(\mathcal{U}, \mathcal{W})$. Тогда, согласно определению 4.1, для любого $\varepsilon > 0$ существует такой номер $N(\varepsilon) \in \mathbb{N}$, что для любых $m > N(\varepsilon)$ и $n > N(\varepsilon)$ выполнено неравенство $\|A_m - A_n\| < \varepsilon$. Выберем произвольный элемент $u \in \mathcal{U}$ и рассмотрим последовательность $\{A_n u\}$ элементов $A_n u \in \mathcal{W}$. Она также фундаментальна в банаховом пространстве $\mathcal{W}$, поскольку, согласно (4.26), для $m > N(\varepsilon)$ и $n > N(\varepsilon)$

$$\|A_m u - A_n u\|_{\mathcal{W}} = \|(A_m - A_n)u\|_{\mathcal{W}} \leq \|A_m - A_n\| \|u\|_{\mathcal{U}} < \varepsilon \|u\|_{\mathcal{U}}.$$

Так как пространство $\mathcal{W}$ банахово, то $\{A_n u\}$ сходится к некоторому элементу w из $\mathcal{W}$. Любому элементу $u \in \mathcal{U}$ поставим в соответствие единственный (в силу единственности предела последовательности) элемент $w \in \mathcal{W}$, который является пределом последовательности $\{A_n u\}$, т.е. определим оператор A , удовлетворяющий равенству

$$\lim_{n \rightarrow \infty} A_n u = Au. \quad (4.27)$$

Оператор A является линейным в силу линейности операторов A_n и свойств предела в банаховых пространствах [IX]. Покажем, что он ограничен. Из неравенства (4.1) имеем

$$|\|A_m\| - \|A_n\|| \leq \|A_m - A_n\|.$$

Следовательно, числовая последовательность $\{\|A_n\|\}$ фундаментальна и, значит, является сходящейся [1]. Поэтому $\{\|A_n\|\}$ ограничена, т.е. существует такое число $c > 0$, что $\|A_n\| \leq c$, $n \in \mathbb{N}$. Отсюда, учитывая (4.26), имеем $\|A_n u\|_{\mathcal{W}} \leq c \|u\|_{\mathcal{U}}$. Предел при $n \rightarrow \infty$ левой части этого неравенства равен $\|Au\|_{\mathcal{W}}$, поскольку в соответствии с (4.1)

$$|\|A_n u\|_{\mathcal{W}} - \|Au\|_{\mathcal{W}}| \leq \|A_n u - Au\|_{\mathcal{W}},$$

а из (4.27) следует, что $\|A_n u - Au\|_{\mathcal{W}} \rightarrow 0$ при $n \rightarrow \infty$. Таким образом, получаем $\|Au\|_{\mathcal{W}} \leq c \|u\|_{\mathcal{U}}$, т.е. оператор A ограничен и $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$.

Докажем, что оператор A является пределом последовательности $\{A_n\}$ в пространстве $\mathcal{L}(\mathcal{U}, \mathcal{W})$, т.е. $\|A_n - A\| \rightarrow 0$ при $n \rightarrow \infty$. Выберем произвольное число $\varepsilon > 0$. Так как последовательность $\{A_n\}$ является фундаментальной в $\mathcal{L}(\mathcal{U}, \mathcal{W})$, то $\|A_n - A_m\| < \varepsilon/2$ для всех натуральных чисел n и m , превышающих некоторое число $N(\varepsilon) \in \mathbb{N}$. Для произвольного элемента $u \in \mathcal{U}$ и $m, n > N$ имеем

$$\|(A_n - A_m)u\|_{\mathcal{W}} \leq \|A_n - A_m\| \|u\|_{\mathcal{U}} < \frac{\varepsilon}{2} \|u\|_{\mathcal{U}}.$$

Так как $\|(A_n - A_m)u - (A_n - A)u\|_{\mathcal{W}} = \|(A_m - A)u\|_{\mathcal{W}} \rightarrow 0$ при $m \rightarrow \infty$, то $\|(A_n - A_m)u\|_{\mathcal{W}} \rightarrow \|(A_n - A)u\|_{\mathcal{W}}$ при $m \rightarrow \infty$. Поэтому, согласно теореме о предельном переходе в неравенстве [I], получаем $\|(A_n - A)u\|_{\mathcal{W}} \leq \frac{\varepsilon}{2} \|u\|_{\mathcal{U}}$. Отсюда находим, что $\|A_n - A\| = \sup_{\|u\|_{\mathcal{U}} \leq 1} \|(A_n - A)u\|_{\mathcal{W}} \leq \frac{\varepsilon}{2} < \varepsilon$. Следовательно,

$A_n \rightarrow A$ при $n \rightarrow \infty$ по норме в $\mathcal{L}(\mathcal{U}, \mathcal{W})$.

Итак, любая фундаментальная последовательность $\{A_n\}$ элементов из $\mathcal{L}(\mathcal{U}, \mathcal{W})$ сходится к элементу этого пространства. Следовательно, $\mathcal{L}(\mathcal{U}, \mathcal{W})$ является банаховым пространством. ►

Последовательность операторов $\{A_n\} \subset \mathcal{L}(\mathcal{U}, \mathcal{W})$ называют **сходящейся поточечно** к оператору A , если для любой точки $u \in \mathcal{U}$ имеем $A_n u \rightarrow Au$ при $n \rightarrow \infty$ по норме в $\mathcal{W}$. Отметим, что если последовательность операторов $\{A_n\} \subset \mathcal{L}(\mathcal{U}, \mathcal{W})$ сходится по норме пространства $\mathcal{L}(\mathcal{U}, \mathcal{W})$ к некоторому оператору $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$, т.е. $\|A_n - A\| \rightarrow 0$ при $n \rightarrow \infty$, то она сходится поточечно к оператору A . Действительно, согласно (4.26),

$$\|A_n u - Au\|_{\mathcal{W}} \leq \|A_n - A\| \|u\|_{\mathcal{U}} \rightarrow 0, \quad n \rightarrow \infty, \quad u \in \mathcal{U}.$$

Обратное утверждение неверно. Например, рассмотрим последовательность операторов $\{A_n\} \subset \mathcal{L}(\mathcal{U}, \mathcal{W})$, $\mathcal{U} = C[0, 1]$, $\mathcal{W} = \mathbb{R}$, причем

$$A_n u = n \int_0^{1/n} u(s) ds, \quad u \in C[0, 1].$$

В силу свойств интеграла с переменным верхним пределом [VI] для любой функции $u \in C[0, 1]$ имеем

$$\begin{aligned}\lim_{n \rightarrow \infty} A_n u &= \lim_{n \rightarrow \infty} n \int_0^{1/n} u(s) ds = \lim_{t \rightarrow 0} \frac{1}{t} \int_0^t u(s) ds = \\ &= \left. \frac{d}{dt} \int_0^t u(s) ds \right|_{t=0} = u(0).\end{aligned}$$

Определим оператор $A: C[0, 1] \rightarrow \mathbb{R}$ равенством $Au = u(0)$, $u \in C[0, 1]$. Нетрудно показать, что $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$. Последовательность $\{A_n\}$ сходится поточечно к оператору A , поскольку $\lim_{n \rightarrow \infty} A_n u = u(0) = Au$, $u \in C[0, 1]$. Однако последовательность $\{A_n\}$ в данном случае не является сходящейся к оператору по норме пространства $\mathcal{L}(\mathcal{U}, \mathcal{W})$, т.е. $\lim_{n \rightarrow \infty} \|A_n - A\| \neq 0$. Действительно, если

$$u_n(s) = \begin{cases} 2n^2 s - 1, & s \in [0, 1/n^2]; \\ 1, & s \in (1/n^2, 1], \end{cases}$$

то $\|u_n\| = \max_{s \in [0, 1]} |u_n(s)| = 1$, и при любом $n \in \mathbb{N}$

$$\begin{aligned}\|A_n - A\| &= \sup_{\|u\| \leq 1} |(A_n - A)u| \geq \\ &\geq |(A_n - A)u_n| = \left| n \int_0^{1/n} u_n(s) ds - u_n(0) \right| = \\ &= n \int_0^{1/n^2} (2n^2 s - 1) ds + n \int_{1/n^2}^{1/n} ds + 1 = 2 - \frac{1}{n} \geq 1.\end{aligned}$$

Следовательно, оператор A не является пределом последовательности $\{A_n\} \subset \mathcal{L}(\mathcal{U}, \mathcal{W})$.

Сформулируем один из основных принципов функционального анализа — **принцип равномерной ограниченности**.

Теорема 4.11 (теорема Банаха — Штейнгауза*). Если последовательность $\{A_n\}$ операторов $A_n \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ ограничена в каждой точке $u \in \mathcal{U}$ банахова пространства $\mathcal{U}$, т.е.

$$\forall u \in \mathcal{U} \exists K_u > 0 \forall n \in \mathbb{N}: \|A_n u\| \leq K_u,$$

то числовая последовательность $\{\|A_n\|\}$ норм этих операторов ограничена.

Доказательство этой теоремы можно найти, например, в учебнике Л.А. Люстерника и В.И. Соболева.

Следствие 4.1. Если последовательность $\{A_n\}$ операторов $A_n \in \mathcal{L}(\mathcal{U}, \mathcal{W})$, где $\mathcal{U}$ — банахово пространство, поточечно сходится к оператору $A: \mathcal{U} \rightarrow \mathcal{W}$, то последовательность $\{\|A_n\|\}$ ограничена.

◀ Так как $A_n u \rightarrow Au$ при $n \rightarrow \infty$, то [IX]

$$\lim_{n \rightarrow \infty} \|A_n u\|_{\mathcal{W}} = \|Au\|_{\mathcal{W}}, \quad u \in \mathcal{U}, \quad (4.28)$$

и числовая последовательность $\{\|A_n u\|\}$ ограничена для любого элемента $u \in \mathcal{U}$. Поэтому в силу теоремы 4.11 последовательность $\{\|A_n\|\}$ ограничена. ▶

Теорема 4.12. Если $\mathcal{U}$ и $\mathcal{W}$ — банаховы пространства, а последовательность операторов $A_n \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ такова, что для каждого элемента $u \in \mathcal{U}$ последовательность $\{A_n u\}$ фундаментальна, то $\{A_n\}$ сходится поточечно к некоторому оператору $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$.

◀ Так как $\{A_n u\}$ — фундаментальная последовательность, а $\mathcal{W}$ — банахово пространство, то для каждого $u \in \mathcal{U}$ существует

*Г. Штейнгауз (1887–1972) — польский математик.

предел $w \in \mathcal{W}$ этой последовательности. Тем самым определен оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, где

$$Au = \lim_{n \rightarrow \infty} A_n u, \quad u \in \mathcal{U}.$$

В силу свойств предела и линейности операторов $A_n \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ оператор A линейный. Согласно следствию 4.1, последовательность $\{\|A_n\|\}$ ограничена, т.е. найдется число $M > 0$, такое, что $\|A_n u\|_{\mathcal{W}} \leq M \|u\|_{\mathcal{U}}$, $u \in \mathcal{U}$, $n \in \mathbb{N}$. Отсюда, переходя к пределу при $n \rightarrow \infty$ и учитывая, что в силу условия $A_n u \rightarrow Au$ при $n \rightarrow \infty$ справедливо (4.28), получаем $\|Au\|_{\mathcal{W}} \leq M \|u\|_{\mathcal{U}}$, $u \in \mathcal{U}$. Следовательно, $\{A_n u\}$ сходится поточечно к оператору $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$. ►

Перейдем к изучению операторов, обратных к линейным ограниченным операторам. Обратный к линейному оператору A является также линейным оператором. Действительно, согласно (4.16), для любых чисел $\alpha_1, \alpha_2 \in \mathbb{R}$

$$A(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 Au_1 + \alpha_2 Au_2, \quad u_1, u_2 \in D(A).$$

Отсюда при $w_1 = Au_1$, $w_2 = Au_2$ и любых $\beta_1, \beta_2 \in \mathbb{R}$ с учетом определения 4.3 имеем

$$\begin{aligned} A^{-1}(\beta_1 w_1 + \beta_2 w_2) &= A^{-1}(\beta_1 Au_1 + \beta_2 Au_2) = \\ &= A^{-1}A(\beta_1 u_1 + \beta_2 u_2) = \beta_1 u_1 + \beta_2 u_2 = \beta_1 A^{-1}w_1 + \beta_2 A^{-1}w_2. \end{aligned}$$

Множество тех элементов $u \in D(A) = \mathcal{U}$, для которых $Au = \tilde{0} \in \mathcal{W}$, называют *ядром оператора* $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ и обозначают $\ker A$. Множество $\ker A$ не пусто, так как $A0 = \tilde{0}$ и $0 \in \ker A$. Нетрудно показать, что множество $\ker A$ является *линейным многообразием*.

Теорема 4.13. Для любого ограниченного линейного оператора $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ ядро $\ker A$ является *подпространством*.

◀ Докажем, что множество $\ker A$ замкнуто в $\mathcal{U}$. Пусть v — предельная точка множества $\ker A$. Тогда существует последовательность $\{v_n\} \subset \ker A$, для которой $v_n \rightarrow v$ при $n \rightarrow \infty$. Так

как A — линейный ограниченный оператор, то, согласно теореме 4.8, он непрерывен. Следовательно, $Av_n \rightarrow Av$ при $n \rightarrow \infty$, но $Av_n = \tilde{0} \in W$, $n \in \mathbb{N}$, и, значит, $Av = \tilde{0}$. Итак, $v \in \ker A$. В силу произвольности предельной точки v многообразие $\ker A$ замкнуто и $\ker A$ — подпространство. ►

Теорема 4.14. Линейный оператор $A: \mathcal{U} \rightarrow \mathcal{W}$, где $\mathcal{U}$ и $\mathcal{W}$ — линейные пространства и $R(A) = \mathcal{W}$, взаимно однозначен тогда и только тогда, когда его ядро состоит лишь из нулевого элемента, т.е. $\ker A = \{0\}$.

◀ Пусть $\ker A = \{0\}$, но допустим, что существует элемент $w \in R(A)$, имеющий два прообраза $u, v \in D(A)$, таких что $u \neq v$. Но тогда $Au = w$ и $Av = w$. Отсюда следует, что $A(u - v) = \tilde{0} \in W$, т.е. существует отличный от нулевого элемент $u - v$, принадлежащий $\ker A$. Это противоречие доказывает взаимную однозначность оператора A при условии $\ker A = \{0\}$.

Пусть теперь оператор A взаимно однозначен. Тогда любой элемент в $\mathcal{W}$ имеет только один прообраз. В частности, элемент $\tilde{0} \in W$ имеет единственный прообраз $0 \in \mathcal{U}$, т.е. ядро $\ker A$ оператора A состоит из одного элемента, $\ker A = \{0\}$. ►

Замечание 4.1. Согласно теореме 4.14 и определению 4.3, для существования у линейного оператора $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$, $R(A) = \mathcal{W}$, обратного оператора A^{-1} необходимо и достаточно, чтобы $\ker A = \{0\}$. #

В функциональном анализе важное место занимает следующая теорема.

Теорема 4.15 (теорема Банаха об обратном операторе). Допустим, что $\mathcal{U}$ и $\mathcal{W}$ — банаховы пространства, а $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ — линейный ограниченный взаимно однозначный оператор, для которого $D(A) = \mathcal{U}$ и $R(A) = \mathcal{W}$. Тогда обратный оператор A^{-1} ограничен. #

Доказательство теоремы можно найти, например, в учебнике А.Н. Колмогорова и С.В. Фомина. Здесь докажем следующую теорему.

Теорема 4.16. Пусть $A \in \mathcal{L}(\mathcal{U}, \mathcal{W})$ — линейный ограниченный оператор и $\mathcal{W} = R(A)$. Тогда для существования и ограниченности в $\mathcal{W}$ оператора A^{-1} , обратного к оператору A , необходимо и достаточно, чтобы для некоторой константы $\tau > 0$ для всех $u \in \mathcal{U}$ было выполнено неравенство

$$\|Au\|_{\mathcal{W}} \geq \tau \|u\|_{\mathcal{U}}. \quad (4.29)$$

◀ **Необходимость.** Пусть оператор A^{-1} существует и ограничен в $\mathcal{W}$. Тогда в соответствии с (4.19) найдется такое число $c > 0$, что $\|A^{-1}v\|_{\mathcal{U}} \leq c\|v\|_{\mathcal{W}}$, $v \in \mathcal{W} = R(A)$. Полагая $v = Au$ и $\tau = 1/c$, приходим к (4.29).

Достаточность. Пусть выполнено (4.29). Тогда $Au = \tilde{0} \in \mathcal{W}$ возможно лишь при $u = 0 \in \mathcal{U}$, т.е. $\ker A = \{0\}$. Поэтому, согласно замечанию 4.1, у оператора A существует обратный оператор A^{-1} . Полагая в (4.29) $u = A^{-1}v$, получаем $\|A^{-1}v\|_{\mathcal{U}} \leq \frac{1}{\tau} \|v\|_{\mathcal{W}}$, т.е. в силу теоремы 4.7 оператор A^{-1} ограничен. ▶

Замечание 4.2. Из (4.26) и хода доказательства теоремы 4.16 следует, что при выполнении условия (4.29) для нормы обратного оператора справедлива оценка

$$\|A^{-1}\| \leq \frac{1}{\tau}. \quad (4.30)$$

Пример 4.16. Найдем решение интегрального уравнения II рода

$$u(t) + \int_0^1 e^{s+t} u(s) ds = w(t), \quad t \in [0, 1], \quad (4.31)$$

где $u, w \in C[0, 1]$ — непрерывные на отрезке $[0, 1]$ неизвестная и заданная функции соответственно. Пусть функция $u \in C[0, 1]$ является решением этого уравнения. Обозначим

$$\int_0^1 e^s u(s) ds = C(u).$$

Покажем, что число $C(u)$ однозначно определено правой частью (4.31). Подставляя функцию $u(t)$ в (4.31), получаем тождество

$$u(t) + C(u)e^t = w(t), \quad t \in [0, 1]. \quad (4.32)$$

Умножим обе его части на функцию e^t и проинтегрируем по отрезку $[0, 1]$:

$$\int_0^1 e^t u(t) dt + C(u) \int_0^1 e^{2t} dt = \int_0^1 e^t w(t) dt.$$

Отсюда находим

$$C(u) + \frac{C(u)}{2}(e^2 - 1) = \int_0^1 e^t w(t) dt,$$

или

$$C(u) = \frac{2}{e^2 + 1} \int_0^1 e^t w(t) dt.$$

Подставив $C(u)$ в (4.32), получим, что любое решение уравнения (4.31) однозначно определено его правой частью:

$$u(t) = w(t) - \frac{2}{e^2 + 1} \int_0^1 e^{t+s} w(s) ds, \quad t \in [0, 1]. \quad (4.33)$$

Нетрудно проверить, что верно и обратное утверждение. Для любой непрерывной на отрезке $[0, 1]$ функции $w(t)$ непрерывная функция $u(t)$, определяемая равенством (4.33), является решением уравнения (4.31).

Заметим, что линейный ограниченный оператор $A: C[0, 1] \rightarrow C[0, 1]$, определяемый в соответствии с (4.31) равенством $Au = w$, имеет линейный ограниченный обратный оператор

$A^{-1}: C[0, 1] \rightarrow C[0, 1]$, определяемый в соответствии с (4.33) равенством $A^{-1}w = u$ (линейность и ограниченность операторов A и A^{-1} несложно проверить). #

Пусть $A \in L = \mathcal{L}(\mathcal{U}, \mathcal{W})$ и $B \in L' = \mathcal{L}(\mathcal{W}, \mathcal{V})$, где $\mathcal{U}, \mathcal{V}, \mathcal{W}$ — нормированные пространства. Композиция линейных операторов $B \circ A$ сохраняет свойства линейности, поскольку для любых элементов $u, v \in \mathcal{U}$ и любых чисел $\lambda, \mu \in \mathbb{R}$

$$(B \circ A)(\lambda u + \mu v) = B(\lambda Au + \mu Av) = \lambda(B \circ A)u + \mu(B \circ A)v.$$

Дважды используя (4.26), для любого элемента $u \in \mathcal{U}$ получаем

$$\|B \circ Au\|_{\mathcal{V}} = \|B(Au)\|_{\mathcal{V}} \leq \|B\|_{L'} \|Au\|_{\mathcal{W}} \leq \|B\|_{L'} \|A\|_{\mathcal{L}} \|u\|_{\mathcal{U}}.$$

Вычисляя при $\|u\|_{\mathcal{U}} \leq 1$ точные верхние грани левой и правой частей этого выражения и учитывая (4.25), приходим к неравенству

$$\|B \circ A\| \leq \|B\|_{L'} \|A\|_{\mathcal{L}} \quad (4.34)$$

для нормы композиции операторов. Таким образом, композиция $B \circ A$ является элементом нормированного пространства $\mathcal{L}(\mathcal{U}, \mathcal{V})$ линейных ограниченных операторов.

Рассмотрим нормированное пространство $\mathcal{L}(\mathcal{U})$ линейных ограниченных (а значит, согласно теореме 4.8, непрерывных) операторов, отображающих нормированное пространство $\mathcal{U}$ в себя. Если $\mathcal{U}$ — банахово пространство, т.е. $\mathcal{U} = \mathcal{B}$, то в силу теоремы 4.10 $\mathcal{L}(\mathcal{B})$ также является банаховым пространством. В $\mathcal{L}(\mathcal{B})$ наряду с операциями (4.24) сложения операторов и умножения оператора на число можно ввести операцию **умножения операторов** как их композицию, отображающую $\mathcal{B}$ в себя. Пусть $A, B \in \mathcal{L}(\mathcal{B})$. Тогда по определению $ABu = (A \circ B)u$, $u \in \mathcal{B}$. В общем случае $ABu \neq BAu$, т.е. умножение операторов в $\mathcal{L}(\mathcal{B})$ не обладает свойством коммутативности. О произведении AB говорят, что оператор B умножен слева на A , или оператор A умножен справа на B . Произведение операторов

$A, B \in \mathcal{L}(B)$, будучи их композицией, ограничено, причем в соответствии с (4.34) имеем

$$\|AB\| \leq \|A\| \|B\| \quad \text{и} \quad \|BA\| \leq \|B\| \|A\| \quad (4.35)$$

(здесь и далее норма $\|\cdot\|$ определена в $\mathcal{L}(B)$). Таким образом, если $A, B \in \mathcal{L}(B)$, то и $AB, BA \in \mathcal{L}(B)$.

Теорема 4.17. Если $A_n, B_n \in \mathcal{L}(B)$, $n \in \mathbb{N}$, и $\lim_{n \rightarrow \infty} A_n = A \in \mathcal{L}(B)$, $\lim_{n \rightarrow \infty} B_n = B \in \mathcal{L}(B)$, то

$$\lim_{n \rightarrow \infty} A_n B_n = AB. \quad (4.36)$$

◀ Используя (4.1) и (4.35), запишем

$$\begin{aligned} \|A_n B_n - AB\| &= \|(A_n B_n - A_n B) + (A_n B - AB)\| \leq \\ &\leq \|A_n B_n - A_n B\| + \|A_n B - AB\| \leq \|A_n\| \|B_n - B\| + \|B\| \|A_n - A\|. \end{aligned}$$

Поскольку последовательность $\{A_n\}$ операторов $A_n \in \mathcal{L}(B)$ сходится в банаховом пространстве $\mathcal{L}(B)$, она ограничена. Следовательно, числовая последовательность $\{\|A_n\|\}$ также ограничена. Так как $\|A_n - A\| \rightarrow 0$ и $\|B_n - B\| \rightarrow 0$ при $n \rightarrow \infty$, то и $\|A_n B_n - AB\| \rightarrow 0$ при $n \rightarrow \infty$, что доказывает утверждение теоремы. ▶

Теорема 4.18. Если $A \in \mathcal{L}(B)$ и $\|A\| \leq q < 1$, то оператор $I - A \in \mathcal{L}(B)$ имеет обратный ограниченный оператор $S = (I - A)^{-1}$, причем

$$\|S\| \leq \frac{1}{1 - q}. \quad (4.37)$$

◀ В банаховом пространстве $\mathcal{L}(B)$ рассмотрим ряд

$$I + A + A^2 + A^3 + \dots + A^n + \dots = \sum_{k=0}^{\infty} A^k. \quad (4.38)$$

В силу (4.35) и условия теоремы для произвольного $k \in \mathbb{N}$ имеем $\|A^k\| \leq \|A\|^k \leq q^k$. Ряд

$$\sum_{k=0}^{\infty} q^k$$

сходится при $q < 1$. Поэтому к некоторому элементу $S \in \mathcal{L}(\mathcal{B})$ сходится и ряд (4.38) [IX].

Для частичной суммы

$$S_n = I + A + A^2 + A^3 + \dots + A^n \quad (4.39)$$

ряда (4.38) находим

$$\begin{aligned} S_n(I - A) &= I + A + A^2 + \dots + A^n - \\ &\quad - A - A^2 - \dots - A^n - A^{n+1} = I - A^{n+1}. \end{aligned}$$

Аналогично $(I - A)S_n = I - A^{n+1}$. Поскольку $\|A^{n+1}\| \leq \|A\|^{n+1} \leq q^{n+1}$, $q < 1$, то $\|A^{n+1}\| \rightarrow 0$ при $n \rightarrow \infty$, а значит, A^{n+1} стремится при $n \rightarrow \infty$ к нулевому оператору $O \in \mathcal{L}(\mathcal{B})$. Переходя в двух последних равенствах к пределу при $n \rightarrow \infty$, получаем $S(I - A) = (I - A)S = I$. В силу теоремы 4.1 заключаем, что оператор $I - A$ имеет обратный оператор, причем $(I - A)^{-1} = S$.

Из (4.39) с учетом (4.1) и формулы для суммы геометрической прогрессии следует

$$\|S_n\| \leq 1 + \|A\| + \|A\|^2 + \dots + \|A\|^n = \frac{1 - \|A\|^{n+1}}{1 - \|A\|}. \quad (4.40)$$

Переходя в (4.40) к пределу при $n \rightarrow \infty$ и учитывая, что $\|A\| \leq q < 1$, приходим к (4.37). ►

Пример 4.17. Рассмотрим интегральное уравнение II рода

$$u(\xi) - \lambda \int_a^b K(\xi, t) u(t) dt = f(\xi), \quad u, f \in C[a, b], \quad (4.41)$$

где $K(\xi, t)$ — **ядро** этого **уравнения**, являющееся функцией, непрерывной на декартовом произведении отрезка $[a, b]$ на себя, а $\lambda \neq 0$ — **параметр** данного **уравнения**. Представим (4.41) в виде операторного уравнения $(I - A)u = f$, где оператор A преобразует искомую функцию $u(\xi)$ в функцию

$$v(\xi) = \lambda \int_a^b K(\xi, t)u(t) dt,$$

т.е. действует из $C[a, b]$ в $C[a, b]$, так как $v(\xi) \in C[a, b]$.

Используя свойства определенного интеграла [VI], для функции $v = Au$ и любой точки $\xi \in [a, b]$ находим

$$\begin{aligned} |v(\xi)| &= \left| \lambda \int_a^b K(\xi, t)u(t) dt \right| \leq |\lambda| \int_a^b |K(\xi, t)u(t)| dt \leq \\ &\leq |\lambda|(b-a) \max_{\xi, t \in [a, b]} |K(\xi, t)| \max_{\xi \in [a, b]} |u(\xi)|. \end{aligned} \quad (4.42)$$

Пусть

$$q = |\lambda|(b-a) \max_{\xi, t \in [a, b]} |K(\xi, t)| < 1. \quad (4.43)$$

Тогда с учетом (4.42) получим

$$\|Au\|_{C[a, b]} = \max_{\xi \in [a, b]} |v(\xi)| \leq q \max_{\xi \in [a, b]} |u(\xi)| = q\|u\|_{C[a, b]}.$$

Отсюда в соответствии с (4.43) и введением нормы оператора при помощи (4.19) следует, что $\|A\| \leq q < 1$.

Таким образом, согласно теореме 4.18, существует оператор S , обратный к оператору $I - A$, т.е. рассматриваемое интегральное уравнение при выполнении условия (4.43) имеет единственное решение $u^0 = Sf$, причем в соответствии с (4.26) и (4.37) $\|u^0\|_{C[a, b]} \leq \|S\| \|f\|_{C[a, b]} \leq \frac{1}{1-q} \|f\|_{C[a, b]}$, где $\|f\|_{C[a, b]} = \max_{\xi \in [a, b]} |f(\xi)|$.

4.6. Спектр линейного оператора

Рассмотрим операторное уравнение

$$Au - \lambda u = f, \quad \text{или} \quad (A - \lambda I)u = f, \quad u, f \in B, \quad (4.44)$$

где A — линейный оператор, отображающий линейное многообразие $D(A)$ банахова пространства B в B , а λ — некоторый числовой параметр. Соответствующее (4.44) *однородное операторное уравнение*

$$Au - \lambda u = 0, \quad \text{или} \quad (A - \lambda I)u = 0 \quad (4.45)$$

всегда имеет *тривиальное решение* $u = 0$.

Пусть для некоторого фиксированного λ оператор $A - \lambda I$ имеет обратный $(A - \lambda I)^{-1} = R_\lambda(A)$, называемый *резольвентным оператором* (или *резольвентой оператора A*). Тогда для этого λ уравнение (4.44) имеет при любом $f \in B$ единственное решение $u = R_\lambda(A)f$, а (4.45) — лишь тривиальное решение $u = 0$. Если в этом случае оператор $R_\lambda(A)$ ограничен, то такое фиксированное λ называют *регулярным значением оператора A* .

Если при фиксированном λ уравнение (4.45) помимо тривиального имеет решение $\bar{u} \in D(A)$, $\bar{u} \neq 0$, то такое λ называют *собственным значением* (или *характеристическим числом*), а $\bar{u}$ — соответствующим этому значению λ *собственным элементом* (или, в случае функциональных пространств, *собственной функцией*) оператора A .

Пусть λ и $\bar{u} \in D(A)$ — некоторое собственное значение оператора A и соответствующий этому λ собственный элемент, а (4.44) имеет при некотором f решение u' , т.е.

$$A\bar{u} - \lambda\bar{u} = 0, \quad Au' - \lambda u' = f.$$

Складывая почленно эти два соотношения, получаем $A(\bar{u} + u') - \lambda(\bar{u} + u') = f$, т.е. $\bar{u} + u'$ также является решением (4.44).

Отсюда видно, что уравнение (4.44) в случае собственного значения λ имеет несколько различных решений.

Совокупность всех значений λ , не являющихся регулярными, называют **спектром оператора** A . В частности, все собственные значения принадлежат спектру. Действительно, если λ — собственное значение оператора A , то **ядро оператора** $A - \lambda I$ содержит наряду с нулевым элементом $0 \in D(A)$ по крайней мере один собственный элемент $\bar{u} \in D(A)$, соответствующий этому собственному значению, т.е. $\ker(A - \lambda I) \neq \{0\}$. Но тогда, согласно теореме 4.14, не существует оператора, обратного к оператору $A - \lambda I$, и λ не является регулярным значением оператора A .

Отметим, что в конечномерном банаховом пространстве спектр любого линейного оператора состоит только из собственных значений. Если линейный оператор A определен в бесконечномерном банаховом пространстве, то его спектр может включать:

1) **точечный спектр**, состоящий из множества собственных значений оператора A ;

2) **непрерывный спектр**, содержащий те значения λ , для которых на множестве $Y \subset B$ значений оператора $A - \lambda I$ определен обратный оператор $(A - \lambda I)^{-1}$, причем $Y \neq B$, но замыкание $\bar{Y} = B$;

3) **остаточный спектр**, состоящий из тех значений λ , для которых оператор $(A - \lambda I)^{-1}$ определен на множестве $Y \subset B$, причем $\bar{Y} \neq B$.

В случаях 2 и 3 оператор $(A - \lambda I)^{-1}$ может быть неограниченным. Таким образом, спектр линейного оператора является объединением трех непересекающихся множеств: точечного, непрерывного и остаточного спектров.

Теорема 4.19. Если A — линейный ограниченный оператор в банаховом пространстве B и $|\lambda| > \|A\|$, то λ — регулярное значение оператора A .

◀ Из условия теоремы ясно, что $\lambda \neq 0$ и $\frac{\|A\|}{|\lambda|} = q < 1$. Оператор $A - \lambda I = -\lambda \left(I - \frac{A}{\lambda} \right)$ в силу теоремы 4.18 имеет обратный ограниченный оператор и при этом

$$R_\lambda(A) = -\frac{1}{\lambda} \left(I + \frac{A}{\lambda} + \frac{A^2}{\lambda^2} + \frac{A^3}{\lambda^3} + \dots \right), \quad (4.46)$$

что доказывает утверждение теоремы. ▶

Следствие 4.2. Если $A \in \mathcal{L}(\mathcal{B})$ и λ — точка спектра оператора A , то $|\lambda| \leq \|A\|$, т.е. точки спектра оператора A принадлежат кругу радиуса $\|A\|$ с центром в нуле.

Теорема 4.20. Регулярные значения линейного ограниченного оператора $A \in \mathcal{L}(\mathcal{B})$, действующего в банаховом пространстве $\mathcal{B}$, образуют открытое числовое множество.

◀ Если λ — регулярное значение оператора A , то существует ограниченный резольвентный оператор $R_\lambda(A) = (A - \lambda I)^{-1}$, определенный на всем банаховом пространстве $\mathcal{B}$. Пусть $|\Delta\lambda| < (\|R_\lambda(A)\|)^{-1}$. Рассмотрим оператор $A - (\lambda + \Delta\lambda)I$, записав его в виде

$$A - \lambda I - \Delta\lambda(A - \lambda I)R_\lambda(A) = (A - \lambda I)(I - \Delta\lambda R_\lambda(A)).$$

Для того чтобы существовал ограниченный обратный оператор $(A - (\lambda + \Delta\lambda)I)^{-1}$, достаточно, чтобы существовал ограниченный обратный оператор $(I - \Delta\lambda R_\lambda(A))^{-1}$. Согласно теореме 4.18, ограниченный обратный оператор $(I - \Delta\lambda R_\lambda(A))^{-1}$ существует, поскольку из условия $|\Delta\lambda| < (\|R_\lambda(A)\|)^{-1}$ следует, что $\|\Delta\lambda R_\lambda(A)\| < 1$. Таким образом, вместе с λ совокупность регулярных значений оператора A включает окрестность λ радиуса $(\|R_\lambda(A)\|)^{-1}$, т.е. совокупность регулярных значений — открытое числовое множество. ▶

Следствие 4.3. Спектр линейного ограниченного оператора $A \in \mathcal{L}(B)$, действующего в банаховом пространстве B , является замкнутым числовым множеством.

◀ Спектр оператора $A \in \mathcal{L}(B)$ является дополнением множества регулярных значений, которое, согласно теореме 4.20, открыто в $\mathbb{R}$ или в $\mathbb{C}$. Дополнение же открытого множества всегда замкнуто [1]. ▶

Пример 4.18. Пусть в банаховом пространстве $C[-\pi, \pi]$ оператор A определен равенством $Au = u(-x)$, $x \in [-\pi, \pi]$. Тогда операторному уравнению $Au - \lambda u = 0$ помимо функции $u(x) \equiv 0$ при $\lambda_1 = 1$ будут удовлетворять все четные функции $u(x) = u(-x)$, принадлежащие $C[-\pi, \pi]$, а при $\lambda_2 = -1$, — все нечетные функции $u(x) = -u(-x)$ из $C[-\pi, \pi]$. Следовательно, $\lambda_1 = 1$ и $\lambda_2 = -1$ являются собственными значениями рассматриваемого оператора и принадлежат его спектру.

Покажем, что спектр оператора A содержит лишь два собственных значения $\lambda_1 = 1$ и $\lambda_2 = -1$ и, следовательно, совпадает с точечным спектром. Пусть $g \in C[-\pi, \pi]$ — произвольная функция из $C[-\pi, \pi]$. Операторное уравнение $Au - \lambda u = g$ эквивалентно уравнению $u(-x) - \lambda u(x) = g(x)$, $x \in [-\pi, \pi]$, или уравнению $u(x) - \lambda u(-x) = g(-x)$, $x \in [-\pi, \pi]$. Умножая последнее уравнение на λ и складывая с последним, получаем

$$(1 - \lambda^2)u(x) = g(-x) + \lambda g(x), \quad x \in [-\pi, \pi].$$

Если $\lambda \neq \pm 1$, то непосредственной проверкой можно установить, что резольвентный оператор $R_\lambda(A)$ определен равенством

$$R_\lambda(A)g = (A - \lambda I)^{-1}g = \frac{g(-x) + \lambda g(x)}{1 - \lambda^2}$$

и является ограниченным в $C[-\pi, \pi]$. Следовательно, любое число $\lambda \neq \pm 1$ является регулярным значением оператора A .

Пример 4.19. В банаховом пространстве $C[0, 1]$ равенством $Au = xu(x)$ определен линейный ограниченный оператор A .

Операторное уравнение $Au - \lambda u = f$, где $f \in C[0, 1]$ — заданная функция, эквивалентно уравнению $xu(x) - \lambda u(x) = f(x)$, $x \in [0, 1]$, где искомой является непрерывная на отрезке $[0, 1]$ функция $u(x)$. При $\lambda \in \mathbb{R} \setminus [0, 1]$ решение этого уравнения существует и единственно для любой функции $f \in C[0, 1]$ и определено равенством $u(x) = \frac{f(x)}{x - \lambda}$. Таким образом, при $\lambda \in \mathbb{R} \setminus [0, 1]$ существует ограниченный резольвентный оператор $R_\lambda(A)$, и λ является регулярным значением оператора A .

Пусть $\lambda \in [0, 1]$. Если непрерывная на отрезке $[0, 1]$ функция $u(x)$ является решением уравнения $(x - \lambda)u(x) = f(x)$, $x \in [0, 1]$, то при $x = \lambda$ левая часть уравнения обращается в нуль. Следовательно, необходимым условием существования решения в этом случае является $f(\lambda) = 0$. Таким образом, при $\lambda \in [0, 1]$ резольвентный оператор $R_\lambda(A)$ определен на множестве $Y_\lambda \subset C[0, 1]$, состоящем из функций, равных нулю в точке $x = \lambda$. Кроме того, можно показать, что оператор $R_\lambda(A)$ не ограничен на Y_λ .

Итак, отрезок $[0, 1]$ является спектром оператора A , причем остаточным, поскольку $\overline{Y_\lambda} \neq C[0, 1]$.

Пример 4.20. Дифференциальный оператор $A = -\frac{d^2}{dx^2}$ определен на линейном многообразии $D(A)$ банахова пространства $C[0, 1]$, состоящем из функций $u(x)$, дважды непрерывно дифференцируемых на отрезке $[0, 1]$ и принимающих значения $u(0) = u(1) = 0$. Операторное уравнение $Au - \lambda u = 0$ является линейным однородным обыкновенным дифференциальным уравнением (ОДУ)

$$u'' + \lambda u = 0 \quad (4.47)$$

с постоянными коэффициентами. Если $\lambda = 0$, то общее решение этого ОДУ имеет вид $u(x) = C_1 + C_2x$. Из условий $u(0) = u(1) = 0$ получаем $C_1 = C_2 = 0$ и $u(x) \equiv 0$, $x \in [0, 1]$. Следовательно, $\lambda = 0$ не является собственным значением оператора A .

Если $\lambda < 0$, то характеристическое уравнение для (4.47) имеет корни $\pm\sqrt{-\lambda} \in \mathbb{R}$, а общее решение ОДУ (4.47) примет

вид

$$u(x) = C_1^* \exp(x\sqrt{-\lambda}) + C_2^* \exp(-x\sqrt{-\lambda}).$$

Так как $u(0) = u(1) = 0$, то $C_1^* + C_2^* = 0$ и $C_1^* \exp(\sqrt{-\lambda}) + C_2^* \exp(-\sqrt{-\lambda}) = 0$. Отсюда находим $C_1^* = C_2^* = 0$, так что $\lambda < 0$ не является собственным значением оператора A .

Если $\lambda > 0$, то, решая для (4.47) характеристическое уравнение, находим его комплексно сопряженные корни $\pm i\omega$, где $\omega > 0$ и $\omega^2 = \lambda$. В этом случае общее решение ОДУ (4.47) имеет вид $u(x) = \tilde{C}_1 \cos \omega x + \tilde{C}_2 \sin \omega x$. Из условий $u(0) = u(1) = 0$ получаем $\tilde{C}_1 = 0$ и $\tilde{C}_1 \cos \omega + \tilde{C}_2 \sin \omega = 0$. Следовательно,

$$\tilde{C}_2 \sin \omega = 0. \quad (4.48)$$

Если в (4.48) принять $\tilde{C}_2 = 0$, то придем к тривиальному решению $u(x) \equiv 0$ ОДУ (4.47). Если же положить $\tilde{C}_2 \neq 0$, то из (4.48) получим $\sin \omega = 0$, т.е. $\omega = k\pi$, $k \in \mathbb{N}$, поскольку $\omega > 0$. Таким образом, имеем бесконечное множество нетривиальных решений

$$u_k(x) = \tilde{C}_2 \sin k\pi x, \quad \tilde{C}_2 \in \mathbb{R} \setminus \{0\}, \quad (4.49)$$

уравнения $Au - \lambda u = 0$ при $\lambda = (k\pi)^2$, $k \in \mathbb{N}$.

Итак, $\lambda_k = (k\pi)^2$, $k \in \mathbb{N}$, являются собственными значениями оператора A , причем каждому из них соответствует одномерное подпространство, состоящее из собственных элементов (собственных функций) $u_k(x)$, определяемых соотношением (4.49), и нулевой функции $u(x) \equiv 0$. #

Рассмотрим спектр *вполне непрерывного* линейного оператора A , действующего в бесконечномерном банаховом пространстве B . Такой оператор, будучи непрерывным, в силу теоремы 4.8 является и ограниченным. Отметим, что для такого оператора нуль будет точкой спектра. Действительно, в противном случае существует непрерывный линейный оператор $A^{-1}: B \rightarrow B$, обратный к A . Тогда, поскольку A и A^{-1} — непрерывные отображения, оператор A открытый единичный

шар переведет в открытое множество из B , обязательно содержащее некоторый замкнутый шар положительного радиуса. Так как оператор A вполне непрерывен, то этот замкнутый шар является относительно компактным, из чего следует конечномерность B (см. 4.2), а это противоречит первоначальному условию.

Таким образом, точка $0 \in \mathbb{R}$ принадлежит спектру вполне непрерывного оператора A . Выясним некоторые свойства точечного спектра такого оператора.

Теорема 4.21. *Вполне непрерывный линейный оператор A в банаховом пространстве B при любом $\delta > 0$ имеет лишь конечное число линейно независимых собственных элементов, отвечающих собственным значениям, превышающим по модулю δ .*

◀ Предположим противное. Пусть для некоторого $\delta > 0$ существует последовательность линейно независимых собственных элементов y_n линейного оператора A , для которых $Ay_n = \lambda_n y_n$ и $|\lambda_n| > \delta$, $n \in \mathbb{N}$. Согласно теореме 4.2, найдется последовательность $\{v_n\} \subset B$, такая, что для любого номера $n \in \mathbb{N}$ справедливы утверждения

$$\|v_n\| = 1; \quad v_n \in B_n; \quad \rho(v_n, B_{n-1}) = \inf_{y \in B_{n-1}} \|v_n - y\| > \frac{1}{2}, \quad (4.50)$$

где B_n — линейная оболочка элементов y_k , $k = \overline{1, n}$. Так как $|\lambda_n| > \delta$, то $\left\| \frac{v_n}{\lambda_n} \right\| = \frac{\|v_n\|}{|\lambda_n|} < \frac{1}{\delta}$, т.е. последовательность $\left\{ \frac{v_n}{\lambda_n} \right\}$ ограничена в B . Рассмотрим последовательность $\{u_n\}$, где $u_n = A\left(\frac{v_n}{\lambda_n}\right)$, $n \in \mathbb{N}$. Так как $v_n \in B_n$, а элементы $y_1, y_2, \dots, y_n$ составляют базис в B_n , то

$$v_n = \sum_{k=1}^n \alpha_k^{(n)} y_k.$$

Кроме того, учитывая соотношения $Ay_k = \lambda_k y_k$, $k = \overline{1, n}$, имеем

$$\begin{aligned} u_n &= A\left(\frac{v_n}{\lambda_n}\right) = \sum_{k=1}^n \frac{\alpha_k^{(n)}}{\lambda_n} A y_k = \\ &= \sum_{k=1}^n \frac{\alpha_k^{(n)}}{\lambda_n} \lambda_k y_k - v_n + v_n = \sum_{k=1}^{n-1} \alpha_k^{(n)} \left(\frac{\lambda_k}{\lambda_n} - 1\right) y_k + v_n. \end{aligned}$$

Поэтому для любого номера $n \in \mathbb{N}$ справедливо представление $u_n = v_n + z_n$, где

$$z_n = \sum_{k=1}^{n-1} \alpha_k^{(n)} \left(\frac{\lambda_k}{\lambda_n} - 1\right) y_k \in \mathcal{B}_{n-1}.$$

Пусть l и m , $l < m$, выбраны произвольно. В силу разложений $u_m = v_m + z_m$, где $z_m \in \mathcal{B}_{m-1}$, и $u_l = v_l + z_l$, где $z_l \in \mathcal{B}_{l-1} \subset \mathcal{B}_{m-1}$, имеем $y = v_l + z_l - z_m \in \mathcal{B}_{m-1}$. Отсюда, учитывая условия (4.50), приходим к соотношениям

$$\|u_m - u_l\| = \|v_m + z_m - v_l - z_l\| = \|v_m - y\| > \frac{1}{2}.$$

Итак, из последовательности $\{u_n\}$, являющейся при действии вполне непрерывного оператора A образом ограниченной последовательности $\{v_n\}$, нельзя выбрать сходящуюся подпоследовательность, т.е. пришли к противоречию, что доказывает теорему. ►

Следствие 4.4. Число линейно независимых собственных элементов, отвечающих данному собственному значению $\lambda \neq 0$ вполне непрерывного оператора $A \in \mathcal{L}(\mathcal{B})$, конечно.

◄ Достаточно применить теорему 4.21 для $\delta = \frac{|\lambda|}{2}$. ►

Следствие 4.5. Для любого вполне непрерывного оператора $A \in \mathcal{L}(\mathcal{B})$ число собственных значений конечно или счетно.

Все собственные значения можно представить в виде конечной или бесконечной последовательности $\{\lambda_n\}$, причем

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \geq |\lambda_{n+1}| \geq \dots$$

(собственные значения в этой последовательности повторяются столько раз, какова размерность собственного подпространства). Если $\{\lambda_n\}$ бесконечна, то $\lambda_n \rightarrow 0$ при $n \rightarrow \infty$. #

Спектр вполне непрерывного линейного оператора в банаховом пространстве состоит только из точечного спектра и точки нуль, которая может принадлежать спектру любого вида. Это утверждение следует из теоремы, которую называют **альтернативой Фредгольма***.

Теорема 4.22. Пусть B - банахово пространство, $A: B \rightarrow B$ - линейный вполне непрерывный оператор и $\lambda \neq 0$. Тогда для операторных уравнений $Au - \lambda u = f$ и $Au - \lambda u = 0$ имеет место одна из следующих возможностей:

1) однородное уравнение имеет только тривиальное решение, а λ является регулярным значением оператора A . В этом случае неоднородное уравнение для любого элемента $f \in B$ имеет решение $u = (A - \lambda I)^{-1}f$ и притом единственное;

2) однородное уравнение имеет хотя бы одно нетривиальное решение. В этом случае неоднородное уравнение либо не имеет решения, либо имеет более одного.

Пример 4.21. Найдем спектр оператора $A: C[0, 1] \rightarrow C[0, 1]$, действующего по правилу

$$\varphi = A(u) \iff \varphi(t) = \int_0^1 ts^2 u(s) ds, \quad t \in [0, 1].$$

Согласно выводам примеров 4.8 и 4.15, оператор A является вполне непрерывным линейным оператором.

*Э.И. Фредгольм (1866–1927) — шведский математик.

Если $\lambda \neq 0$ является собственным значением оператора A , то существует отличное от нулевого решение $u \in C[0, 1]$ уравнения

$$t \int_0^1 s^2 u(s) ds = \lambda u(t), \quad t \in [0, 1]. \quad (4.51)$$

Этим решением будет собственная функция, отвечающая λ и имеющая вид $u(t) = \frac{at}{\lambda}$, где $a \neq 0$, что следует из (4.51), если обозначить

$$a = \int_0^1 s^2 u(s) ds.$$

Подставляя $u(t) = \frac{at}{\lambda}$ в последний интеграл, получаем уравнение относительно a и λ :

$$a = \int_0^1 \frac{a}{\lambda} s^3 ds.$$

Решая его, находим $a \left(1 - \frac{1}{4\lambda}\right) = 0$. Так как $a \neq 0$, то $\lambda = 1/4$. Если функцию $u(t) = t$ подставить в (4.51) при $\lambda = 1/4$, то получим верное тождество.

Таким образом, единственное собственное значение оператора A , отличное от нуля, равно $1/4$, а собственные функции, отвечающие этому значению, имеют вид Ct , где $C \neq 0$, и вместе с нулевой функцией образуют одномерное подпространство в $C[0, 1]$. Следовательно, спектр оператора A состоит из двух значений 0 и $1/4$, принадлежащих точечному спектру. Собственному значению $\lambda = 0$ отвечает любая ненулевая функция $u(t)$, $t \in [0, 1]$, для которой

$$\int_0^1 t^2 u(t) dt = 0,$$

например, $u(t) = 4t - 3$.

4.7. Пополнение нормированного пространства

Если последовательность $\{u_n\}$ элементов банахова пространства B такова, что любой элемент $u \in B$ можно единственным образом представить в виде сходящегося ряда

$$u = \sum_{n=1}^{\infty} a_n u_n, \quad a_n \in \mathbb{R}, \quad (4.52)$$

то последовательность $\{u_n\}$ называют *счетным базисом* в B . При этом равенство (4.52) называют *разложением элемента u по базису $\{u_n\}$* , а коэффициенты a_n в этом равенстве — *координатами элемента u в данном базисе*.

Счетный базис является последовательностью линейно независимых элементов. Банахово пространство со счетным базисом сепарабельно [IX], но не всякое сепарабельное пространство имеет счетный базис.

Обсуждение *математических моделей* физических процессов в части I показывает, что в этих моделях, как правило, приходится иметь дело с решением *операторного уравнения* вида

$$P(u) = f, \quad (4.53)$$

где P — некоторый оператор в банаховом пространстве B , $u \in D(P)$, $f \in R(P)$. На практике элементами банахова пространства являются функции $u(x)$, определенные в некоторой области $V \subset \mathbb{R}^m$ и удовлетворяющие краевым условиям на границе ∂V этой области. Большинство приближенных методов решения подобных уравнений основано на построении такой последовательности $\{\tilde{u}_N\}$, что

$$\lim_{N \rightarrow \infty} P(\tilde{u}_N) = f. \quad (4.54)$$

При этом изучают вопросы, имеет ли последовательность $\{\tilde{u}_N\}$ предел, принадлежит ли этот предел $D(P)$ и является ли он решением уравнения $P(u) = f$.

Учитывая (4.52), можно положить

$$\tilde{u}_N = \sum_{n=1}^N a_n(N) u_n, \quad a_n(N) \in \mathbb{R}, \quad (4.55)$$

где функции $u_n \in \mathcal{B}$, $n = \overline{1, N}$, являются элементами счетного базиса $\{u_n\}$ в банаховом пространстве $\mathcal{B}$.

Если область определения $D(P)$ оператора P является подпространством и $u_n \in D(P)$, $n \in \mathbb{N}$, то при выполнении условия (4.54) в некоторых случаях (см. 5.2) можно ожидать, что последовательность $\{\tilde{u}_N\}$ будет сходиться к искомой функции $u_0 \in D(P)$, удовлетворяющей (4.53). Но если $u_n \notin D(P)$, то и $\tilde{u}_N \notin D(P)$, так что выражение $P(\tilde{u}_N)$ в (4.54) не будет определено. Поэтому следует использовать такие функции $u_n \in D(P)$, $n \in \mathbb{N}$, которые бы составляли счетный базис $\{u_n\}$ в $D(P)$.

Однако $D(P)$ может и не быть подпространством. Тогда последовательность $\{\tilde{u}_N\} \subset D(P)$ может сходиться к элементу $u \notin D(P)$, не являющемуся решением уравнения (4.53) в обычном смысле. Кроме того, (4.53) не имеет решения в $D(P)$, если $f \notin R(P)$, что характерно для прикладных задач.

Эти трудности можно преодолеть, расширив понятие решения операторного уравнения. Такое **решение**, называемое **обобщенным** в отличие от **классического решения** $u_0 \in D(P)$, рассматривают в пространстве, получаемом расширением области $D(P)$ определения оператора P , пополняя ее некоторыми особыми элементами.

Напомним, что *линейные пространства* D и D' называют изоморфными, если существует такое взаимно однозначное отображение $\varphi: D \rightarrow D'$, что $\varphi(\lambda u + \mu v) = \lambda \varphi(u) + \mu \varphi(v)$, $\lambda, \mu \in \mathbb{R}$, $u, v \in D$. Такое отображение называют изоморфизмом этих пространств.

Определение 4.10. *Нормированные пространства* D и D' называют **изометричными** и говорят, что D' изометрично D (и наоборот), если существует их изоморфизм φ как

линейных пространств, удовлетворяющий условию $\|\varphi(u)\|_{D'} = \|u\|_D$, $u \in D$. При этом изоморфизм φ называют **изометрией**.

Определение 4.11. Банахово пространство B называют **пополнением нормированного пространства D** , если существует линейное многообразие $D' \subset B$, изометричное D и являющееся множеством, всюду плотным в B ($\overline{D'} = B$).

Теорема 4.23. Любое нормированное пространство D имеет некоторое пополнение B . Любые два пополнения B_1 и B_2 нормированного пространства D изометричны.

◀ Две фундаментальные последовательности $\{u_n\}$ и $\{u'_n\}$ элементов из D назовем эквивалентными и будем писать $\{u_n\} \sim \{u'_n\}$, если

$$\lim_{n \rightarrow \infty} \|u_n - u'_n\|_D = 0. \quad (4.56)$$

Для отношения $\sim$ из (4.56) следуют свойства *рефлексивности* ($\{u_n\} \sim \{u_n\}$, так как $\|u_n - u_n\|_D \equiv 0$), *симметричности* (если $\{u_n\} \sim \{u'_n\}$, то $\{u'_n\} \sim \{u_n\}$ в силу $\|u_n - u'_n\|_D = \|u'_n - u_n\|_D$) и *транзитивности*: если $\{u_n\} \sim \{u'_n\}$ и $\{u'_n\} \sim \{u''_n\}$, то $\{u_n\} \sim \{u''_n\}$, так как из (4.1) имеем

$$\begin{aligned} 0 \leq \|u_n - u''_n\|_D &= \|(u_n - u'_n) + (u'_n - u''_n)\|_D \leq \\ &\leq \|u_n - u'_n\|_D + \|u'_n - u''_n\|_D \end{aligned}$$

и после перехода в этом неравенстве к пределу при $n \rightarrow \infty$ получаем

$$\lim_{n \rightarrow \infty} \|u_n - u''_n\|_D = 0.$$

Следовательно, отношение $\sim$ является отношением *эквивалентности*. Поэтому множество всех фундаментальных последовательностей элементов из D распадается на непересекающиеся подмножества, каждое из которых составляет некоторый класс эквивалентности. Множество классов эквивалентности фундаментальных последовательностей из D обозначим B и

покажем, что его можно наделить структурой банахова пространства.

Начнем с определения в B операции сложения. Пусть $\tilde{u}$ и $\tilde{v}$ — классы эквивалентности по отношению, введенному в D . Выберем из этих классов эквивалентности фундаментальные в D последовательности $\{u_n\} \in \tilde{u}$ и $\{v_n\} \in \tilde{v}$. Нетрудно проверить, что последовательность $\{u_n + v_n\}$ также фундаментальна в D . Поэтому она входит в некоторый класс эквивалентности, который обозначим $\tilde{u} + \tilde{v}$. Покажем, что определение этого класса корректно, т.е. оно не зависит от выбора последовательностей из классов $\tilde{u}$ и $\tilde{v}$. Пусть $\{u_n\} \sim \{u'_n\}$ и $\{v_n\} \sim \{v'_n\}$. Докажем, что $\{u_n + v_n\} \sim \{u'_n + v'_n\}$. Наряду с (4.56) будет верно и

$$\lim_{n \rightarrow \infty} \|v_n - v'_n\|_D = 0. \quad (4.57)$$

Используя (4.1), получаем

$$\begin{aligned} 0 \leq \| (u_n + v_n) - (u'_n + v'_n) \|_D &= \| (u_n - u'_n) + (v_n - v'_n) \|_D \leq \\ &\leq \| u_n - u'_n \|_D + \| v_n - v'_n \|_D. \end{aligned}$$

Переходя в этом неравенстве к пределу при $n \rightarrow \infty$ и учитывая (4.56) и (4.57), имеем

$$\lim_{n \rightarrow \infty} \| (u_n + v_n) - (u'_n + v'_n) \|_D = 0,$$

т.е. $\{u_n + v_n\} \sim \{u'_n + v'_n\}$, а значит, и $\{u'_n + v'_n\} \in \tilde{u} + \tilde{v}$, что доказывает корректность определения в B операции сложения элементов.

Введем операцию умножения элементов из B на числа. Если последовательность $\{u_n\} \in \tilde{u}$ фундаментальна в D , то легко проверить, что последовательность $\{\lambda u_n\}$, где $\lambda \in \mathbb{R}$, также фундаментальна в D . Следовательно, она входит в некоторый класс эквивалентности, который обозначим $\lambda \tilde{u}$. Проверим, что определение этого класса корректно, т.е. оно не зависит от выбора последовательностей из класса $\tilde{u}$. Пусть $\{u_n\} \sim \{u'_n\}$.

Докажем, что $\{\lambda u_n\} \sim \{\lambda u'_n\}$ для любого $\lambda \in \mathbb{R}$. Используя (4.56), находим

$$\lim_{n \rightarrow \infty} \|\lambda u_n - \lambda u'_n\|_D = |\lambda| \lim_{n \rightarrow \infty} \|u_n - u'_n\|_D = 0,$$

т.е. $\{\lambda u_n\} \sim \{\lambda u'_n\}$, а значит, и $\{\lambda u'_n\} \in \tilde{u}$, что доказывает корректность определения в B операции умножения элемента на число.

Итак, в B введены линейные операции. Так как они определены через линейные операции в D , то можно показать, что в B выполнены все аксиомы линейного пространства. Поэтому B является линейным пространством. Роль нулевого элемента в B выполняет класс $\tilde{0}$, определяемый условием $\tilde{u} + \tilde{0} = \tilde{u}$, $\tilde{u} \in B$. Представителем этого класса является фундаментальная последовательность, все члены которой равны нулевому элементу $0 \in D$. Отсюда, учитывая (4.56), получаем, что последовательность $\{u_n\} \in \tilde{0}$ тогда и только тогда, когда $\|u_n\| \rightarrow 0$ при $n \rightarrow \infty$.

Введем норму в линейном пространстве B . Для любого класса эквивалентности $\tilde{u}$ выберем фундаментальную последовательность $\{u_n\} \in \tilde{u}$ и положим

$$\|\tilde{u}\|_B = \lim_{n \rightarrow \infty} \|u_n\|_D. \quad (4.58)$$

Так как в соответствии с (4.1)

$$|\|u_m\|_D - \|u_n\|_D| \leq \|u_m - u_n\|_D,$$

то числовая последовательность $\{\|u_n\|\}$ в силу определения 4.1 фундаментальна и, согласно критерию Коши [1], сходится к некоторому пределу. Значение этого предела не зависит от выбора последовательности из класса $\tilde{u}$. В самом деле, для произвольных последовательностей $\{u_n\}$ и $\{u'_n\}$ из класса эквивалентности $\tilde{u}$ в соответствии с (4.1) имеем $|\|u_n\|_D - \|u'_n\|_D| \leq \|u_n - u'_n\|_D$ и, используя (4.56), получаем $\|u\|_D - \|u'_n\|_D \rightarrow 0$ при $n \rightarrow \infty$, или $\lim_{n \rightarrow \infty} \|u'_n\|_D = \lim_{n \rightarrow \infty} \|u_n\|_D$.

Для (4.58) выполнены аксиомы нормы (см. 4.1). Действительно, $\|\tilde{u}\|_B \geq 0$, причем если $\|\tilde{u}\|_B = 0$, то для любой последовательности $\{u_n\} \in \tilde{u}$ в соответствии с (4.58) $\|u_n\|_D \rightarrow 0$ при $n \rightarrow \infty$. Таким образом, $\{u_n\} \in \tilde{0}$, т.е. $\tilde{u} = \tilde{0}$. Очевидно и обратное: если $\tilde{u} = \tilde{0}$, то $\|\tilde{u}\|_B = 0$. Из аксиом нормы имеем $\|\lambda u_n\|_D = |\lambda| \|u_n\|_D$ и $\|u_n + v_n\|_D \leq \|u_n\|_D + \|v_n\|_D$, где $\lambda \in \mathbb{R}$, $\{u_n\} \in \tilde{u}$ и $\{v_n\} \in \tilde{v}$. Переходя в этих соотношениях к пределу при $n \rightarrow \infty$, получаем $\|\lambda \tilde{u}\|_B = |\lambda| \|\tilde{u}\|_B$ и $\|\tilde{u} + \tilde{v}\|_B \leq \|\tilde{u}\|_B + \|\tilde{v}\|_B$.

Таким образом, (4.58) определяет в линейном пространстве B норму, т.е. B является нормированным пространством. Покажем, что оно включает линейное многообразие D' , изометричное D . Прежде всего убедимся, что существует инъективное отображение φ , переводящее D в B . Любому элементу $u \in D$ поставим в соответствие класс $\varphi(u) = \tilde{u}' \in B$ фундаментальных последовательностей, которому принадлежит стационарная последовательность $\{u_n\}$ с элементами $u_n = u \in D$, $n \in \mathbb{N}$. Ясно, что эта последовательность сходится к элементу $u \in D$. Но если в классе эквивалентности $\tilde{u}' \in B$ хотя бы одна последовательность $\{u_n\} \in \tilde{u}'$ сходится к некоторому элементу $u \in D$, то все последовательности этого класса сходятся, причем к этому же элементу $u \in D$. В самом деле, пусть $\{u'_n\} \in \tilde{u}'$. Тогда с учетом аксиомы 3 нормы (см. 4.1) запишем

$$\|u'_n - u\|_D \leq \|u'_n - u_n\|_D + \|u_n - u\|_D.$$

Переходя в этом неравенстве к пределу при $n \rightarrow \infty$ и учитывая (4.56), получаем, что $\|u'_n - u\|_D \rightarrow 0$ при $n \rightarrow \infty$. Поэтому, если $\varphi(u) = \varphi(w)$ для некоторых $u, w \in D$, $u \neq w$, то класс эквивалентности $\varphi(u)$ состоит из последовательностей, которые сходятся как к u , так и к w , что невозможно. Следовательно, отображение φ инъективно.

Из (4.58) вытекает, что если $u \in D$ и $\tilde{u}' = \varphi(u)$, то

$$\|\tilde{u}'\|_B = \|u\|_D. \quad (4.59)$$

Обозначим область значений отображения φ через D' . Соответствие φ между элементами $u \in D$ и классами $\tilde{u}' \in D' \subset B$,

удовлетворяющее (4.59), является биективным (взаимно однозначным) отображением D в D' . Это отображение линейно в силу определения в $\mathcal{B}$ линейных операций с классами эквивалентности, и поэтому D' — линейное многообразие в $\mathcal{B}$, а отображение φ — изоморфизм между D и D' . В качестве нормы в D' используем норму в $\mathcal{B}$. С учетом (4.59) $\|\varphi(u)\|_{D'} = \|\varphi(u)\|_{\mathcal{B}} = \|u\|_D$. Согласно определению 4.10, это означает, что D и D' изометричны.

Итак, множество D' всех классов $\tilde{u}'$, содержащих стационарные последовательности, члены каждой из которых равны соответствующему элементу из D , является линейным многообразием в $\mathcal{B}$, изометричным D . Покажем, что это многообразие всюду плотно в $\mathcal{B}$. Выберем произвольное $\varepsilon > 0$ и в любом классе $\tilde{u} \in \mathcal{B}$ рассмотрим некоторую фундаментальную последовательность $\{u_m\} \in \tilde{u}$ элементов из D . В силу определения 4.1 фундаментальной последовательности существует такой номер $N \in \mathbb{N}$, что

$$\|u_n - u_m\|_D < \frac{\varepsilon}{2} \quad (4.60)$$

при $m > N$ и $n > N$. Зафиксируем номер $m > N$. Тогда класс $\varphi(u_m) = \tilde{u}_m \in D'$ содержит стационарную последовательность, все элементы которой равны u_m . В соответствии с (4.58) и (4.60) получим

$$\|\tilde{u} - \tilde{u}_m\|_{\mathcal{B}} = \lim_{n \rightarrow \infty} \|u_n - u_m\|_D \leq \frac{\varepsilon}{2} < \varepsilon. \quad (4.61)$$

Это означает, что в любой ε -окрестности точки $\tilde{u} \in \mathcal{B}$ найдется хотя бы одна точка из D' . Таким образом, согласно определению 4.2, линейное многообразие D' всюду плотно в $\mathcal{B}$.

Наконец, покажем, что нормированное пространство $\mathcal{B}$ является полным, т.е. банаховым. Пусть задана фундаментальная в $\mathcal{B}$ последовательность $\{\tilde{u}_n\}$ элементов из $\mathcal{B}$. Так как D' всюду плотно в $\mathcal{B}$, то для каждого $\tilde{u}_n \in \mathcal{B}$ можно найти такой элемент $\tilde{v}_n \in D'$, что

$$\|\tilde{v}_n - \tilde{u}_n\|_{\mathcal{B}} < \frac{1}{n}. \quad (4.62)$$

Тогда, используя (4.1), получаем

$$\begin{aligned}\|\tilde{v}_m - \tilde{v}_n\|_B &= \|(\tilde{v}_m - \tilde{u}_m) + (\tilde{u}_n - \tilde{v}_n) + (\tilde{u}_m - \tilde{u}_n)\|_B \leq \\ &\leq \|\tilde{v}_m - \tilde{u}_m\|_B + \|\tilde{u}_n - \tilde{v}_n\|_B + \|\tilde{u}_m - \tilde{u}_n\|_B < \\ &< \frac{1}{m} + \frac{1}{n} + \|\tilde{u}_m - \tilde{u}_n\|_B. \quad (4.63)\end{aligned}$$

Согласно определению 4.1 фундаментальной последовательности, для произвольного $\varepsilon > 0$ существует такой номер N , что $\|\tilde{u}_m - \tilde{u}_n\|_B < \frac{\varepsilon}{2}$ при $m > N$ и $n > N$. Выберем $m > \max\left\{\frac{4}{\varepsilon}, N\right\}$ и $n > \max\left\{\frac{4}{\varepsilon}, N\right\}$. Тогда из (4.63) имеем

$$\|\tilde{v}_m - \tilde{v}_n\|_B < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{2} = \varepsilon, \quad (4.64)$$

что, согласно определению 4.1, означает фундаментальность последовательности $\{\tilde{v}_n\}$ в B . Благодаря изометричности D' и D каждому элементу $\tilde{v}_n \in D'$ соответствует единственный элемент $v_n \in D$, причем в силу фундаментальности последовательности $\{\tilde{v}_n\}$ последовательность $\{v_n\}$ также фундаментальна и ей соответствует некоторый класс эквивалентности $\tilde{v} \in B$.

Учитывая (4.1) и (4.62), запишем

$$0 \leq \|\tilde{u}_n - \tilde{v}\|_B \leq \|\tilde{u}_n - \tilde{v}_n\|_B + \|\tilde{v} - \tilde{v}_n\|_B < \frac{1}{n} + \|\tilde{v} - \tilde{v}_n\|_B. \quad (4.65)$$

Так как $\|\tilde{v} - \tilde{v}_n\|_B = \lim_{m \rightarrow \infty} \|\tilde{v}_m - \tilde{v}_n\|_B$, то $\lim_{m \rightarrow \infty} \|\tilde{v} - \tilde{v}_n\|_B = 0$. Переходя в (4.65) к пределу при $n \rightarrow \infty$, получаем

$$\lim_{n \rightarrow \infty} \|\tilde{u}_n - \tilde{v}\|_B = 0,$$

т.е. на основании (4.2) заключаем, что $\lim_{n \rightarrow \infty} \tilde{u}_n = \tilde{v}$. Это означает, что произвольная фундаментальная последовательность элементов из B сходится по норме $\|\cdot\|_B$ к элементу, также принадлежащему B , т.е. нормированное пространство B является полным (банаховым) пространством (см. 4.1).

Таким образом, все условия определения 4.11 выполнены, а это значит, что B является пополнением нормированного пространства D . ►

Вопросы и задачи

4.1. Доказать, что в функциональном пространстве $\mathcal{U}$ непрерывно дифференцируемых на отрезке $[a, b]$ функций $f(x)$ можно ввести следующие нормы:

$$\|f\|_1 = |f(a)| + \max_{x \in [a, b]} |f'(x)|, \quad \|f\|_2 = \int_a^b |f(x)| dx + \max_{x \in [a, b]} |f'(x)|.$$

Показать, что нормы $\|\cdot\|_1$ и $\|\cdot\|_2$ эквивалентны норме

$$\|f\|_0 = \max \left\{ \max_{x \in [a, b]} |f(x)|, \max_{x \in [a, b]} |f'(x)| \right\}, \quad f \in \mathcal{U},$$

т.е. найдутся числа $\alpha_i, \beta_i > 0$, $i = 1, 2$, такие, что

$$\alpha_i \|f\|_0 \leq \|f\|_i \leq \beta_i \|f\|_0, \quad f \in \mathcal{U}.$$

4.2. Доказать неравенство Гельдера* для интегралов при $p > 1$, $1/p + 1/q = 1$:

$$\int_a^b |f(t)g(t)| dt \leq \left(\int_a^b |f(t)|^p dt \right)^{1/p} \left(\int_a^b |g(t)|^q dt \right)^{1/q}.$$

4.3. Доказать неравенство Минковского для интегралов при $p \geq 1$:

$$\left(\int_a^b |f(t) + g(t)|^p dt \right)^{1/p} \leq \left(\int_a^b |f(t)|^p dt \right)^{1/p} + \left(\int_a^b |g(t)|^p dt \right)^{1/p}$$

*О Гельдер (1859–1937) — немецкий математик.

4.4. Показать, что множество $L_p[a, b]$, $p \geq 1$, функций f , определенных на отрезке $[a, b]$, для которых существует конечный интеграл Лебега $\int_a^b |f(t)|^p dt$, является нормированным пространством с нормой

$$\|f\|_p = \left(\int_a^b |f(t)|^p dt \right)^{1/p}.$$

4.5. Доказать, что нормированное пространство $C^k[a, b]$ функций f , имеющих на отрезке $[a, b]$ непрерывные производные до k -го порядка включительно, с нормой

$$\|f\| = \sum_{m=0}^k \max_{x \in [a, b]} |f^{(m)}(x)|, \quad f \in C^k[a, b],$$

является банаховым.

4.6. Доказать, что оператор $A: C^1[0, 1] \rightarrow C[0, 1]$, действующий по правилу $\varphi = Af \iff \varphi(x) = \frac{df}{dx}$, $x \in [0, 1]$, является непрерывным.

4.7. Является ли оператор $A: \mathcal{U} \rightarrow C[a, b]$, где $\mathcal{U}$ — нормированное пространство непрерывно дифференцируемых на отрезке $[a, b]$ функций с нормой

$$\|f\| = \max_{x \in [a, b]} |f(x)|, \quad f \in \mathcal{U},$$

действующий по правилу $\varphi = Af \iff \varphi(x) = \frac{df}{dx}$, $x \in [a, b]$, непрерывным?

4.8. Пусть $Q[0, 1]$ — нормированное пространство непрерывных на отрезке $[0, 1]$ функций f (с обычными операциями сложения и умножения на число) с нормой

$$\|f\| = \left(\int_0^1 |f(x)|^2 dx \right)^{1/2}, \quad f \in Q[0, 1].$$

Доказать, что оператор $A: Q[0, 1] \rightarrow Q[0, 1]$, действующий по правилу $\varphi = Af \iff \varphi(x) = f^2(x)$, $x \in [0, 1]$, не является непрерывным, а оператор $A: C[0, 1] \rightarrow Q[0, 1]$, действующий по тому же правилу, непрерывен.

4.9. Доказать, что оператор $A: C[0, 1] \rightarrow C[0, 1]$, действующий по правилу

$$\varphi = Af \iff \varphi(t) = \int_0^1 e^{t^2+s} f(s) ds + tf(0) + t^2 f(1/2) + t^3 f(1)$$

при $t \in [0, 1]$, является вполне непрерывным.

4.10. Является ли оператор $A: C[0, 1] \rightarrow C[0, 1]$, действующий по правилу $\varphi = Af \iff \varphi(t) = t^3 f(t)$, $t \in [0, 1]$, вполне непрерывным?

4.11. Выяснить, будет ли вполне непрерывным оператор $A: C^1[0, 1] \rightarrow C[0, 1]$, определяемый равенством $(Af)(t) = \frac{df(t)}{dt}$, $t \in [0, 1]$.

4.12. Доказать, что оператор $A: L_p[a, b] \rightarrow L_p[a, b]$, $p \geq 1$ (см. задачу 4.4), действующий по правилу

$$\varphi = Af \iff \varphi(t) = \int_a^t f(t) dt, \quad f \in L_p[a, b],$$

является вполне непрерывным.

4.13. Доказать, что для линейного ограниченного функционала $F: \mathcal{U} \rightarrow \mathbb{R}$, определенного в нормированном пространстве $\mathcal{U}$, норма $\|F\| = \inf \mathcal{M} = \inf \mathcal{K}$, где $\mathcal{M}$ и $\mathcal{K}$ — множества всех чисел M и K , таких, что $|Fu| \leq M\|u\|_{\mathcal{U}}$ для всех $u \in \mathcal{U}$ и $|Fu| \leq K$ для всех $u \in \mathcal{U}$, удовлетворяющих условию $\|u\|_{\mathcal{U}} \leq 1$.

4.14. Найти норму линейного функционала F , определенного в нормированном пространстве $\mathcal{U}$:

а) $F(x) = 2x(1) - x(0)$, $x = x(t)$, $x \in \mathcal{U} = C[-1, 1]$;

б) $F(x) = \int_0^1 x(t)(2t - 1) dt$, $x = x(t)$, $x \in \mathcal{U} = C[0, 1]$;

в) $F(x) = \int_0^1 tx(t) dt - x(0)$, $x = x(t)$, $x \in \mathcal{U} = C[0, 1]$;

г) $F(x) = \int_{-1}^1 t^2 x(t) dt$, $x = x(t)$, $x \in \mathcal{U} = L_2[-1, 1]$;

д) $F(x) = \int_{-2}^2 t^4 x(t) dt$, $x = x(t)$, $x \in \mathcal{U} = L_1[-2, 2]$.

4.15. Доказать, что оператор $A: \mathcal{U} \rightarrow \mathcal{V}$ ($\mathcal{U}, \mathcal{V}$ — нормированные пространства) является линейным ограниченным, и найти его норму, если:

а) $(Au)(t) = \int_0^t u(s) ds$, $\mathcal{U} = C[0, 1]$, $\mathcal{V} = C[0, 1]$;

б) $(Au)(t) = u(t)$, $\mathcal{U} = C[-1, 1]$, $\mathcal{V} = C[0, 1]$;

в) $(Au)(t) = t^2 u(0)$, $\mathcal{U} = C[0, 1]$, $\mathcal{V} = C[0, 1]$;

г) $(Au)(t) = u(t^2)$, $\mathcal{U} = C[0, 1]$, $\mathcal{V} = C[0, 1]$.

4.16. Доказать, что оператор $A: C[0, 1] \rightarrow C[0, 1]$, определяемый равенством

$$(Au)(t) = u(t) + \int_0^1 ts^2 u(s) ds, \quad t \in [0, 1],$$

является линейным ограниченным. Доказать, что существует оператор A^{-1} , обратный к оператору A , причем этот оператор также является линейным ограниченным. Найти A^{-1} .

4.17. Для каждого значения параметра $\alpha \in \mathbb{R}$ найти спектр оператора $A: C[0, 1] \rightarrow C[0, 1]$, определяемого для любой функции $f \in C[0, 1]$ равенством $(Af)(x) = xf(x) + \alpha f(0)$, $x \in [0, 1]$. Какие точки спектра принадлежат точечному, остаточному, непрерывному спектрам? Имеет ли этот оператор собственные функции?

4.18. Доказать, что операторы, определяемые соотношениями (4.31) и (4.33), являются линейными и ограниченными.

4.19. Доказать, что следующие уравнения имеют единственное решение в банаховом пространстве $C[0, 1]$:

$$\text{а) } f(x) + \frac{1}{2} \int_0^1 \frac{f(y) \operatorname{arctg}(e^{x+y^3} - \ln(1 + \sqrt{xy}))}{x^7 + y^2 + 1} dy = 0;$$

$$\text{б) } \int_0^1 \frac{f(y) dy}{f^2(y) + x^4 + y^2 + 9} - f(x) = e^x.$$

4.20. Доказать, что оператор $A: L_p[a, b] \rightarrow L_p[a, b]$, действующий по правилу

$$A\varphi = f \iff f(t) = \int_a^b K(t, s) \varphi(s) ds, \quad t \in [a, b],$$

где $K(t, s)$ — функция, непрерывная в квадрате $[a, b]^2$, является вполне непрерывным. Найти спектр этого оператора, если:

$$\text{а) } K(t, s) = ts^4, \quad a = 0, \quad b = 1;$$

$$\text{б) } K(t, s) = t - s, \quad a = 0, \quad b = 1;$$

$$\text{в) } K(t, s) = ts(1 - ts), \quad a = 0, \quad b = 1;$$

$$\text{г) } K(t, s) = \sum_{k=1}^{\infty} \frac{\sin kt \sin ks}{k^2}, \quad a = 0, \quad b = \pi.$$

5. ОПЕРАТОРЫ В ГИЛЬБЕРТОВЫХ ПРОСТРАНСТВАХ

5.1. Гильбертово пространство

Бесконечномерное *банахово пространство*, в котором введена операция *скалярного умножения*, индуцирующая норму этого пространства, называют *гильбертовым**. Мы будем обозначать его, как правило, символом $\mathcal{H}$ (по первой букве фамилии Д. Гильберта). Скалярное произведение элементов u и v будем обозначать $\langle u, v \rangle$.

Напомним, что операция скалярного умножения удовлетворяет следующим *аксиомам скалярного умножения* (в их формулировках u, v, w — произвольные элементы линейного пространства, а α — произвольное действительное число):

- 1) $\langle u, v \rangle = \langle v, u \rangle$ (симметрия);
- 2) $\langle u + w, v \rangle = \langle u, v \rangle + \langle w, v \rangle$ (дистрибутивность);
- 3) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$ (однородность);
- 4) $\langle u, u \rangle \geq 0$, причем $\langle u, u \rangle = 0$ тогда и только тогда, когда $u = 0$ (неотрицательность скалярного квадрата).

В случае комплексного гильбертова пространства аксиома симметрии принимает вид $\langle u, v \rangle = \overline{\langle v, u \rangle}$.

Норма в гильбертовом пространстве, порожденная скалярным умножением, выражается через *скалярный квадрат* $\langle u, u \rangle$:

$$\|u\| = \sqrt{\langle u, u \rangle}. \quad (5.1)$$

Неравенство Коши — Буняковского

$$|\langle u, v \rangle| \leq \sqrt{\langle u, u \rangle} \sqrt{\langle v, v \rangle} \quad (5.2)$$

* Д. Гильберт (1862–1943) — немецкий математик.

справедливо для произвольного евклидова пространства, в том числе и для гильбертова пространства (в последнем случае его иногда называют **неравенством Шварца***). Отметим, что неравенство Коши — Буняковского превращается в равенство тогда и только тогда, когда элементы u и v линейно зависимы, т.е. один из них может быть получен умножением другого на число. В частности, это верно, когда хотя бы один из элементов u, v является нулевым.

С учетом (5.1) неравенство (5.2) принимает вид

$$|\langle u, v \rangle| \leq \|u\| \|v\|. \quad (5.3)$$

Пример 5.1. В линейном пространстве X функций, непрерывных на отрезке $[0, 1]$, скалярное умножение можно ввести соотношением $[IX]$

$$\langle f, g \rangle = \int_0^1 f(x) g(x) dx, \quad f, g \in X. \quad (5.4)$$

Несложно проверить, что для формулы (5.4) выполнены все аксиомы скалярного умножения. Это пространство, являясь линейным (см. пример 2.1), будет нормированным. При этом норма, индуцированная скалярным умножением, определена соотношением

$$\|f\| = \left(\int_0^1 f^2(x) dx \right)^{1/2}, \quad f \in X.$$

Последовательность функций из X , фундаментальная в нормированном пространстве X , может не иметь предела в X , т.е. нормированное пространство X со скалярным умножением (5.4) не является полным (а значит, и гильбертовым). Например, нетрудно показать, что последовательность $\{\varphi_n\}$ не-

*Г Шварц (1843–1921) — немецкий математик.

прерывных на отрезке $[0, 1]$ функций $\varphi_n(x)$, рассмотренных в примере 4.3 (см. рис. 4.1), фундаментальна в X , но не является сходящейся в X .

Известно*, что гильбертовым является пространство функций $f(x)$, суммируемых на отрезке $[0, 1]$ с квадратом. Это пространство обозначают $L_2[0, 1]$. Скалярное умножение в нем определяют соотношением

$$\langle f, g \rangle = \int_0^1 f(x)g(x)dx, \quad f, g \in L_2[0, 1], \quad (5.5)$$

а норму — соотношением

$$\|f\| = \left(\int_0^1 f^2(x)dx \right)^{1/2}, \quad f \in L_2[0, 1]. \quad (5.6)$$

Равенство (5.5) аналогично (5.4), но теперь интегралы в (5.6) и (5.5) следует понимать как интегралы Лебега. Нетрудно убедиться, используя свойства интеграла Лебега [IX] (они аналогичны свойствам определенного интеграла), что для (5.5) выполнены все аксиомы скалярного умножения.

Сходимость в среднем квадратичном фундаментальной последовательности $\{g_n(x)\}$ функций $g_n(x) \in L_2[0, 1]$ к функции $g(x) \in L_2[0, 1]$ означает, что

$$\lim_{n \rightarrow \infty} \int_0^1 (g_n(x) - g(x))^2 dx = 0.$$

Отметим, что непрерывные, непрерывно дифференцируемые заданное число раз, кусочно постоянные и измеримые ограниченные функции (в том числе принимающие на $[0, 1]$ конечное число значений) составляют всюду плотные подмножества гильбертова пространства $L_2[0, 1]$ [IX]. Поскольку в $L_2[0, 1]$ су-

*См., например: Колмогоров А. Н., Фомин С. В.

ществуют счетные всюду плотные подмножества (например, многочлены с рациональными коэффициентами), то пространство $L_2[0, 1]$ будет сепарабельным. #

Гильбертовым пространством является множество $L_2(\Omega)$ действительных функций, суммируемых с квадратом на измеримом множестве $\Omega \subset \mathbb{R}^n$ [XV]. Скалярное умножение в $L_2(\Omega)$ вводят соотношением

$$\langle f, g \rangle = \int_{\Omega} f(x) g(x) dx, \quad f, g \in L_2(\Omega).$$

Для этих функций конечен интеграл Лебега

$$\|f\|^2 = \int_{\Omega} |f(x)|^2 dx,$$

который определяет норму в $L_2(\Omega)$.

Обобщением пространства $L_2(\Omega)$ является гильбертово пространство $L_2(\Omega, \sigma)$ функций, суммируемых с квадратом и с весом σ , где $\sigma(x)$ — неотрицательная измеримая на Ω действительная функция. Для функций из $L_2(\Omega, \sigma)$ конечен интеграл

$$\|f\|^2 = \int_{\Omega} |f(x)|^2 \sigma(x) dx.$$

Скалярное произведение в $L_2(\Omega)$ имеет вид [XV]

$$\langle f, g \rangle = \int_{\Omega} f(x) g(x) \sigma(x) dx, \quad f, g \in L_2(\Omega, \sigma).$$

Гильбертовым будет и линейное пространство $L_2^{(m)}(\Omega)$ векторных функций $f: \Omega \rightarrow \mathbb{R}^m$, определенных на измеримом мно-

жестве $\Omega \subset \mathbb{R}^n$, для которых конечен интеграл

$$\|f\|^2 = \int_{\Omega} (f(x), f(x)) dx,$$

где $(\cdot, \cdot)$ обозначает операцию *стандартного скалярного умножения* векторов m -мерного евклидова арифметического пространства. Правило скалярного умножения в $L_2^{(m)}(\Omega)$ задает формула [XV]

$$\langle f, g \rangle = \int_{\Omega} (f(x), g(x)) dx, \quad f, g \in L_2^{(m)}(\Omega).$$

Отметим, что в $L_2[0, 1]$, $L_2(\Omega)$ и $L_2^{(m)}(\Omega)$ считают равными любые две функции, отличающиеся на множестве, для которого *мера Лебега* равна нулю (говорят также „на множестве меры нуль“). Это обеспечивает выполнение аксиомы скалярного умножения (и соответственно нормы), согласно которой $\langle u, u \rangle = 0$ только в случае, если u является нулевым элементом линейного пространства.

Напомним, что если *элементы* u и v гильбертова пространства $\mathcal{H}$ связаны соотношением $\langle u, v \rangle = 0$, то их называют *ортogonalными*. В этом случае используют запись $u \perp v$ (или $v \perp u$). Если элемент $w \in \mathcal{H}$ ортogonalен каждому элементу u *подпространства* $M \subset \mathcal{H}$, т.е. $w \perp u$, $u \in M$, то этот элемент называют *ортogonalным подпространству* M и пишут $w \perp M$.

Если последовательности $\{u_n\}$ и $\{v_n\}$ элементов гильбертова пространства $\mathcal{H}$ по норме сходятся к элементам u и v , то существует предел [IX]

$$\lim_{n \rightarrow \infty} \langle u_n, v_n \rangle = \langle u, v \rangle, \quad u, v \in \mathcal{H}. \quad (5.7)$$

Полагая в (5.7) сначала $v_n = v$, $n \in \mathbb{N}$, а затем $v = u$ и $v_n = u_n$, $n \in \mathbb{N}$, получаем

$$\lim_{n \rightarrow \infty} \langle u_n, v \rangle = \langle u, v \rangle \quad \text{и} \quad \lim_{n \rightarrow \infty} \|u_n\| = \|u\|, \quad (5.8)$$

в частности при $v = u$ из первого равенства (5.8) имеем

$$\lim_{n \rightarrow \infty} \langle u_n, u \rangle = \langle u, u \rangle = \|u\|^2. \quad (5.9)$$

Замечание 5.1. Элемент $u \in \mathcal{H}$, ортогональный множеству M , всюду плотному в гильбертовом пространстве $\mathcal{H}$, является нулевым элементом $\mathcal{H}$, т.е. $u = 0 \in \mathcal{H}$. Действительно, если $u \perp M$ и $\overline{M} = \mathcal{H}$, то существует последовательность $\{u_n\} \subset M$, такая, что $u_n \rightarrow u$ при $n \rightarrow \infty$, и для любого $n \in \mathbb{N}$ имеем $\langle u_n, u \rangle = 0$. Тогда

$$0 = \lim_{n \rightarrow \infty} \langle u_n, u \rangle = \left\langle \lim_{n \rightarrow \infty} u_n, u \right\rangle = \langle u, u \rangle = \|u\|^2,$$

т.е. $u = 0$. Таким образом, если для некоторого элемента $u \in \mathcal{H}$ при любом $v \in M$ справедливо $\langle u, v \rangle = 0$, то $u = 0 \in \mathcal{H}$.

Теорема 5.1. Пусть последовательность $\{v_k\}$ в гильбертовом пространстве $\mathcal{H}$ образует счетный базис. Если для некоторого элемента $u \in \mathcal{H}$ выполнены равенства $\langle u, v_k \rangle = 0$, $k \in \mathbb{N}$, то $u = 0$.

◀ Множество M всех элементов вида

$$v = \sum_{k=1}^n b_k v_k \quad (5.10)$$

будет всюду плотным в $\mathcal{H}$, так как эти элементы образуют линейную оболочку счетного базиса [IX]. По условию теоремы $\langle u, v_k \rangle = 0$ при произвольном $k \in \mathbb{N}$. Поэтому с учетом линейности скалярного умножения, согласно (5.10),

$$\langle u, v \rangle = \left\langle u, \sum_{k=1}^n b_k v_k \right\rangle = \sum_{k=1}^n b_k \langle u, v_k \rangle = 0, \quad n \in \mathbb{N}, \quad v \in M.$$

Отсюда в соответствии с замечанием 5.1 следует, что $u = 0$. ▶

В любом сепарабельном гильбертовом пространстве $\mathcal{H}$ существует ортонормированный базис [IX]. Напомним схему построения ортонормированного базиса в $\mathcal{H}$. Если $\mathcal{H}$ сепарабельно, то можно выделить счетное множество $M \subset \mathcal{H}$, такое, что $\overline{M} = \mathcal{H}$. Располагая элементы из M в виде последовательности и удаляя из нее все элементы, являющиеся линейными комбинациями предыдущих, получаем линейно независимую систему $\{u_n\} \subset M$, замкнутую в $\mathcal{H}$.

Используя процесс ортогонализации Грама — Шмидта, эту систему можно ортонормировать и получить ортонормированную систему $\{w_k\}$ функций $w_k \in M$, $k \in \mathbb{N}$, причем также замкнутую в $\mathcal{H}$. Ортонормированная замкнутая система и является ортонормированным базисом в $\mathcal{H}$. Отметим также, что эта система будет *полной ортонормированной системой*. Для такой системы справедливо, что в $\mathcal{H}$ нет элемента, кроме нулевого, ортогонального всем элементам этой системы.

5.2. Операторы и функционалы в гильбертовом пространстве

Рассмотрим линейный оператор $A: D(A) \rightarrow \mathcal{H}$, область определения $D(A)$ которого является всюду плотным множеством в гильбертовом пространстве $\mathcal{H}$.

Определение 5.1. Линейный оператор A называют **симметрическим**, если для произвольных элементов $u, v \in D(A)$ справедливо равенство $\langle Au, v \rangle = \langle u, Av \rangle$, и **положительным**, если к тому же $\langle Au, u \rangle \geq 0$ для любого элемента $u \in D(A)$, причем $\langle Au, u \rangle = 0$ тогда и только тогда, когда $u = 0$. При выполнении неравенства

$$\langle Au, u \rangle \geq \gamma^2 \|u\|^2, \quad u \in D(A), \quad (5.11)$$

где $\gamma \neq 0$, положительный оператор A называют **положительно определенным**.

Таким образом, любой положительно определенный оператор является положительным, а значит, и симметрическим, но обратное, вообще говоря, неверно. Отметим, что для положительно определенного оператора A ядро $\ker A = \{0\}$.

Пример 5.2. Рассмотрим оператор $A = -\frac{d^2}{dx^2}$ на множестве $D(A)$ дважды непрерывно дифференцируемых на отрезке $[0, 1]$ функций $w(x)$, удовлетворяющих однородным граничным условиям

$$w(0) = w(1) = 0. \quad (5.12)$$

Множество $D(A)$ является *линейным многообразием* в гильбертовом пространстве $L_2[0, 1]$ функций, суммируемых с квадратом и определенных на отрезке $[0, 1]$. В самом деле, для любых функций $f(x), g(x) \in D(A)$ и произвольных чисел $\alpha, \beta \in \mathbb{R}$ имеем $(\alpha f(x) + \beta g(x)) \in D(A)$, поскольку любая линейная комбинация дважды непрерывно дифференцируемых функций дважды непрерывно дифференцируема, а

$$\alpha f(0) + \beta g(0) = \alpha f(1) + \beta g(1) = 0,$$

если $f(0) = g(0) = f(1) = g(1) = 0$. Множество $D(A)$ является *всюду плотным* в $L_2[0, 1]$ (см. пример 5.1).

Отметим, что если в (5.12) хотя бы одно из граничных условий неоднородное, то множество $D(A)$ уже не будет линейным многообразием. Пусть, например, вместо (5.12) заданы граничные условия $w(0) = 0$ и $w(1) = 1$. Тогда для $f(x), g(x) \in D(A)$ имеем $f(1) = g(1) = 1$. Можно подобрать такие числа $\alpha, \beta \in \mathbb{R}$, что $(\alpha f(1) + \beta g(1)) = \alpha + \beta \neq 1$, т.е. $(\alpha f(x) + \beta g(x)) \notin D(A)$, а это означает, что множество $D(A)$ не является линейным многообразием.

Линейность оператора A следует из свойств линейности множества $D(A)$ и линейности операции дифференцирования. Убедимся, что оператор A симметрический. Ясно, что для функций из $D(A)$ интеграл в (5.5) достаточно рассматривать в смысле Римана. Используя (5.5) и интегрируя по частям с

учетом условия вида (5.12), для функций $u, v \in D(A)$ получаем

$$\begin{aligned}\langle Au, v \rangle &= - \int_0^1 u'' v \, dx = -u'v \Big|_0^1 + \int_0^1 u'v' \, dx = \int_0^1 u'v' \, dx = \\ &= uv' \Big|_0^1 - \int_0^1 uv'' \, dx = \int_0^1 u(-v'') \, dx = \langle u, Av \rangle, \quad (5.13)\end{aligned}$$

т.е. оператор A — симметрический в $D(A)$.

Из (5.13) следует, что

$$\langle Aw, w \rangle = \int_0^1 (w')^2 \, dx \geq 0, \quad w \in D(A),$$

причем если $\langle Aw, w \rangle = 0$, то $w'(x) = 0$, $x \in [0, 1]$, а значит, w — постоянная на отрезке $[0, 1]$ функция, и в соответствии с заданными граничными условиями имеем $w(x) = 0$, $x \in [0, 1]$. Таким образом, оператор A является положительным в $D(A)$. В примере 5.10 показано, что этот оператор является и положительно определенным в $D(A)$, т.е. для него справедливо неравенство (5.11), которое в данном случае имеет вид

$$\langle Aw, w \rangle = \int_0^1 (w')^2 \, dx \geq \gamma^2 \int_0^1 w^2 \, dx, \quad w \in D(A). \quad (5.14)$$

Если расширить область определения оператора A , сняв граничные условия (5.12), то он утратит свойство симметричности. В самом деле, рассмотрим функции $u(x) = x$ и $v(x) = x(1-x)$, которые дважды непрерывно дифференцируемы на отрезке $[0, 1]$, но функция $u(x)$ не удовлетворяет условию (5.12). Тогда вместо (5.13) получим

$$\langle Au, v \rangle = - \int_0^1 u'' v \, dx = 0, \quad \langle u, Av \rangle = - \int_0^1 uv'' \, dx = \int_0^1 2x \, dx = 1 \neq 0,$$

т.е. $\langle Au, v \rangle \neq \langle u, Av \rangle$. #

Выясним, каков общий вид *линейного ограниченного функционала* в гильбертовом пространстве $\mathcal{H}$. Для такого функционала справедлива следующая теорема.

Теорема 5.2 (теорема Ф. Рисса*). Для любого линейного ограниченного функционала J в гильбертовом пространстве $\mathcal{H}$ существует единственный элемент $v \in \mathcal{H}$ с нормой $\|v\| = \|J\|$, такой, что

$$Ju = \langle u, v \rangle \in \mathbb{R}, \quad u \in \mathcal{H}. \quad (5.15)$$

◀ Рассмотрим ядро $\ker J$ линейного ограниченного функционала J , являющегося частным случаем линейного оператора. Согласно теореме 4.13, $\ker J$ является *подпространством* в $\mathcal{H}$. Если $\ker J = \mathcal{H}$, то $J = 0$ и равенство (5.15) справедливо при $v = 0$. В случае $\ker J \neq \mathcal{H}$ существует отличный от нулевого элемент $v_0 \perp \ker J$ [IX]. Тогда для любого $u \in \mathcal{H}$ выберем элемент

$$w = (Ju)v_0 - (Jv_0)u,$$

для которого

$$Jw = (Ju)(Jv_0) - (Jv_0)(Ju) = 0,$$

т.е. $w \in \ker J$, и поэтому $w \perp v_0$, или

$$\langle w, v_0 \rangle = 0 = (Ju) \langle v_0, v_0 \rangle - (Jv_0) \langle u, v_0 \rangle.$$

Отсюда

$$Ju = \frac{Jv_0}{\langle v_0, v_0 \rangle} \langle u, v_0 \rangle.$$

При $v = \frac{(Jv_0)v_0}{\langle v_0, v_0 \rangle}$, согласно аксиоме однородности скалярного умножения (см. 5.1), приходим к (5.15).

Если бы существовало два элемента v и v_1 , удовлетворяющих (5.15), т.е. $Ju = \langle u, v \rangle = \langle u, v_1 \rangle$, то в силу дистрибутивно-

*Ф. Рисс (1880–1956) — венгерский математик.

сти скалярного умножения для всех $u \in \mathcal{H}$ было бы $\langle u, v - v_1 \rangle = 0$ и $v_1 = v$ (см. замечание 5.1).

Из (5.15) с учетом неравенства Коши — Буняковского (5.3) получаем $|Ju| = |\langle u, v \rangle| \leq \|u\| \|v\|$, причем неравенство переходит в равенство при $u = \lambda v$. Отсюда, согласно (4.20), находим

$$\|J\| = \sup_{\|u\| \leq 1} |Ju| = \|v\|. \quad \blacktriangleright$$

Отображение $F: \mathcal{L} \times \mathcal{L} \rightarrow \mathbb{R}$, где $\mathcal{L}$ — линейное пространство, называют функционалом, зависящим от двух элементов линейного пространства L . **Функционал** $J[u, v]$, зависящий от двух элементов $u, v \in L$ линейного пространства L , называют **билинейным**, если при любом фиксированном v он является линейным по u , а при любом фиксированном u — линейным по v , т.е. для произвольных действительных чисел α_1 и α_2

$$J[\alpha_1 u_1 + \alpha_2 u_2, v] = \alpha_1 J[u_1, v] + \alpha_2 J[u_2, v],$$

$$J[u, \alpha_1 v_1 + \alpha_2 v_2] = \alpha_1 J[u, v_1] + \alpha_2 J[u, v_2].$$

Полагая в билинейном функционале $v = u$, получаем **квадратичный функционал** $J[u] = J[u, u]$. Этот функционал называют положительно определенным, если для любого ненулевого элемента $u \in L$ выполнено неравенство $J[u] > 0$. Частным случаем квадратичного функционала является функционал вида $J[u] = \langle Au, u \rangle$, $u \in D(A) \subset \mathcal{H}$, который, согласно (5.11), будет положительно определенным, если линейный оператор A — положительный. Квадратичный функционал $J[u]$ сильно положителен, если существует такое постоянное число $k > 0$, что $J[u] \geq k \|u\|^2$ для любого элемента $u \in D(J)$.

В гильбертовом пространстве $\mathcal{H}$ рассмотрим *операторное уравнение*

$$Au = f, \quad (5.16)$$

где A — положительный оператор. Такое уравнение характерно для линейных задач математической физики.

Теорема 5.3. Если оператор A положительный, то уравнение $Au = f$ имеет не более одного решения в $D(A)$.

◀ Допустим, что операторное уравнение имеет два решения u_1 и u_2 , т.е. выполнены равенства

$$Au_1 = f \quad \text{и} \quad Au_2 = f.$$

Вычитая второе равенство из первого и учитывая, что A является положительным (и поэтому линейным) оператором, получаем $A(u_1 - u_2) = 0$, что равносильно равенству

$$\langle A(u_1 - u_2), (u_1 - u_2) \rangle = 0.$$

Отсюда, согласно определению 5.1 положительного оператора, $u_1 - u_2 = 0$, или $u_1 = u_2$. ▶

Функционал вида

$$J[u] = \langle Au, u \rangle - 2\langle f, u \rangle \quad (5.17)$$

с квадратичным слагаемым $\langle Au, u \rangle$ и линейным по u вторым слагаемым обычно также называют квадратичным. Связь между решением операторного уравнения (5.16) и точкой минимума функционала (5.17) устанавливает следующая **теорема о квадратичном функционале**.

Теорема 5.4. Пусть оператор A , определенный на всюду плотном множестве $D(A)$ в гильбертовом пространстве $\mathcal{H}$, является положительным. Элемент $u_0 \in D(A)$ является решением уравнения $Au = f$ тогда и только тогда, когда квадратичный функционал $J[u] = \langle Au, u \rangle - 2\langle f, u \rangle$ принимает на u_0 минимальное значение в $D(A)$, т.е. $J[u] > J[u_0]$ для любого элемента $u \in D(A) \setminus \{u_0\}$.

◀ Функционал $J[u]$ определен при всех $u \in D(A)$. Пусть элемент $u_0 \in D(A)$ удовлетворяет (5.16), так что $f = Au_0$. Подставляя это выражение для f в (5.17), с учетом симметрии

и дистрибутивности скалярного умножения (см. 5.1) и симметричности положительного оператора A (см. определение 5.1) получаем

$$\begin{aligned} J[u] &= \langle Au, u \rangle - 2\langle Au_0, u \rangle = \langle Au, u \rangle - \langle Au_0, u \rangle - \langle u, Au_0 \rangle = \\ &= \langle Au, u \rangle - \langle Au_0, u \rangle - \langle Au, u_0 \rangle + \langle Au_0, u_0 \rangle - \langle Au_0, u_0 \rangle = \\ &= \langle A(u - u_0), u - u_0 \rangle - \langle Au_0, u_0 \rangle. \end{aligned}$$

Следовательно, $J[u_0] = -\langle Au_0, u_0 \rangle$, а

$$J[u] - J[u_0] = \langle A(u - u_0), u - u_0 \rangle \geq 0$$

в силу положительности оператора A , причем $Ju = Ju_0$ лишь при $u = u_0$.

Пусть теперь функционал (5.17) принимает минимальное значение в $D(A)$ на элементе u_0 , т.е.

$$J[u_0 + tv] \geq J[u_0], \quad v \in D(A), \quad t \in \mathbb{R}.$$

Используя аксиомы скалярного умножения и симметричность положительного оператора A , получаем

$$\begin{aligned} J[u_0 + tv] &= \langle A(u_0 + tv), u_0 + tv \rangle - 2\langle f, u_0 + tv \rangle = \\ &= \langle Au_0 + tAv, u_0 + tv \rangle - 2\langle f, u_0 \rangle - 2t\langle f, v \rangle = \\ &= \langle Au_0, u_0 \rangle + 2t\langle Au_0, v \rangle + t^2\langle Av, v \rangle - 2t\langle f, v \rangle - 2\langle f, u_0 \rangle. \end{aligned}$$

При $v \neq 0$ правая часть этого равенства является квадратным трехчленом по t и при фиксированных u_0 , f и v достигает минимума в точке $t = 0$. Следовательно,

$$\left. \frac{d}{dt} J[u_0 + tv] \right|_{t=0} = 2\langle Au_0, v \rangle - 2\langle f, v \rangle = 0,$$

или $\langle Au_0 - f, v \rangle = 0$. Поскольку элемент $v \in D(A)$ выбран произвольно и $D(A)$ всюду плотно в $\mathcal{H}$, то, согласно замечанию 5.1,

$Au_0 - f = 0$, т.е. элемент $u_0 \in D(A)$ является решением уравнения $Au = f$. ►

Пример 5.3. Обыкновенное дифференциальное уравнение

$$\frac{d^2}{dx^2} \left(EI(x) \frac{d^2 w(x)}{dx^2} \right) = q(x), \quad x \in (0, l), \quad (5.18)$$

с граничными условиями

$$w(0) = w(l) = 0, \quad w'(0) = w'(l) = 0 \quad (5.19)$$

описывает изменение вдоль жестко закрепленной на концах балки ее прогиба $w(x)$, вызванного распределенной поперечной нагрузкой $q(x)$ (рис. 5.1), непрерывной в интервале $(0, l)$. Балка имеет длину l , модуль $E = \text{const} > 0$ упругости материала, а момент $I(x)$ инерции ее поперечного сечения является дважды непрерывно дифференцируемой на отрезке $[0, l]$ функцией координаты x , причем $I(x) > 0$, $x \in [0, l]$.

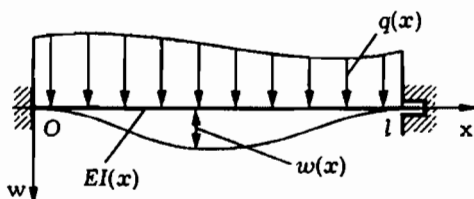


Рис. 5.1

Задачу (5.18), (5.19) можно представить операторным уравнением $Aw = q$, где оператор

$$A = \frac{d^2}{dx^2} \left(EI(x) \frac{d^2}{dx^2} \right) \quad (5.20)$$

определен на множестве $D(A) \subset L_2[0, l]$ функций, непрерывных на $[0, l]$ вместе со своими производными до четвертого порядка включительно и удовлетворяющих (5.19). Множество $D(A)$

является линейным многообразием, всюду плотным в гильбертовом пространстве $L_2[0, l]$ (см. пример 5.1).

Ясно, что оператор A является линейным в силу линейности операции дифференцирования. Используя (5.4) и интегрируя по частям с учетом (5.19), для любых $f, g \in D(A)$ имеем

$$\begin{aligned}\langle Af, g \rangle &= \int_0^l (EI f'')'' g \, dx = (EI f'')' g \Big|_0^l - \int_0^l (EI f'')' g' \, dx = \\ &= -EI f'' g' \Big|_0^l + \int_0^l EI f'' g'' \, dx = \int_0^l EI f'' g'' \, dx. \quad (5.21)\end{aligned}$$

Аналогично получаем

$$\langle f, Ag \rangle = \int_0^l f (EI g'')'' \, dx = \int_0^l EI f'' g'' \, dx,$$

т.е. линейный оператор A , согласно определению 5.1, является симметрическим в $D(A)$.

Из (5.21) следует, что

$$\langle Aw, w \rangle = \int_0^l EI (w'')^2 \, dx \geq 0, \quad w \in D(A), \quad (5.22)$$

причем если $\langle Aw, w \rangle = 0$, то $w''(x) \equiv 0$ на $[0, l]$. Отсюда $w(x) = Ax + B$, но, согласно (5.19), $A = B = 0$, а значит, $w(x) \equiv 0$, т.е. в соответствии с определением 5.1 симметрический оператор A является в $D(A)$ положительным.

Квадратичный функционал (5.17) в данном случае с учетом (5.22) имеет вид

$$J[w] = \langle Aw, w \rangle - 2 \langle q, w \rangle = \int_0^l EI (w'')^2 \, dx - 2 \int_0^l q w \, dx \quad (5.23)$$

и при заданной функции $w \in D(A)$ прогиба балки равен удвоенной полной потенциальной энергии этой балки. Первое слагаемое в правой части (5.23) отвечает удвоенной упругой энергии балки, а второе — взятой с обратным знаком удвоенной работе внешних сил, т.е. заданной распределенной нагрузке $q(x)$. В связи с этим (5.23) обычно называют **функционалом энергии**. Это название часто переносят и на общий случай квадратичного функционала (5.17).

Пусть $w_0 \in D(A)$ — решение (5.18). Подставляя $q = (EIw_0)''$ в (5.23) и учитывая, согласно (5.21), что

$$\int_0^l (EIw_0)'' w \, dx = \int_0^l EIw_0'' w'' \, dx, \quad w \in D(A),$$

получаем

$$\begin{aligned} J[w] &= \int_0^l EI(w'')^2 \, dx - 2 \int_0^l EIw_0'' w'' \, dx = \\ &= \int_0^l EI(w'' - w_0'')^2 \, dx - \int_0^l EI(w_0'')^2 \, dx. \end{aligned}$$

Поскольку $(w'' - w_0'')^2 \geq 0$, то функционал $J[w]$ достигает минимума в $D(A)$ лишь при условии $w'' - w_0'' = 0$, т.е. при $w(x) = w_0(x) + A_1 x + B_1$. Но в силу (5.19) $A_1 = B_1 = 0$, и (как это и следует из теоремы 5.4) функционал $J[w]$ минимален в $D(A)$ при $w(x) = w_0(x)$, $x \in [0, l]$.

Элемент $w_0 \in D(A)$, минимизирующий (5.23), является решением уравнения (5.18), что следует из теоремы 5.4.

Замечание 5.2. Теорема 5.4 утверждает лишь эквивалентность задач нахождения решения $u_0 \in D(A)$ операторного уравнения (5.16) и поиска элемента u^* , минимизирующего функционал (5.17). При этом предполагается, что $D(A)$ всюду плотно

в $\mathcal{H}$. Но нередко возникают ситуации, когда элемент f в (5.16) не принадлежит области $R(A)$ значений этого оператора и решения (5.16) в обычном смысле не существует. Так, в случае разрывной функции $q(x)$ в (5.18) (см. пример 5.3), что характерно для прикладных задач, это уравнение не имеет решения в $D(A)$, а значит, и функционал (5.23) не достигает минимума в $D(A)$.

Пусть $\{u_n\}$ — последовательность элементов $u_n \in D(A)$, минимизирующая функционал (5.17). Эта последовательность, будучи фундаментальной в $D(A)$, может сходиться к элементу $u^* \notin D(A)$. Тогда можно рассмотреть некоторое расширение функционала (5.17). В этом случае

$$J[u^*] = \lim_{n \rightarrow \infty} (\langle Au_n, u_n \rangle - 2\langle f, u_n \rangle),$$

и такой функционал может принимать минимальное значение в некотором расширенном по отношению к $D(A)$ множестве. Элемент $u^* \notin D(A)$, минимизирующий этот функционал, является примером *обобщенного решения* операторного уравнения (5.16) в отличие от его *классического решения* $u_0 \in D(A)$. Необходимость построения такого расширенного множества может возникнуть, например, в случае, когда в (5.16) элемент f не принадлежит области значений оператора A ($f \notin R(A)$).

В более общем случае операторному уравнению

$$P(u) = f, \quad f \in \mathcal{H}, \quad (5.24)$$

с оператором P , не обязательно положительным, но непрерывным, не удастся поставить в соответствие минимизируемый функционал. Если $u_0 \in D(P)$ — классическое решение (5.24), то оно при $u = u_0$ удовлетворяет равенству

$$\langle P(u), v \rangle = \langle f, v \rangle, \quad v \in \mathcal{H}. \quad (5.25)$$

Но (5.24) может и не иметь классического решения (например, при $f \notin R(P)$). В этом случае может существовать элемент

$u_* \notin D(P)$, являющийся пределом последовательности $\{u_n\}$ с элементами $u_n \in D(P)$, причем

$$\lim_{n \rightarrow \infty} \langle P(u_n), v \rangle = \langle f, v \rangle, \quad v \in \mathcal{H}.$$

Такой элемент $u_* \in \mathcal{H}$ называют **слабым решением** операторного уравнения (5.24). Если $u_* \in D(P)$, то слабое решение совпадает с классическим. Слабое решение можно рассматривать как частный случай обобщенного и для его нахождения необходимо каким-то образом расширить область возможных решений уравнения (5.24).

5.3. Энергетическое пространство

Пусть на линейном многообразии $D(A) \subset H$, всюду плотном в гильбертовом пространстве $\mathcal{H}$, задан положительно определенный оператор A . Определим на $D(A)$ новое скалярное произведение

$$\langle u, v \rangle_A = \langle Au, v \rangle, \quad u, v \in D(A). \quad (5.26)$$

Нетрудно проверить, что для него выполнены все аксиомы скалярного умножения. В самом деле, в силу симметричности, а значит, и линейности положительно определенного оператора (см. определение 5.1) для произвольных $u, v \in D(A)$

$$\langle u, v \rangle_A = \langle Au, v \rangle = \langle u, Av \rangle = \langle Av, u \rangle = \langle v, u \rangle_A,$$

что доказывает симметричность нового скалярного умножения. Аналогично для любых $u, v, u_1 \in D(A)$ и $\alpha, \alpha_1 \in \mathbb{R}$

$$\begin{aligned} \langle \alpha u + \alpha_1 u_1, v \rangle_A &= \langle A(\alpha u + \alpha_1 u_1), v \rangle = \langle \alpha Au + \alpha_1 Au_1, v \rangle = \\ &= \alpha \langle Au, v \rangle + \alpha_1 \langle Au_1, v \rangle = \alpha \langle u, v \rangle_A + \alpha_1 \langle u_1, v \rangle_A, \end{aligned}$$

что доказывает дистрибутивность и однородность нового скалярного умножения. Неотрицательность скалярного квадрата и его невырожденность следуют из определения 5.1 и (5.26).

Скалярное произведение (5.26) определяет новую норму

$$\|u\|_A = \sqrt{\langle u, u \rangle_A} = \sqrt{\langle Au, u \rangle}, \quad u \in D(A), \quad (5.27)$$

и расстояние

$$\rho_A(u, v) = \|u - v\|_A, \quad u, v \in D(A). \quad (5.28)$$

Скалярное произведение $\langle u, u \rangle_A$ и **норму** $\|u\|_A$ обычно называют **энергетическими**.

Для положительно определенного оператора A связь между новой и прежней нормами устанавливает соотношение

$$\|u\|_A^2 \geq \gamma^2 \|u\|^2, \quad \text{или} \quad \|u\| \leq \frac{1}{\gamma} \|u\|_A, \quad (5.29)$$

с той же самой постоянной $\gamma \neq 0$, что и в (5.11).

Квадратичный функционал $J[u] = \langle u, u \rangle_A - 2\langle f, u \rangle$ (5.17), или **функционал энергии** (см. пример 5.3), с учетом (5.26) и (5.27) можно записать в виде

$$J[u] = \|u\|_A^2 - 2\langle f, u \rangle. \quad (5.30)$$

Пример 5.4. В гильбертовом пространстве $L_2[0, 1]$ функций, суммируемых с квадратом на отрезке $[0, 1]$, рассмотрим оператор $A = -\frac{d^2}{dx^2}$, область определения $D(A)$ которого является линейным многообразием дважды непрерывно дифференцируемых функций $u(x)$, удовлетворяющих условию $u(0) = u(1) = 0$ (см. пример 5.2). В примере 5.10 показано, что этот оператор положительно определен, а из (5.13) и (5.26) следует, что

$$\langle u, v \rangle_A = \langle Au, v \rangle = \int_0^1 u'v' dx = \langle u', v' \rangle, \quad u, v \in D(A). \quad (5.31)$$

Таким образом, новое скалярное произведение функций из $D(A)$ совпадает с прежним скалярным произведением их производных. Ясно, что энергетическое скалярное произведение,

заданное равенством (5.31), можно распространить на более широкое по сравнению с $D(A)$ линейное многообразие, содержащее, например, функции, имеющие кусочно непрерывную производную на отрезке $[0, 1]$.

Для новых нормы и расстояния, согласно (5.27), запишем

$$\|u\|_A = \sqrt{\langle u', u' \rangle} = \left(\int_0^1 (u')^2 dx \right)^{1/2}, \quad (5.32)$$

$$\rho_A(u, v) = \|u - v\|_A = \left(\int_0^1 (u' - v')^2 dx \right)^{1/2}. \quad (5.33)$$

Функционал энергии (5.30), заданный на множестве $D(A)$, в данном случае имеет вид

$$J[u] = \int_0^1 (u')^2 dx - 2 \int_0^1 f u dx.$$

Отсюда видно, что функционал $J[u]$, как энергетические скалярное произведение (5.31) и норма (5.33), может быть определен на более широком, чем $D(A)$, линейном многообразии. #

Множество $D(A)$ с энергетическим скалярным произведением (5.26) является евклидовым пространством. Обозначим его $\mathcal{U}_A$. Из (5.29) следует, что если *последовательность* функций $u_n \in \mathcal{U}_A$, $n \in \mathbb{N}$, является *фундаментальной* в пространстве $\mathcal{U}_A$, то она фундаментальна и в $\mathcal{H}$, а если она сходится к элементу $u \in \mathcal{U}_A$, то, конечно, она сходится к этому элементу и в $\mathcal{H}$, но обратные утверждения, вообще говоря, неверны.

Последовательность $\{u_n\} \subset D(A)$ может быть фундаментальной в гильбертовом пространстве $\mathcal{H}$ (относительно нормы $\|\cdot\|$ этого пространства), но не быть фундаментальной в евклидовом пространстве $\mathcal{U}_A$ (относительно нормы $\|\cdot\|_A$). Действительно, вернемся к примеру 5.4, в котором $\mathcal{H} = L_2[0, 1]$, $A = -\frac{d^2}{dx^2}$ и $D(A)$ — линейное многообразие дважды непрерывно

дифференцируемых на отрезке $[0, 1]$ функций $u(x)$, удовлетворяющих условиям $u(0) = u(1) = 0$. Последовательность функций $u_n(x) = \frac{1}{n} \sin n\pi x$, $n \in \mathbb{N}$, принадлежащих $D(A)$, является фундаментальной в $L_2[0, 1]$, поскольку сходится к функции $u(x) \equiv 0$:

$$\|u_n\|^2 = \int_0^1 \frac{\sin^2 n\pi x}{n^2} dx \leq \frac{1}{n^2} \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty.$$

Однако эта последовательность не является фундаментальной в евклидовом пространстве $\mathcal{U}_A$, поскольку с учетом равенства (5.32) при $n \in \mathbb{N}$

$$\begin{aligned} \|u_n - u_{2n}\|_A^2 &= \int_0^1 (u'_n(x) - u'_{2n}(x))^2 dx = \\ &= \int_0^1 (\pi \cos n\pi x - \pi \cos 2n\pi x)^2 dx = \pi^2 \int_0^1 \left(\frac{1 + \cos 2n\pi x}{2} - \right. \\ &\quad \left. - 2 \cos n\pi x \cos 2n\pi x + \frac{1 + \cos 4n\pi x}{2} \right) dx = \pi^2. \end{aligned}$$

В частности, последовательность $\{u_n\}$ не сходится по энергетической норме, и функция $u(x) = 0$, $x \in [0, 1]$, не является ее пределом по энергетической норме.

*Нормированное пространство $\mathcal{U}_A$ с нормой $\|u\|_A$ (5.27) не обязательно будет полным. Пополнение $\mathcal{U}_A$ называют **энергетическим пространством** и обозначают $\mathcal{H}_A$. Если же $\mathcal{U}_A$ полное, то оно совпадает с энергетическим пространством, но этот случай обычно не представляет практического интереса.*

Перейдем к более общему случаю, когда нормированное пространство $\mathcal{U}_A$ с нормой $\|u\|_A$ (5.27) не является полным. Воспользуемся рассмотренной в теореме 4.23 процедурой построения *пополнения нормированного пространства*. Пространство

$\mathcal{U}_A$ со скалярным произведением $\langle \cdot, \cdot \rangle_A$ является частным случаем нормированного пространства с нормой $\|u\|_A = \sqrt{\langle u, u \rangle_A}$. Пополнив $\mathcal{U}_A$ как нормированное пространство, согласно теореме 4.23, придем к *банахову пространству* $\mathcal{B}$, элементами которого являются классы $\tilde{u}$ эквивалентных фундаментальных относительно энергетической нормы последовательностей $\{u_n\}$ элементов из $D(A)$. В силу полноты $\mathcal{B}$ это пространство станет гильбертовым (обозначим его $\mathcal{H}_A$), если в нем операцию скалярного умножения ввести равенством

$$\langle \tilde{u}, \tilde{v} \rangle_{\mathcal{H}_A} = \lim_{n \rightarrow \infty} \langle u_n, v_n \rangle_A, \quad \tilde{u}, \tilde{v} \in \mathcal{H}_A, \quad (5.34)$$

где $\{u_n\}$ и $\{v_n\}$ — произвольные последовательности из классов эквивалентности $\tilde{u}$ и $\tilde{v}$.

Покажем, что предел в (5.34) существует и конечен, т.е. последовательность $\{\langle u_n, v_n \rangle_A\}$ фундаментальна. Используя *неравенство треугольника* и *аксиомы скалярного умножения* (см. 5.1), запишем

$$\begin{aligned} & |\langle u_m, v_m \rangle_A - \langle u_n, v_n \rangle_A| \leq \\ & \leq |\langle u_m, v_m \rangle_A - \langle u_m, v_n \rangle_A| + |\langle u_m, v_n \rangle_A - \langle u_n, v_n \rangle_A| = \\ & = |\langle u_m, v_m - v_n \rangle_A| + |\langle u_m - u_n, v_n \rangle_A| \leq \\ & \leq \|u_m\|_A \|v_m - v_n\|_A + \|u_m - u_n\|_A \|v_n\|_A. \end{aligned} \quad (5.35)$$

Фундаментальные последовательности $\{u_n\}$ и $\{v_n\}$ ограничены, т.е. существует такое $c > 0$ и такой номер $N \in \mathbb{N}$, что $\|u_n\|_A \leq c$ и $\|v_n\|_A \leq c$ при $n > N$. Тогда при $m > N$ и $n > N$ будем иметь

$$|\langle u_m, v_m \rangle_A - \langle u_n, v_n \rangle_A| \leq c(\|v_m - v_n\|_A + \|u_m - u_n\|_A). \quad (5.36)$$

Таким образом, из фундаментальности $\{u_n\}$ и $\{v_n\}$ следует, согласно определению 4.1, фундаментальность последовательности $\{\langle u_n, v_n \rangle_A\}$.

Покажем также, что значение предела в (5.34) не зависит от выбора последовательностей из классов $\tilde{u}$ и $\tilde{v}$. Выберем произвольные последовательности $\{u'_n\} \in \tilde{u}$ и $\{v'_n\} \in \tilde{v}$. Тогда аналогично (5.35) и (5.36) имеем

$$\begin{aligned} |\langle u_n, v_n \rangle_A - \langle u'_n, v'_n \rangle_A| &\leq \|u_n\|_A \|v_n - v'_n\|_A + \\ &+ \|u_n - u'_n\|_A \|v'_n\|_A \leq c(\|v_n - v'_n\|_A + \|u_n - u'_n\|_A). \end{aligned} \quad (5.37)$$

Так как $\{u_n\} \sim \{u'_n\}$ и $\{v_n\} \sim \{v'_n\}$ (см. теорему 4.23), то справедливы (4.56) и (4.57) в виде

$$\lim_{n \rightarrow \infty} \|u_n - u'_n\|_A = 0 \quad \text{и} \quad \lim_{n \rightarrow \infty} \|v_n - v'_n\|_A = 0. \quad (5.38)$$

Переходя в (5.37) к пределу при $n \rightarrow \infty$ и учитывая (5.38), получаем

$$\lim_{n \rightarrow \infty} |\langle u_n, v_n \rangle_A - \langle u'_n, v'_n \rangle_A| = 0.$$

Отсюда с учетом (5.34) следует $\lim_{n \rightarrow \infty} \langle u_n, v_n \rangle_A = \lim_{n \rightarrow \infty} \langle u'_n, v'_n \rangle_A$.

Для (5.34) аксиома симметрии скалярного умножения выполнена в силу симметрии $\langle \cdot, \cdot \rangle_A$ (см. 5.1), а аксиомы дистрибутивности и однородности — в силу соответствующих свойств $\langle \cdot, \cdot \rangle_A$ и свойств пределов. Ясно, что аксиома неотрицательности скалярного квадрата $\langle \tilde{u}, \tilde{u} \rangle_{\mathcal{H}_A} \geq 0$ верна, поскольку левая часть неравенства является пределом последовательности неотрицательных чисел $\langle u_n, u_n \rangle_A$. Если $\tilde{u} = \tilde{0}$, т.е. $\tilde{u}$ — нулевой элемент в $\mathcal{H}_A$, то такой класс содержит стационарную последовательность, все члены которой равны нулевому элементу 0 в $D(A)$. Поэтому, согласно (5.34), $\langle \tilde{u}, \tilde{u} \rangle_{\mathcal{H}_A} = 0$. Обратно, если

$$0 = \langle \tilde{u}, \tilde{u} \rangle_{\mathcal{H}_A} = \lim_{n \rightarrow \infty} \langle u_n, u_n \rangle_A = \lim_{n \rightarrow \infty} \|u_n\|_A^2,$$

то это означает, что последовательность $\{u_n\}$ эквивалентна стационарной последовательности, все члены которой равны нулевому элементу 0 в $D(A)$, т.е. $\{u_n\} \in \tilde{u} = \tilde{0}$.

Отождествим элемент $u \in D(A)$ с классом $\tilde{u} \in \mathcal{H}_A$, содержащим стационарную последовательность, все члены которой

равны u . Тогда можно считать, что $D(A) \subset \mathcal{H}_A$, причем

$$\langle u, v \rangle_{\mathcal{H}_A} = \langle \tilde{u}, \tilde{v} \rangle_{\mathcal{H}_A} = \lim_{n \rightarrow \infty} \langle u, v \rangle_A = \langle u, v \rangle_A, \quad u, v \in D(A),$$

т.е. для любых элементов из $D(A)$ операции энергетического скалярного умножения, определенные в $D(A)$ и $\mathcal{H}_A$, совпадают. Отметим, что из теоремы 4.23 следует, что $D(A)$ всюду плотно в $\mathcal{H}_A$.

Таким образом, для $\mathcal{H}_A$ выполнены все условия определения 4.11 пополнения нормированного пространства, т.е. построенное гильбертово пространство $\mathcal{H}_A$ является пополнением евклидова пространства $\mathcal{U}_A$.

Замечание 5.3. Описанная процедура пополнения евклидова пространства до гильбертова применима для любого (не только энергетического) евклидова пространства. #

Из процедуры этого построения следует, что $D(A) = \mathcal{U}_A \subset \mathcal{H}_A$, $\overline{\mathcal{U}_A} = \mathcal{H}_A$ (замыкание по энергетической норме). Если рассмотреть евклидово пространство $D(A)$ со скалярным произведением $\langle \cdot, \cdot \rangle$ гильбертова пространства $\mathcal{H}$ и повторить описанную выше процедуру пополнения (относительно нормы в $\mathcal{H}$), то получим само гильбертово пространство $\mathcal{H}$. Поскольку множество всех фундаментальных последовательностей элементов из $D(A)$ относительно нормы в $\mathcal{H}$ включает множество всех фундаментальных последовательностей элементов из $D(A)$ относительно энергетической нормы, то $\mathcal{H}_A \subset \mathcal{H}$.

Теорема 5.5. Если положительно определенный оператор A определен на множестве $D(A)$, всюду плотном в энергетическом пространстве $\mathcal{H}_A$ со скалярным произведением $\langle \cdot, \cdot \rangle_A$ (5.26), то функционал энергии $J[u] = \langle u, u \rangle_A - 2 \langle f, u \rangle$ (5.30) достигает минимума на элементе $u^* \in \mathcal{H}_A$, однозначно определяемом равенством

$$\langle u^*, u \rangle_A = \langle f, u \rangle, \quad u \in \mathcal{H}_A. \quad (5.39)$$

◀ Функционал $\langle f, u \rangle$ при любом фиксированном элементе $f \in \mathcal{H}$ является линейным ограниченным в гильбертовом пространстве $\mathcal{H}$, причем

$$|\langle f, u \rangle| \leq \|f\| \|u\|, \quad u \in \mathcal{H}. \quad (5.40)$$

Но при любом $u \in \mathcal{H}_A$ справедливо (5.29). Тогда вместо (5.40) можно написать $|\langle f, u \rangle| \leq \frac{\|f\|}{\gamma} \|u\|_A$, т.е. этот функционал, если его сузить на $\mathcal{H}_A$, будет ограниченным линейным функционалом в $\mathcal{H}_A$. Поэтому, согласно теореме 5.2 Ф. Рисса, существует такой элемент $u^* \in \mathcal{H}_A$, однозначно определяемый элементом $f \in \mathcal{H}$, что справедливо (5.39). Тогда вместо (5.30) получим

$$\begin{aligned} J[u] &= \langle u, u \rangle_A - 2\langle u^*, u \rangle_A = \langle u, u \rangle_A - 2\langle u^*, u \rangle_A + \langle u^*, u^* \rangle_A - \\ &\quad - \langle u^*, u^* \rangle_A = \langle u - u^*, u - u^* \rangle_A - \langle u^*, u^* \rangle_A = \|u - u^*\|_A^2 - \|u^*\|_A^2, \end{aligned}$$

т.е. функционал энергии достигает минимума при условии $\|u - u^*\|_A^2 = 0$, что в силу аксиомы 3 нормы (см. 4.1) отвечает равенству $u = u^*$ в $\mathcal{H}_A$. ►

Замечание 5.4. Из хода доказательства теоремы 5.5 следует, что функционал энергии (5.30) имеет в энергетическом пространстве $\mathcal{H}_A$ единственный минимум, достигаемый на элементе $u^* \in \mathcal{H}_A$, являющемся единственным обобщенным решением операторного уравнения (5.16) с положительно определенным оператором A .

Напомним, что если $u^* \notin D(A)$, то элемент u^* называют обобщенным решением операторного уравнения (5.16). Если же $u^* \in D(A)$, то элемент u^* совпадает с классическим решением этого уравнения. Если $u^* \notin D(A)$, то (5.16) не имеет классического решения $u_0 \in D(A)$. В самом деле, элемент u_0 , удовлетворяющий уравнению $Au = f$, согласно теореме 5.4, есть точка минимума функционала (5.17), который совпадает с (5.30) в $D(A)$, т.е. должны совпадать элементы $u_0 \in D(A)$ и $u^* \notin D(A)$, что невозможно.

5.4. Однородное операторное уравнение

Перейдем к рассмотрению *однородного операторного уравнения*

$$Au - \lambda Bu = 0, \quad (5.41)$$

где A и B — симметрический и положительно определенный операторы соответственно, действующие в гильбертовом пространстве $\mathcal{H}$, причем $D(A) \subset D(B)$. Число λ назовем **собственным значением** этого однородного **уравнения**, если при этом λ оно имеет решение, отличное от *тривиального решения* $u = 0$. Такое нетривиальное решение называют **собственным элементом операторного уравнения** (5.41), соответствующим собственному значению λ .

В частном случае $B = I$, где I — действующий в $\mathcal{H}$ тождественный оператор, (5.41) переходит в (4.45). Напомним, что в этом случае значение λ и соответствующее ему решение (4.45) называют **собственными значением и элементом оператора** A , а не операторного уравнения.

Пусть $\tilde{u} \in D(A)$ — собственный элемент уравнения (5.41), соответствующий собственному значению $\tilde{\lambda}$, т.е.

$$A\tilde{u} = \tilde{\lambda}B\tilde{u}. \quad (5.42)$$

Если умножить обе части (5.42) скалярно на $\tilde{u} \in D(A)$, то из полученного равенства $\langle A\tilde{u}, \tilde{u} \rangle = \tilde{\lambda} \langle B\tilde{u}, \tilde{u} \rangle$ следует, что в случае **положительного** (а тем более, положительно определенного) оператора A собственные значения λ уравнения (5.41) **положительны**.

Некоторому собственному значению $\tilde{\lambda}$ операторного уравнения (5.41) могут соответствовать несколько линейно независимых собственных элементов, число которых определяет **кратность** этого **собственного значения**. **Собственное значение** называют **простым**, если его кратность равна единице, а в противном случае — **кратным**. Если **линейная**

оболочка собственных элементов, соответствующих некоторому собственному значению $\tilde{\lambda}$, является замкнутым множеством в $\mathcal{H}$, т.е. подпространством (так, в частности, будет при конечной кратности $\tilde{\lambda}$), то говорят о **собственном подпространстве операторного уравнения**.

Пусть $\mathcal{H}_B$ — энергетическое пространство, являющееся пополнением нормированного пространства $D(B)$ по энергетической норме

$$\|u\|_B = \sqrt{\langle u, u \rangle_B} = \sqrt{\langle Bu, u \rangle}, \quad (5.43)$$

индуцированной энергетическим скалярным произведением

$$\langle u, v \rangle_B = \langle Bu, v \rangle. \quad (5.44)$$

Рассмотрим два различных собственных значения λ_1 и λ_2 с соответствующими собственными подпространствами S_1 и S_2 . Эти подпространства будут ортогональными в $\mathcal{H}_B$, т.е.

$$\langle u, v \rangle_B = \langle Bu, v \rangle = \langle Bv, u \rangle = 0, \quad u \in S_1, \quad v \in S_2. \quad (5.45)$$

Действительно, умножив скалярно обе части равенств

$$Au = \lambda_1 Bu, \quad u \in S_1, \quad \text{и} \quad Av = \lambda_2 Bv, \quad v \in S_2,$$

на $v \in S_2$ и $u \in S_1$ соответственно, получим

$$\langle Au, v \rangle = \lambda_1 \langle Bu, v \rangle \quad \text{и} \quad \langle Av, u \rangle = \lambda_2 \langle Bv, u \rangle.$$

В силу симметричности оператора A имеем $\langle Au, v \rangle = \langle Av, u \rangle$. Поэтому $\lambda_1 \langle Bu, v \rangle = \lambda_2 \langle Bv, u \rangle$, или $(\lambda_1 - \lambda_2) \langle u, v \rangle_B = 0$. Поскольку $\lambda_1 \neq \lambda_2$, то приходим к (5.45).

Предположим, что у уравнения (5.41) все собственные значения простые, а их множество счетно. Тогда набор собственных

значений можно представить в виде возрастающей последовательности $\{\lambda_n\}$. Пусть каждому из этих собственных значений λ_n отвечает одномерное собственное подпространство S_n , содержащее собственный элемент $v_n \in S_n$, $n \in \mathbb{N}$. Тогда в соответствии с (5.45) будем иметь $\langle v_m, v_n \rangle_B = 0$ при $m \neq n$, т.е. система $\{v_n\}$ ортогональна в энергетическом пространстве $\mathcal{H}_B$. Если эта система к тому же является *полной* в $\mathcal{H}_B$, то она будет в $\mathcal{H}_B$ *счетным базисом*.

Ясно, что если v_n — собственный элемент уравнения (5.41), соответствующий собственному значению λ_n , то $w_n = \frac{v_n}{\|v_n\|_B}$ также является его собственным элементом, соответствующим λ_n . Если система $\{v_n\}$ полна в $\mathcal{H}_B$, то система $\{w_n\}$ собственных элементов w_n будет в $\mathcal{H}_B$ ортонормированным базисом, т.е. $\langle w_m, w_n \rangle = 0$ при $m \neq n$ и $\|w_n\| = 1$, $n \in \mathbb{N}$. В этом случае любой элемент $u \in D(A)$ можно представить в $\mathcal{H}_B$ рядом Фурье

$$u = \sum_{n=1}^{\infty} a_n w_n \quad (5.46)$$

с коэффициентами $a_n = \langle u, w_n \rangle_B = \langle B u, w_n \rangle$. Согласно *равенству Парсеваля*, получим

$$\|u\|_B^2 = \langle u, u \rangle_B = \langle B u, u \rangle = \sum_{n=1}^{\infty} a_n^2. \quad (5.47)$$

Отметим, что ряд в (5.46) сходится не только в $\mathcal{H}_B$, но и в $\mathcal{H}$, поскольку для нормы в $\mathcal{H}$ справедливо неравенство $\|u\| \leq \gamma \|u\|_B$, $\gamma > 0$.

Для фиксированного $N \in \mathbb{N}$ и произвольного $u \in D(A)$ с учетом равенства $A w_n = \lambda_n B w_n$, $n \in \mathbb{N}$, а также симметричности операторов A и B и линейности скалярного произведения имеем

$$\left\langle A u, \sum_{n=1}^N a_n w_n \right\rangle = \sum_{n=1}^N a_n \langle u, A w_n \rangle = \sum_{n=1}^N a_n \langle u, B w_n \rangle \lambda_n = \sum_{n=1}^N a_n^2 \lambda_n.$$

Отсюда при $N \rightarrow \infty$ в силу непрерывности скалярного произведения, используя (5.46), получаем

$$\langle Au, u \rangle = \sum_{n=1}^{\infty} a_n^2 \lambda_n. \quad (5.48)$$

Замечание 5.5. Любой собственный элемент $\tilde{u}$ уравнения (5.41), соответствующий собственному значению $\tilde{\lambda}$, является собственным и для уравнения

$$(A + cB)u = \lambda Bu, \quad (5.49)$$

где $c \in \mathbb{R}$, но будет соответствовать собственному значению $\lambda(c) = \tilde{\lambda} + c$. Действительно, подставляя в (5.49) $\tilde{u}$ и $\lambda(c)$ вместо u и λ соответственно и учитывая (5.42), приходим к тождеству. Таким образом, путем выбора константы $c > 0$ можно добиться, чтобы в неубывающей последовательности $\{\lambda_n(c)\}$ собственных значений $\lambda_n(c)$ уравнения (5.49) наименьшее из них было бы положительным, т.е. $\lambda_1(c) > 0$, а оператор $A + cB$ в (5.49) — положительно определенным. В самом деле, используя (5.47) и (5.48), при $\lambda_1(c) > 0$ запишем

$$\langle (A + cB)u, u \rangle = \sum_{n=1}^{\infty} a_n^2 \lambda_n(c) \geq \lambda_1(c) \sum_{n=1}^{\infty} a_n^2 = \lambda_1(c) \langle Bu, u \rangle.$$

Но для положительно определенного оператора B верно неравенство $\langle Bu, u \rangle \geq \gamma^2 \|u\|^2$. Поэтому при выполнении условия $\lambda_1(c) > 0$ оператор $A + cB$ будет положительно определенным.

Нахождение собственных значений и соответствующих им собственных элементов операторного уравнения составляет содержание **задачи на собственные значения**. Эти задачи для (5.41) с симметрическим оператором A и для (5.49) с положительно определенным оператором $A + cB$ эквивалентны в том смысле, что собственные значения связаны простым соотношением $\lambda_n(c) = \lambda_n + c$, $n \in \mathbb{N}$. Поэтому без потери общности

примем, что оператор A в (5.41) является положительно определенным. Тогда наряду с энергетическим пространством $\mathcal{H}_B$ можно ввести энергетическое пространство $\mathcal{H}_A$ с энергетической нормой $\|u\|_A^2 = \langle u, u \rangle_A = \langle Au, u \rangle$, в котором система $\{w_n\}$ будет ортогональной:

$$\begin{aligned} \langle w_m, w_n \rangle_A &= \langle Aw_m, w_n \rangle = \lambda_m \langle Bw_m, w_n \rangle = \\ &= \lambda_m \langle w_m, w_n \rangle_B = \lambda_m \delta_{mn}. \quad \# \end{aligned}$$

Важные свойства собственных значений для операторного уравнения (5.41) с положительно определенными операторами A и B устанавливает следующая теорема.

Теорема 5.6. Пусть все собственные значения однородного операторного уравнения $Au - \lambda Bu = 0$ (5.41) образуют возрастающую последовательность $\{\lambda_n\}$, а соответствующие собственные подпространства S_n одномерны. Тогда для любого собственного значения этого уравнения с положительно определенными операторами A и B справедливы равенства

$$\lambda_1 = \min_{u \in D(A) \setminus \{0\}} \frac{\langle Au, u \rangle}{\langle Bu, u \rangle}, \quad \lambda_n = \min_{u \in \mathcal{U}_n} \frac{\langle Au, u \rangle}{\langle Bu, u \rangle}, \quad n > 1, \quad (5.50)$$

где $\mathcal{U}_n$ — множество ненулевых элементов из $D(A)$, ортогональных в $\mathcal{H}_B$ собственным элементам w_k , $k = \overline{1, n-1}$, т.е. удовлетворяющих условиям $\langle Bu, w_k \rangle = 0$. При этом минимум в (5.50) достигается тогда и только тогда, когда u — любой элемент, принадлежащий собственному подпространству S_n уравнения (5.41), соответствующему λ_n .

◀ Рассмотрим сначала в (5.50) случай $n = 1$. Используя (5.47) и (5.48), запишем

$$\frac{\langle Au, u \rangle}{\langle Bu, u \rangle} = \frac{\sum_{n=1}^{\infty} a_n^2 \lambda_n}{\sum_{n=1}^{\infty} a_n^2} \geq \lambda_1,$$

причем равенство возможно лишь при условии $a_n = 0$, $n > 1$. Но тогда, согласно (5.46), $u = a_1 w_1 \in S_1$, т.е. минимум в (5.50) может быть достигнут только на элементе, принадлежащем одномерному собственному подпространству S_1 , соответствующему собственному значению λ_1 .

При $n > 1$ для элемента $u \in \mathcal{U}_n \subset D(A)$, удовлетворяющего условиям теоремы, в (5.46) коэффициенты $a_k = \langle u, w_k \rangle = 0$, $k = \overline{1, n-1}$, так что

$$u = \sum_{k=n}^{\infty} a_k w_k. \quad (5.51)$$

Тогда, снова используя (5.47) и (5.48), запишем

$$\frac{\langle Au, u \rangle}{\langle Bu, u \rangle} = \frac{\sum_{k=n}^{\infty} a_k^2 \lambda_k}{\sum_{k=n}^{\infty} a_k^2} \geq \lambda_n,$$

причем равенство возможно лишь при $a_k = 0$, $k > n$, т.е., согласно (5.51), минимум в (5.50) может быть достигнут только на элементе $u = a_n w_n \in S_n$, принадлежащем одномерному собственному подпространству S_n , соответствующему собственному значению λ_n . ►

Теорема 5.6 устанавливает свойства собственных значений уравнения (5.41) в предположении, что они существуют. Рассмотрим частный случай (5.41), когда $B = I$. Тогда эта теорема устанавливает условия существования последовательности $\{\lambda_n\}$ собственных значений λ_n положительно определенного оператора A и полноту системы $\{w_n\}$ соответствующих им собственных элементов w_n . В этом случае собственные элементы оператора A можно рассматривать как нетривиальные обобщенные решения w_n операторного уравнения $Au = \lambda_n u$ в энергетическом пространстве $\mathcal{H}_A$. Равенство $Aw_n = \lambda_n w_n$ справедливо тогда и только тогда, когда

$$\langle w_n, u \rangle_A = \lambda_n \langle w_n, u \rangle, \quad u \in \mathcal{H}_A. \quad (5.52)$$

Теорема 5.7. Пусть любое множество элементов $u \in \mathcal{H}_A$ энергетического пространства $\mathcal{H}_A$ положительно определенно-го оператора A , ограниченное по энергетической норме $\|u\|_A = \langle Au, u \rangle$, является относительно компактным в гильбертовом пространстве $\mathcal{H}$. Тогда этот оператор имеет неограниченную неубывающую последовательность $\{\lambda_n\}$ положительных собственных значений λ_n , а соответствующие им собственные элементы $w_n \in \mathcal{H}_A$ образуют систему $\{w_n\}$, полную как в $\mathcal{H}_A$, так и в $\mathcal{H}$.

◀ Рассмотрим ограниченный снизу функционал

$$J_0[u] = \frac{\|u\|_A^2}{\|u\|^2}, \quad u \in \mathcal{H}_A, \quad \|u\| > 0, \quad (5.53)$$

и обозначим через λ_1 его точную нижнюю грань, причем $\lambda_1 \geq \gamma^2 > 0$, где γ^2 — константа в неравенстве $\langle Au, u \rangle \geq \gamma^2 \|u\|^2$ для положительно определенного оператора A . Сначала докажем, что существует элемент $w_1 \in \mathcal{H}_A$, для которого $J_0[w_1] = \lambda_1$.

По свойству точной нижней грани для любого $m \in \mathbb{N}$ можно найти такой элемент $u_m \in \mathcal{H}_A$, что

$$\lambda_1 \leq J_0[u_m] = \frac{\|u_m\|_A^2}{\|u_m\|^2} \leq \lambda_1 + \frac{1}{m}.$$

Переходя в неравенстве $\lambda_1 \leq J_0[u_m] \leq \lambda_1 + 1/m$ к пределу при $m \rightarrow \infty$, находим

$$\lim_{m \rightarrow \infty} J_0[u_m] = \lim_{m \rightarrow \infty} \frac{\|u_m\|_A^2}{\|u_m\|^2} = \lambda_1. \quad (5.54)$$

В неравенстве

$$J_0[u_m + tu] = \frac{\|u_m + tu\|_A^2}{\|u_m + tu\|^2} \geq \lambda_1, \quad u \in \mathcal{H}_A, \quad t \in \mathbb{R},$$

заменим квадраты норм скалярными произведениями и запишем

$$\langle u_m + tu, u_m + tu \rangle_A \geq \lambda_1 \langle u_m + tu, u_m + tu \rangle,$$

или, используя свойства скалярного умножения,

$$t^2 (\|u\|_A^2 - \lambda_1 \|u\|^2) + 2t (\langle u_m, u \rangle_A - \lambda_1 \langle u_m, u \rangle) + \\ + \|u_m\|_A^2 - \lambda_1 \|u_m\|^2 \geq 0.$$

Если квадратный трехчлен относительно произвольного $t \in \mathbb{R}$ не меняет знака, то его дискриминант неположителен, т.е.

$$(\langle u_m, u \rangle_A - \lambda_1 \langle u_m, u \rangle)^2 \leq \\ \leq (\|u\|_A^2 - \lambda_1 \|u\|^2) (\|u_m\|_A^2 - \lambda_1 \|u_m\|^2) \leq \\ \leq \|u\|_A^2 (\|u_m\|_A^2 - \lambda_1 \|u_m\|^2) = \|u\|_A^2 \left(\frac{\|u_m\|_A^2}{\|u_m\|^2} - \lambda_1 \right) \|u_m\|^2.$$

Очевидно, что это неравенство справедливо и для элементов $v_m = \frac{u_m}{\|u_m\|} \in \mathcal{H}_A$ последовательности $\{v_n\}$, т.е.

$$|\langle v_m, u \rangle_A - \lambda_1 \langle v_m, u \rangle| \leq \|u\|_A \sqrt{\|v_m\|_A^2 - \lambda_1}. \quad (5.55)$$

При этом $\|v_m\|_A^2 \geq \lambda_1$. Равенство (5.54) запишем в виде

$$\lim_{m \rightarrow \infty} J_0[u_m] = \lim_{m \rightarrow \infty} J_0[v_m \|u_m\|] = \lim_{m \rightarrow \infty} \|v_m\|_A^2 = \lambda_1. \quad (5.56)$$

Положим $u = v_m - v_k$ и подставим в (5.55):

$$|\langle v_m, v_m - v_k \rangle_A - \lambda_1 \langle v_m, v_m - v_k \rangle| \leq \|v_m - v_k\|_A \sqrt{\|v_m\|_A^2 - \lambda_1}.$$

Меняя в этом неравенстве местами индексы m и k , запишем

$$|\langle v_k, v_k - v_m \rangle_A - \lambda_1 \langle v_k, v_k - v_m \rangle| \leq \|v_k - v_m\|_A \sqrt{\|v_k\|_A^2 - \lambda_1}.$$

Сумма левых частей этих неравенств с учетом неравенства треугольника дает

$$|\langle v_m, v_m - v_k \rangle_A - \lambda_1 \langle v_m, v_m - v_k \rangle| + \\ + |\langle v_k, v_k - v_m \rangle_A - \lambda_1 \langle v_k, v_k - v_m \rangle| \geq \\ \geq |\langle v_m - v_k, v_m - v_k \rangle_A - \lambda_1 \langle v_m - v_k, v_m - v_k \rangle| = \\ = \left| \|v_m - v_k\|_A^2 - \lambda_1 \|v_m - v_k\|^2 \right|.$$

Итак, после почленного сложения упомянутых неравенств получаем

$$\begin{aligned} & \left| \|v_m - v_k\|_A^2 - \lambda_1 \|v_m - v_k\|^2 \right| \leq \\ & \leq \|v_m - v_k\|_A \left(\sqrt{\|v_m\|_A^2 - \lambda_1} + \sqrt{\|v_k\|_A^2 - \lambda_1} \right). \end{aligned} \quad (5.57)$$

Согласно (5.56), последовательность $\{v_m\}$ ограничена по энергетической норме, т.е. $\|v_m\|_A < c = \text{const}$, $m \in \mathbb{N}$. Значит, в соответствии с неравенством треугольника $\|v_m - v_k\|_A \leq \|v_m\|_A + \|v_k\|_A < 2c$. Из (5.57) с учетом (5.56) вытекает, что для произвольно выбранного числа $\varepsilon > 0$ существует такой номер N_1 , что при $k > m > N_1$ справедливо неравенство

$$\left| \|v_m - v_k\|_A^2 - \lambda_1 \|v_m - v_k\|^2 \right| < \frac{\varepsilon}{2}. \quad (5.58)$$

По условию теоремы ограниченное по норме $\|\cdot\|_A$ множество элементов $v_m \in \mathcal{H}_A$ относительно компактно в $\mathcal{H}$, т.е. из любого подмножества этого множества можно выделить последовательность, сходящуюся по норме $\|\cdot\|$. Такую подпоследовательность обозначим $\{\tilde{v}_m\}$. Тогда $\{\tilde{v}_m\}$ является фундаментальной последовательностью в $\mathcal{H}$, и поэтому для выбранного числа $\varepsilon > 0$ существует такой номер $N_2 > N_1$, для которого $\|\tilde{v}_m - \tilde{v}_k\| < \frac{\varepsilon}{2\lambda_1}$ при $k > m > N_2$. Но из неравенства (5.58) вытекает, что $\|\tilde{v}_m - \tilde{v}_k\|_A < \varepsilon$ при $k > m > N_2$, т.е. последовательность $\{\tilde{v}_m\}$ фундаментальна в $\mathcal{H}_A$ и в силу полноты $\mathcal{H}_A$ сходится к некоторому элементу $w_1 \in \mathcal{H}_A$, так что $\|\tilde{v}_m - w_1\|_A \rightarrow 0$ при $m \rightarrow \infty$. Поскольку $\|\tilde{v}_m - w_1\|_A \geq \gamma \|\tilde{v}_m - w_1\|$, то эта последовательность сходится к тому же элементу w_1 и в $\mathcal{H}$. Так как $\|\tilde{v}_m\| = 1$ и $\|\tilde{v}_m\|_A^2 \geq \lambda_1$, то с учетом (5.56) имеем

$$\|w_1\| = \lim_{m \rightarrow \infty} \|\tilde{v}_m\| = 1 \quad \text{и} \quad \|w_1\|_A^2 = \lim_{m \rightarrow \infty} \|\tilde{v}_m\|_A^2 = \lambda_1,$$

т.е. $J_0[w_1] = \lambda_1$.

Покажем теперь, что элемент $w_1 \in \mathcal{H}_A$ является обобщенным решением операторного уравнения $Au = \lambda_1 u$ в энергетическом пространстве $\mathcal{H}_A$. Полагая в (5.55) $v_m = \tilde{v}_m$, при $m \rightarrow \infty$

с учетом (5.56) и $\tilde{v}_m \rightarrow w_1$ как в $\mathcal{H}_A$, так и в $\mathcal{H}$, находим

$$\langle w_1, u \rangle_A = \lambda_1 \langle w_1, u \rangle, \quad u \in \mathcal{H}_A, \quad (5.59)$$

что совпадает с (5.52) при $n = 1$. Таким образом, ненулевой элемент $w_1 \in \mathcal{H}_A$ является нормированным собственным элементом оператора A , соответствующим собственному значению λ_1 .

Если $\tilde{\lambda}$ и $\tilde{v}$ — собственное значение и соответствующий ему собственный элемент оператора A , то справедливо неравенство

$$\tilde{\lambda} = \frac{\|\tilde{v}\|_A^2}{\|\tilde{v}\|^2} \geq \inf_{u \in \mathcal{H}_A} \frac{\|u\|_A^2}{\|u\|^2} = \lambda_1,$$

поскольку λ_1 по предположению является точной нижней гранью функционала (5.53). Следовательно, λ_1 является наименьшим собственным значением оператора A и оно положительно.

Обозначим теперь через λ_2 точную нижнюю грань функционала $J_0[u]$ при дополнительном ограничении $\langle u, w_1 \rangle = 0$. Это ограничение сужает множество элементов $u \in \mathcal{H}_A$, на котором следует искать минимум в (5.53). Поэтому $\lambda_2 \geq \lambda_1$. Повторяя с учетом дополнительного ограничения предшествующие рассуждения, можно установить, что λ_2 является вторым собственным значением оператора A и что этому значению соответствует нормированный собственный элемент w_2 , ортогональный w_1 .

Продолжая этот процесс, построим неубывающую последовательность положительных собственных значений λ_n , $n \in \mathbb{N}$, и соответствующую им последовательность $\{w_n\}$ нормированных собственных элементов w_n оператора A . В общем случае в последовательности $\{\lambda_n\}$ некоторые элементы могут совпадать, т.е. соответствующее им собственное подпространство оператора A может быть неодномерно, причем собственные элементы с этими номерами образуют в нем ортонормированный базис, если оно конечномерно.

Это подпространство не может быть бесконечномерным, поскольку $\lambda_n \rightarrow \infty$ при $n \rightarrow \infty$. Докажем это от противного.

Допустим, что последовательность $\{\lambda_n\}$ ограничена, т.е. $\lambda_n \leq \Lambda^2 = \text{const}$, $n \in \mathbb{N}$. Тогда для любого номера $n \in \mathbb{N}$ из (5.52) при $u = w_n \in \mathcal{H}_A$ имеем

$$\langle w_n, w_n \rangle_A = \|w_n\|_A^2 = \lambda_n \langle w_n, w_n \rangle = \lambda_n \|w_n\|^2 = \lambda_n,$$

т.е. $\|w_n\|_A = \sqrt{\lambda_n} \leq \Lambda$. Таким образом, элементы последовательности $\{w_n\}$ образуют в $\mathcal{H}_A$ ограниченное по энергетической норме множество, которое по условию теоремы компактно в $\mathcal{H}$, т.е. из любой части этого множества можно выделить сходящуюся в $\mathcal{H}$ подпоследовательность, которую обозначим $\{\tilde{w}_m\}$. Она фундаментальна в $\mathcal{H}$, и поэтому $\|\tilde{w}_m - \tilde{w}_k\|^2 < 1$ для достаточно больших номеров m и k . Но это невозможно, поскольку собственные элементы $\tilde{w}_m$ ортонормированы в $\mathcal{H}$ и

$$\begin{aligned} \|\tilde{w}_m - \tilde{w}_k\|^2 &= \langle \tilde{w}_m - \tilde{w}_k, \tilde{w}_m - \tilde{w}_k \rangle = \\ &= \|\tilde{w}_m\|^2 - 2\langle \tilde{w}_m, \tilde{w}_k \rangle + \|\tilde{w}_k\|^2 = 2. \end{aligned}$$

Отсюда следует, что $\lambda_n \rightarrow \infty$ при $n \rightarrow \infty$.

Покажем, что система $\{w_n\}$ ортогональных в $\mathcal{H}_A$ собственных элементов оператора A полна в $\mathcal{H}_A$, т.е. является в $\mathcal{H}_A$ счетным базисом. Любое собственное значение λ_n является точной нижней гранью функционала $J_0[u]$ на множестве элементов $u \in \mathcal{H}_A$, удовлетворяющих дополнительным ограничениям $\langle u, w_k \rangle = 0$, $k = \overline{1, n-1}$. Если бы система $\{w_n\}$ была неполна в $\mathcal{H}_A$, то нашлись бы отличные от нулевого элементы, ортогональные в $\mathcal{H}_A$ всем w_n . Обозначая через $\bar{\lambda}$ точную нижнюю грань $J_0[u]$ на указанных ненулевых элементах, из предыдущих этапов доказательства этой теоремы получаем, что $\bar{\lambda}$ — собственное значение оператора A , большее любого λ_n , $n \in \mathbb{N}$. Но это невозможно, поскольку $\lambda_n \rightarrow \infty$ при $n \rightarrow \infty$.

Система $\{w_n\}$ собственных элементов $w_n \in \mathcal{H}_A$ оператора A полна и в $\mathcal{H}$. Действительно, энергетическое пространство $\mathcal{H}_A$ шире области $D(A)$ определения оператора A , всюду плотной в $\mathcal{H}$. Поэтому $\mathcal{H}_A$ также всюду плотно в $\mathcal{H}$. Для произвольных $\varepsilon > 0$ и $f \in \mathcal{H}$ выберем элемент $g \in \mathcal{H}_A$ так, чтобы выполнялось

неравенство $\|f - g\| < \frac{\varepsilon}{2}$. Для этого элемента выберем такую линейную комбинацию $b_1 w_1 + b_2 w_2 + \dots + b_N w_N$ с коэффициентами $b_n \in \mathbb{R}$, $n = \overline{1, N}$, что выполняется условие

$$\left\| g - \sum_{n=1}^N b_n w_n \right\|_A < \frac{\varepsilon \gamma}{2},$$

где $\gamma > 0$ — константа в неравенстве $\|u\|_A \geq \gamma \|u\|$. Тогда, используя неравенство треугольника, можно записать

$$\begin{aligned} \left\| f - \sum_{n=1}^N b_n w_n \right\| &= \left\| (f - g) + g - \sum_{n=1}^N b_n w_n \right\| \leq \|f - g\| + \\ &+ \left\| g - \sum_{n=1}^N b_n w_n \right\| \leq \|f - g\| + \frac{1}{\gamma} \left\| g - \sum_{n=1}^N b_n w_n \right\|_A < \frac{\varepsilon}{2} + \frac{\varepsilon \gamma}{2\gamma} = \varepsilon. \end{aligned}$$

Итак, для произвольного $\varepsilon > 0$ существует линейная комбинация элементов системы $\{w_n\}$, которая отличается от f по норме пространства $\mathcal{H}$ меньше чем на ε . Это означает, что система $\{w_n\}$ собственных элементов оператора A является замкнутой в $\mathcal{H}$. Но замкнутость ортонормированной системы в гильбертовом пространстве равносильна ее полноте [IX]. Значит, система $\{w_n\}$ полна в $\mathcal{H}$. Она также полна и в $\mathcal{H}_A$, поскольку $\mathcal{H}_A \subset \mathcal{H}$. ►

Если положительно определенный оператор A удовлетворяет условию теоремы 5.7, то говорят, что он имеет **дискретный спектр**. Можно показать*, что эта теорема верна, если оператор A симметрический и положительный, а для общего случая операторного уравнения (5.41) она имеет следующую формулировку.

Теорема 5.8. Пусть положительно определенные операторы A и B в гильбертовом пространстве H таковы, что любое множество элементов энергетического пространства $\mathcal{H}_A$, ограниченное по энергетической норме $\|\cdot\|_A$, является относительно

*См.: Мизлин С.Г., 1970.

компактным в энергетическом пространстве $\mathcal{H}_B$. Тогда собственные значения уравнения (5.41) образуют неограниченную неубывающую последовательность $\{\lambda_n\}$ положительных чисел, а из собственных элементов этого уравнения можно сформировать систему $\{v_n\}$, полную как в $\mathcal{H}_A$, так и в $\mathcal{H}_B$.

5.5. Уравнения с вполне непрерывными симметрическими операторами

Пусть $\mathcal{H}$ — гильбертово пространство, а A — вполне непрерывный симметрический оператор, действующий в $\mathcal{H}$. Рассмотрим операторные уравнения

$$Au - \lambda u = f \quad (5.60)$$

и

$$Au - \lambda u = 0. \quad (5.61)$$

В случае конечномерного евклидова пространства для любого симметрического оператора существует ортонормированный базис, состоящий из собственных векторов данного оператора, и решение уравнений (5.60) и (5.61) при их записи в этом базисе не вызывает затруднений. В бесконечномерном пространстве для произвольного линейного симметрического оператора такое утверждение сделать нельзя. Однако если оператор к тому же является *вполне непрерывным*, то аналогичное утверждение можно сделать и в случае бесконечномерных пространств. Итак, выясним свойства *собственных элементов и собственных значений* симметрического вполне непрерывного оператора в гильбертовом пространстве.

Теорема 5.9. Если A — симметрический линейный ограниченный оператор в гильбертовом пространстве $\mathcal{H}$, то наименьшее число $C \geq 0$, для которого выполнено неравенство

$$|\langle Au, u \rangle| \leq C \|u\|^2, \quad u \in \mathcal{H}, \quad (5.62)$$

равно $\|A\|$.

◀ Если наименьшее число, для которого выполнено неравенство (5.62), обозначить C_A , то, имея, согласно (4.26) и (5.2), неравенство

$$|\langle Au, u \rangle| \leq \|Au\| \|u\| \leq \|A\| \|u\|^2,$$

приходим к соотношению $C_A \leq \|A\|$. Отметим, что последнее неравенство справедливо для любого ограниченного оператора в $\mathcal{H}$.

Для симметрического оператора A и любых $u, v, w \in \mathcal{H}$ имеем $\langle A^2 u, u \rangle = \langle Au, Au \rangle$ и

$$\begin{aligned} \langle A(w+v), w+v \rangle - \langle A(w-v), w-v \rangle &= \langle Aw, w \rangle + \langle Av, w \rangle + \\ &+ \langle Aw, v \rangle + \langle Av, v \rangle - \langle Aw, w \rangle + \langle Av, w \rangle + \langle Aw, v \rangle - \langle Av, v \rangle = \\ &= 2\langle Av, w \rangle + 2\langle Aw, v \rangle = 2\langle v, Aw \rangle + 2\langle Aw, v \rangle = 4\langle Aw, v \rangle. \end{aligned}$$

Положив $w = \lambda u$, $v = (1/\lambda)Au$, $\lambda \neq 0$, получим

$$\langle Aw, v \rangle = \left\langle A\lambda u, \frac{1}{\lambda}Au \right\rangle = \|Au\|^2.$$

Используя (5.62) и равенство параллелограмма [IX], имеем

$$\begin{aligned} \|Au\|^2 &= \frac{1}{4} \left\langle A\left(\lambda u + \frac{1}{\lambda}Au\right), \lambda u + \frac{1}{\lambda}Au \right\rangle - \\ &- \frac{1}{4} \left\langle A\left(\lambda u - \frac{1}{\lambda}Au\right), \lambda u - \frac{1}{\lambda}Au \right\rangle \leq \frac{1}{4} C_A \left\| \lambda u + \frac{1}{\lambda}Au \right\|^2 + \\ &+ \frac{1}{4} C_A \left\| \lambda u - \frac{1}{\lambda}Au \right\|^2 = \frac{1}{2} C_A \left(\lambda^2 \|u\|^2 + \frac{1}{\lambda^2} \|Au\|^2 \right). \end{aligned}$$

Для произвольного $u \neq 0$ положим $\lambda = \sqrt{\|Au\|/\|u\|}$. Тогда $\|Au\|^2 \leq C_A \|Au\| \|u\|$, откуда $\|Au\| \leq C_A \|u\|$. Следовательно, $\|A\| \leq C_A$. Имея ранее полученное неравенство $C_A \leq \|A\|$, заключаем, что $\|A\| = C_A$. ►

В предыдущей главе (см. 4.6) рассмотрены свойства спектра вполне непрерывного оператора в банаховом пространстве. Все отличные от нуля точки спектра для такого оператора являются собственными значениями. Однако существование отличных от нуля собственных значений не гарантировалось. Если вполне непрерывный оператор в гильбертовом пространстве $\mathcal{H}$ является к тому же симметрическим, то он всегда имеет ненулевое собственное значение.

Теорема 5.10. Всякий вполне непрерывный симметрический линейный ненулевой оператор A имеет хотя бы одно отличное от нуля собственное значение λ , причем $|\lambda| = \|A\|$.

◀ Согласно теореме 5.9, имеем

$$\sup_{\|u\|=1} |\langle Au, u \rangle| \leq \|A\|.$$

Обозначим через C левую часть записанного неравенства. Тогда C — это наименьшее из всех чисел, для которых $|\langle Au, u \rangle| \leq C$ при $\|u\| = 1$. Пусть v — произвольный элемент, отличный от нуля. Тогда для $u = v/\|v\|$ имеем

$$\left| \left\langle A \frac{v}{\|v\|}, \frac{v}{\|v\|} \right\rangle \right| \leq C,$$

или $|\langle Av, v \rangle| \leq C\|v\|^2$. Согласно теореме 5.9, $C \geq \|A\|$. Таким образом,

$$\sup_{\|u\|=1} |\langle Au, u \rangle| = \|A\|,$$

причем $\|A\| > 0$, так как $A \neq 0$. В силу свойств точной верхней грани существует последовательность $\{u_n\} \subset \mathcal{H}$, такая, что $\|u_n\| = 1$, $n \in \mathbb{N}$, и $|\langle Au_n, u_n \rangle| \rightarrow \|A\|$ при $n \rightarrow \infty$. Из последовательности $\{u_n\}$ можно выбрать подпоследовательность $\{u_{n_k}\}_{k=1}^\infty$, для которой $\langle Au_{n_k}, u_{n_k} \rangle \rightarrow \lambda$ при $k \rightarrow \infty$, причем λ равно либо $\|A\|$, либо $-\|A\|$.

Обозначим $\mathbf{v}_k = \mathbf{u}_{n_k}$, $k \in \mathbb{N}$. Тогда, учитывая равенство $\|\mathbf{v}_k\| = 1$, $k \in \mathbb{N}$, имеем

$$0 \leq \|A\mathbf{v}_k - \lambda\mathbf{v}_k\|^2 = \|A\mathbf{v}_k\|^2 - 2\lambda \langle A\mathbf{v}_k, \mathbf{v}_k \rangle + \lambda^2. \quad (5.63)$$

Поскольку

$$2\lambda \langle A\mathbf{v}_k, \mathbf{v}_k \rangle - \lambda^2 \leq \|A\mathbf{v}_k\|^2 \leq \|A\|^2 = \lambda^2$$

и

$$\lim_{k \rightarrow \infty} \langle A\mathbf{v}_k, \mathbf{v}_k \rangle = \lambda,$$

существует $\lim_{k \rightarrow \infty} \|A\mathbf{v}_k\| = \lambda$. Переходя к пределу в (5.63) при $k \rightarrow \infty$, получаем

$$\lim_{k \rightarrow \infty} \|A\mathbf{v}_k - \lambda\mathbf{v}_k\|^2 = 0,$$

или

$$A\mathbf{v}_k - \lambda\mathbf{v}_k \rightarrow \mathbf{0} \quad \text{при} \quad k \rightarrow \infty. \quad (5.64)$$

Так как оператор A вполне непрерывный, а последовательность $\{\mathbf{v}_k\}$ ограничена в $\mathcal{H}$, то последовательность $\{A\mathbf{v}_k\}$ образует множество, относительно компактное в $\mathcal{H}$. Поэтому из нее можно извлечь сходящуюся подпоследовательность $\{A\mathbf{v}_{k_m}\}$. Учитывая (5.64), имеем сходимость последовательности $\{\mathbf{v}_{k_m}\}$. Обозначим ее предел $\mathbf{v}$. Тогда $A\mathbf{v}_{k_m} \rightarrow A\mathbf{v}$ при $m \rightarrow \infty$, причем $\|\mathbf{v}\| = \lim_{m \rightarrow \infty} \|\mathbf{v}_{k_m}\| = 1$. Согласно (5.64), $A\mathbf{v} = \lambda\mathbf{v}$, $\mathbf{v} \neq \mathbf{0}$, т.е. λ — собственное значение оператора A и $|\lambda| = \|A\|$. ►

Согласно теореме 4.21 и ее следствиям, вполне непрерывный симметрический оператор в гильбертовом пространстве $\mathcal{H}$ имеет конечную или счетную последовательность $\{\lambda_n\}$ собственных значений $\lambda_n \neq 0$, $n \in \mathbb{N}$, где каждое из них повторяется столько раз, какова размерность собственного подпространства оператора (все его собственные подпространства конечно-

мерны), причем

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n| \geq |\lambda_{n+1}| \geq \dots \quad \text{и} \quad \lim_{n \rightarrow \infty} \lambda_n = 0,$$

если последовательность $\{\lambda_n\}$ бесконечна. В силу теоремы 5.10 и следствия 4.2 $\|A\| = |\lambda_1|$. Последовательности $\{\lambda_n\}$ отвечает ортонормированная система $\{\varphi_n\}$ собственных элементов. Ортонормированности всегда можно добиться, поскольку собственные элементы, отвечающие различным собственным значениям, ортогональны, а собственные подпространства, отвечающие различным собственным значениям, конечномерны и в каждом из них можно выбрать ортонормированный базис, применяя процесс ортогонализации Грама — Шмидта.

Теорема 5.11. Пусть A — вполне непрерывный симметрический линейный оператор, действующий в гильбертовом пространстве $\mathcal{H}$, а $\{x_n\}$ — ортонормированная система собственных элементов (не обязательно бесконечная), отвечающая последовательности $\{\lambda_n\}$ собственных значений $\lambda_n \neq 0$. Тогда для любого элемента $u \in \mathcal{H}$ справедливо разложение

$$Au = \sum_n \lambda_n \langle u, x_n \rangle x_n.$$

◀ Пусть сначала система $\{x_n\}$ бесконечна. Для произвольного элемента $u \in \mathcal{H}$ положим

$$v_k = u - \sum_{n=1}^k \langle u, x_n \rangle x_n, \quad k \in \mathbb{N}. \quad (5.65)$$

Так как $\{x_n\}$ — ортонормированная система, то $\langle v_k, x_n \rangle = 0$, $n = \overline{1, k}$, т.е. v_k принадлежит подпространству $\mathcal{H}_k$, образованному элементами, ортогональными $x_1, x_2, \dots, x_k$. Очевидно, что $\mathcal{H}_k$ — линейное многообразие, а его замкнутость следует из непрерывности скалярного умножения. Значит, $\mathcal{H}_k$ можно

рассматривать как гильбертово пространство. Если A_k — сужение оператора A на подпространство $\mathcal{H}_k$, то оператор A_k действует из $\mathcal{H}_k$ в $\mathcal{H}_k$. Действительно, так как $A_k v = Av$ для любого $v \in \mathcal{H}_k$, то при $n = \overline{1, k}$ и $v \in \mathcal{H}_k$ имеем

$$\langle A_k v, x_n \rangle = \langle Av, x_n \rangle = \langle v, Ax_n \rangle = \langle v, \lambda_n x_n \rangle = \lambda_n \langle v, x_n \rangle = 0.$$

Согласно теореме 5.10, верно равенство $\|A_k\| = |\lambda_{k+1}|$. Поэтому в соответствии с (5.65) для v_k получаем

$$\|Av_k\| = \|A_k v_k\| \leq \|A_k\| \|v_k\| = |\lambda_{k+1}| \|v_k\|.$$

Кроме того, в силу ортонормированности системы $\{x_n\}$ имеем

$$\|v_k\|^2 = \|u\|^2 - \sum_{n=1}^k |\langle u, x_n \rangle|^2 \leq \|u\|^2.$$

Поскольку $\lambda_{k+1} \rightarrow 0$ при $k \rightarrow \infty$, а последовательность $\{\|v_k\|\}$ ограничена, имеем

$$\left\| Au - \sum_{n=1}^k \langle u, x_n \rangle Ax_n \right\| = \|Av_k\| \rightarrow 0 \quad \text{при} \quad k \rightarrow \infty.$$

Следовательно,

$$\sum_{n=1}^k \lambda_n \langle u, x_n \rangle x_n \rightarrow Au \quad \text{при} \quad k \rightarrow \infty,$$

и справедливо разложение

$$Au = \sum_{n=1}^{\infty} \lambda_n \langle u, x_n \rangle x_n.$$

Пусть теперь система $\{x_n\}$ конечна, т.е. $\{x_n\} = \{x_n\}_{n=1}^k$. Тогда элемент

$$v = u - \sum_{n=1}^k \langle u, x_n \rangle x_n$$

принадлежит подпространству $\mathcal{H}_k$, которое образуют элементы, ортогональные $x_1, x_2, \dots, x_k$. Если A_k — сужение оператора A на $\mathcal{H}_k$, то в силу теоремы 5.10 оператор A_k является нулевым, поскольку ни один собственный элемент оператора A , отвечающий отличному от нуля собственному значению, не принадлежит $\mathcal{H}_k$. Таким образом, $\mathcal{H}_k = \ker A$ и $v \in \ker A$. Тогда

$$Au = \sum_{n=1}^k \lambda_n \langle u, x_n \rangle x_n,$$

что завершает доказательство теоремы. ►

Теорема 5.12 (теорема Гильберта — Шмидта). Для любого вполне непрерывного симметрического линейного оператора A в гильбертовом пространстве $\mathcal{H}$ существует ортонормированная система $\{x_n\}$ собственных элементов, отвечающих собственным значениям $\lambda_n \neq 0$, такая, что любой элемент $u \in \mathcal{H}$ единственным образом можно представить в виде

$$u = \sum_n c_n x_n + v,$$

где $v \in \ker A$. При этом $c_n = \langle u, x_n \rangle$ для всех возможных номеров n .

◄ Для любого элемента $u \in \mathcal{H}$ его ряд Фурье по ортонормированной системе $\{x_n\}$ собственных элементов сходится к некоторому элементу $w \in \mathcal{H}$ [IX]:

$$w = \sum_n \langle u, x_n \rangle x_n.$$

Учитывая, что $Ax_n = \lambda_n x_n$, получим

$$Aw = \sum_n \lambda_n \langle u, x_n \rangle x_n,$$

поскольку оператор A непрерывный. Согласно теореме 5.11, также имеем

$$Au = \sum_n \lambda_n \langle u, x_n \rangle x_n.$$

Следовательно, $Au - Aw = 0 \in \mathcal{H}$ и элемент $u - w = v \in \ker A$. Таким образом,

$$u = w + v = \sum_n \langle u, x_n \rangle x_n + v = \sum_n c_n x_n + v,$$

где $v \in \ker A$, а $w \perp v$, поскольку v как собственный элемент, отвечающий нулевому собственному значению, ортогонален любому собственному элементу x_n . Итак, искомое представление элемента $u \in \mathcal{H}$ получено.

Докажем единственность этого представления. Пусть существует еще одно представление $u = w' + v'$, где $w' \perp v'$, $v' \in \ker A$. Тогда $w + v = w' + v'$ и $w - w' = v' - v$, причем $\langle w - w', v' - v \rangle = 0$. Следовательно, $w - w' = v' - v = 0 \in \mathcal{H}$, а для элемента w разложение по ортонормированной системе $\{x_n\}$ совпадает с рядом Фурье элемента w [IX], что и доказывает единственность полученного выше представления. ►

Следствие 5.1. Если $\mathcal{H}$ — гильбертово сепарабельное пространство, A — вполне непрерывный симметрический линейный оператор в $\mathcal{H}$, то в $\mathcal{H}$ существует ортонормированный базис, состоящий из собственных элементов оператора A .

◄ Для доказательства утверждения достаточно к ортонормированной системе $\{x_n\}$ собственных элементов оператора A , отвечающих его ненулевым собственным значениям, добавить

счетный или конечный ортонормированный базис подпространства $\ker A$, который в сепарабельном гильбертовом пространстве всегда существует [IX]. ►

Следствие 5.2. Если нуль не является собственным значением вполне непрерывного симметрического линейного оператора A в гильбертовом пространстве $\mathcal{H}$, то в $\mathcal{H}$ существует ортонормированный базис, состоящий из собственных элементов оператора A . #

Перейдем к изучению неоднородного операторного уравнения (5.60), в котором $\lambda \neq 0$ и A — вполне непрерывный симметрический оператор в гильбертовом пространстве $\mathcal{H}$. Пусть собственные значения $\lambda_n \neq 0$ и соответствующие собственные элементы x_n оператора A известны, и элемент $u \in \mathcal{H}$ является решением уравнения (5.60). Тогда, согласно теореме 5.11, имеем

$$\sum_n \lambda_n \langle u, x_n \rangle x_n - \lambda u = f. \quad (5.66)$$

Умножим скалярно обе части (5.66) на x_k :

$$\lambda_k \langle u, x_k \rangle - \lambda \langle u, x_k \rangle = \langle f, x_k \rangle. \quad (5.67)$$

Если λ не равно ни одному собственному значению λ_k , то для любого возможного k получим

$$\langle u, x_k \rangle = \frac{\langle f, x_k \rangle}{\lambda_k - \lambda}.$$

Тогда в соответствии с (5.66) искомое решение имеет вид

$$u = \frac{1}{\lambda} \sum_n \frac{\lambda_n}{\lambda_n - \lambda} \langle f, x_n \rangle x_n - \frac{1}{\lambda} f. \quad (5.68)$$

Наоборот, если ряд в (5.68) сходится, то элемент u в (5.68) является решением уравнения (5.60). Ряд в (5.68) действительно сходится для любого $f \in \mathcal{H}$, поскольку последовательность

$\{S_n\}$ его частичных сумм фундаментальна. Так как последовательность $\{\lambda_n\}$ сходится к нулю при $n \rightarrow \infty$, то сходится и последовательность $\left\{\left(\frac{\lambda_n}{\lambda(\lambda_n - \lambda)}\right)^2\right\}$ и, следовательно, она является ограниченной. Поэтому для $l > k$ имеем

$$\|S_l - S_k\|^2 = \frac{1}{\lambda^2} \sum_{n=k+1}^l \left| \frac{\lambda_n}{\lambda_n - \lambda} \right|^2 |\langle f, x_n \rangle|^2 \leq M \sum_{n=k+1}^l |\langle f, x_n \rangle|^2,$$

где M — верхняя грань последовательности $\left\{\left(\frac{\lambda_n}{\lambda(\lambda_n - \lambda)}\right)^2\right\}$. Ряд

$$\sum_{n=1}^{\infty} |\langle f, x_n \rangle|^2$$

сходится [IX]. Поэтому для произвольного числа $\varepsilon > 0$ найдется такой номер N , что

$$\sum_{n=k+1}^l |\langle f, x_n \rangle|^2 < \frac{\varepsilon}{M} \quad \text{при } l > k > N,$$

а значит, и $\|S_l - S_k\|^2 < \varepsilon$ при $l > k > N$.

Если λ совпадает с некоторым собственным значением, которому отвечает собственное подпространство размерности m , то $\lambda = \lambda_n$, $n = \overline{k, k+m-1}$. В соответствии с (5.67) элемент f должен быть ортогонален собственным элементам x_n , $n = \overline{k, k+m-1}$, а значит, и собственному подпространству, отвечающему λ . В этом случае решение имеет вид

$$u = \sum_n c_n x_n - \frac{1}{\lambda} f, \quad (5.69)$$

где c_n — произвольные действительные числа для $n = \overline{k, k+m-1}$ и $c_n = \frac{\lambda_n}{\lambda} \frac{\langle f, x_n \rangle}{\lambda_n - \lambda}$ для остальных номеров n . Наоборот, если $\langle f, x_n \rangle = 0$, $n = \overline{k, k+m-1}$, коэффициенты c_n для номеров $n = \overline{k, k+m-1}$ выбраны произвольно, а для остальных номеров

$c_n = \frac{\lambda_n \langle f, x_n \rangle}{\lambda_n - \lambda}$, то ряд в (5.69) сходится, и элемент u , определяемый (5.69), является решением уравнения (5.60).

В случае, когда λ совпадает с некоторым собственным значением, но элемент f не ортогонален соответствующим собственным элементам, уравнение (5.60) не имеет решений.

Пример 5.5. Найдем решение интегрального уравнения II рода

$$\int_{-1}^1 (1+ts)u(s) ds - \lambda u(t) = 2t^2 - \frac{t}{3} - \frac{2}{3}. \quad (5.70)$$

Оно эквивалентно операторному уравнению $Au - \lambda u = f$, определяемому функцией $f = 2t^2 - \frac{t}{3} - \frac{2}{3}$ и оператором $A: L_2[-1, 1] \rightarrow L_2[-1, 1]$, действующим по правилу

$$(Au)(t) = \int_{-1}^1 K(t, s)u(s) ds, \quad u \in L_2[-1, 1],$$

где $K(t, s) = 1 + ts$ — непрерывная в квадрате $[-1, 1]^2$ функция, удовлетворяющая условию $K(t, s) = K(s, t)$. Нетрудно доказать, что в этом случае A является симметрическим вполне непрерывным линейным оператором.

Сначала найдем ненулевые собственные значения оператора A . Так как

$$(Au)(t) = \int_{-1}^1 u(s) ds + t \int_{-1}^1 su(s) ds = a + bt, \quad u \in L_2[-1, 1],$$

то собственные функции необходимо искать в виде $\varphi(t) = a + bt$. Подставляя функцию $\varphi(t) = a + bt$ в уравнение $A\varphi = \lambda\varphi$, получаем

$$\int_{-1}^1 (a + bs) ds + t \int_{-1}^1 s(a + bs) ds = \lambda(a + bt),$$

или $(\lambda - 2)a + (\lambda - 2/3)bt = 0$, $t \in [-1, 1]$. Таким образом, собственным значениям $\lambda_1 = 2$ и $\lambda_2 = 2/3$ отвечают собственные функции $\varphi_1(t) = \frac{1}{\sqrt{2}}$ и $\varphi_2(t) = \sqrt{\frac{3}{2}}t$ (при условии $\|\varphi_1\| = \|\varphi_2\| = 1$) и соответствующие одномерные собственные подпространства. Тогда при $\lambda \neq 2, 2/3$ интегральное уравнение (5.70) имеет единственное решение, определяемое (5.68):

$$\begin{aligned} u(t) &= \frac{1}{\lambda} \left(\frac{2}{2-\lambda} \frac{1}{\sqrt{2}} \int_{-1}^1 \left(2s^2 - \frac{s}{3} - \frac{2}{3} \right) \frac{ds}{\sqrt{2}} + \right. \\ &\quad \left. + \frac{2/3}{2/3-\lambda} \sqrt{\frac{3}{2}} t \int_{-1}^1 \left(2s^2 - \frac{s}{3} - \frac{2}{3} \right) \sqrt{\frac{3}{2}} s ds - 2t^2 + \frac{t}{3} + \frac{2}{3} \right) = \\ &= \frac{1}{\lambda} \left(\frac{1}{2-\lambda} \left(\frac{2s^3}{3} - \frac{2s}{3} \right) \Big|_{-1}^1 - \frac{3t}{2-3\lambda} \frac{s^3}{9} \Big|_{-1}^1 - 2t^2 + \frac{t}{3} + \frac{2}{3} \right) = \\ &= \frac{1}{\lambda} \left(-2t^2 - \frac{\lambda t}{2-3\lambda} + \frac{2}{3} \right). \end{aligned}$$

При $\lambda = 2$ функция $f(t) = 2t^2 - \frac{t}{3} - \frac{2}{3}$ ортогональна функции $\varphi_1(t) = \frac{1}{\sqrt{2}}$, т.е. $\langle f, \varphi_1 \rangle = 0$. Следовательно, в соответствии с (5.69) имеем решение

$$\begin{aligned} u(t) &= \frac{c_1}{\sqrt{2}} + \frac{2/3}{2(2/3-2)} \sqrt{\frac{3}{2}} t \int_{-1}^1 \left(2s^2 - \frac{s}{3} - \frac{2}{3} \right) \sqrt{\frac{3}{2}} s ds - \\ &\quad - \frac{1}{2} \left(2t^2 - \frac{t}{3} - \frac{2}{3} \right) = \frac{c_1}{\sqrt{2}} + \frac{t}{12} - t^2 + \frac{t}{6} + \frac{1}{3} = C - t^2 + \frac{t}{4}, \end{aligned}$$

где $c_1 \in \mathbb{R}$ и $C = \frac{c_1}{\sqrt{2}} + \frac{1}{3}$ — произвольное действительное число.

При $\lambda = \frac{2}{3}$ функция $f(t) = 2t^2 - \frac{t}{3} - \frac{2}{3}$ не ортогональна функции $\varphi_2(t) = \sqrt{\frac{3}{2}}t$, так что интегральное уравнение (5.70) решений не имеет. #

Линейные дифференциальные операторы, как правило, являются неограниченными, и, следовательно, не являются непрерывными. Однако в ряде случаев эти операторы имеют обратные операторы, оказывающиеся вполне непрерывными. Сформулируем теорему, затрагивающую такие случаи*.

Теорема 5.13. Пусть A — неограниченный линейный симметрический оператор в сепарабельном гильбертовом пространстве $\mathcal{H}$, а его обратный оператор A^{-1} является вполне непрерывным. Тогда всегда можно выбрать ортонормированный базис $\{x_n\}$ в $\mathcal{H}$, состоящий из собственных элементов оператора A^{-1} , причем нуль не является собственным значением этого оператора, а последовательность $\{\mu_n\}$ собственных значений A^{-1} ($A^{-1}x_n = \mu_n x_n$) содержит бесконечно много различных значений. Спектр оператора A совпадает с последовательностью $\{\lambda_n\}$, $\lambda_n = \frac{1}{\mu_n}$, а x_n — собственные элементы оператора A , отвечающие собственным значениям λ_n . Если λ — *регулярное значение оператора A* , то для всех $u \in \mathcal{H}$ справедливо равенство

$$(A - \lambda I)^{-1}u = \sum_{n=1}^{\infty} \frac{1}{\lambda_n - \lambda} \langle u, x_n \rangle x_n.$$

Дополнение 5.1. Сопряженные пространства и сопряженные операторы

Рассматривая *линейные ограниченные функционалы*, определенные в нормированном пространстве $\mathcal{U}$, как частный случай *линейных ограниченных операторов* с областью значений, лежащей в *банаховом пространстве* $\mathbb{R}$, заключаем, что множество линейных ограниченных функционалов образует банахово пространство (см. 4.5). Его обозначают $\mathcal{U}^*$ и называют *сопряженным* с $\mathcal{U}$. Норму любого функционала $F \in \mathcal{U}^*$ определена

* Доказательство теоремы см., например: Хатсон В., Пим Дж.

равенством

$$\|F\| = \sup_{\|u\| \leq 1} |F(u)|.$$

Пример 5.6. Найдем пространство, сопряженное с $\mathbb{R}^n$, где норма задана равенством

$$\|x\|_{\mathbb{R}^n} = \sqrt{\sum_{i=1}^n x_i^2}, \quad x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n.$$

Пусть e_i , $i = \overline{1, n}$, — стандартный базис в $\mathbb{R}^n$. Тогда для любого элемента $x \in \mathbb{R}^n$ справедливо разложение

$$x = \sum_{i=1}^n x_i e_i.$$

Если F — произвольный линейный функционал, действующий в $\mathbb{R}^n$, то

$$F(x) = F\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i F(e_i)$$

и, согласно неравенству Коши — Буняковского,

$$\begin{aligned} \|F\| &= \sup_{\|x\|_{\mathbb{R}^n} \leq 1} |F(x)| = \sup_{\|x\|_{\mathbb{R}^n} \leq 1} \left| \sum_{i=1}^n x_i F(e_i) \right| \leq \\ &\leq \sup_{\|x\|_{\mathbb{R}^n} \leq 1} \|x\|_{\mathbb{R}^n} \sqrt{\sum_{i=1}^n (F(e_i))^2} = \sqrt{\sum_{i=1}^n (F(e_i))^2}. \end{aligned}$$

Причем, если

$$x_0 = \left(\sum_{i=1}^n (F(e_i))^2 \right)^{-1/2} \sum_{i=1}^n F(e_i) e_i,$$

то $\|x_0\|_{\mathbb{R}^n} = 1$ и

$$F(x_0) = \sqrt{\sum_{i=1}^n (F(e_i))^2},$$

а значит, для любого $F \in \mathcal{U}^* = (\mathbb{R}^n)^*$ имеем

$$\|F\| = \sqrt{\sum_{i=1}^n (F(e_i))^2}.$$

Таким образом, любой линейный функционал $F \in (\mathbb{R}^n)^*$ однозначно определен вектором $(F(e_1), F(e_2), \dots, F(e_n))$. Наоборот, любой вектор $(f_1, f_2, \dots, f_n) \in \mathbb{R}^n$ задает линейный функционал $F \in (\mathbb{R}^n)^*$, действующий по правилу

$$F(x) = \sum_{i=1}^n x_i f_i, \quad x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n,$$

причем

$$\|F\| = \sqrt{\sum_{i=1}^n f_i^2}.$$

Нетрудно показать, что соответствие между функционалами F и $(F(e_1), F(e_2), \dots, F(e_n))$, где $F \in (\mathbb{R}^n)^*$, является *изоморфизмом*, сохраняющим нормы соответствующих элементов, т.е. *пространства* $(\mathbb{R}^n)^*$ и $\mathbb{R}^n$ *изометричны*. Поэтому можно считать, что с точностью до *изометрии* $(\mathbb{R}^n)^* = \mathbb{R}^n$. #

Рассмотрим банахово *пространство* $L_p[a, b]$, $p > 1$, *функций, суммируемых с p -й степенью*, норма в котором определяется равенством

$$\|x\|_p = \left(\int_a^b |x(t)|^p dt \right)^{1/p}.$$

Можно доказать*, что банахово пространство, сопряженное с $L_p[a, b]$, изометрично банахову пространству $L_q[a, b]$, где $1/p + 1/q = 1$, т.е. $L_p^*[a, b] = L_q[a, b]$. Любой линейный ограниченный функционал $F \in L_p^*[a, b]$ можно представить в виде

$$F(x) = \int_a^b x(t) f(t) dt, \quad x \in L_p[a, b],$$

где f — некоторая функция из $L_q[a, b]$, причем $\|F\| = \|f\|_q$. В частности, $L_2^*[a, b] = L_2[a, b]$.

С банаховым пространством $L_1[a, b]$ суммируемых функций сопряжено банахово пространство $L_\infty[a, b]$ функций f , почти всюду ограниченных на отрезке $[a, b]$, с нормой $\|f\|_\infty$, которая определяется формулой

$$\|f\|_\infty = \inf_{\mu E=0} \sup_{t \in [a, b] \setminus E} |f(t)|,$$

где μE обозначает меру Лебега множества $E \subset [a, b]$. Любой функционал $F \in L_1^*[a, b]$ имеет вид

$$F(x) = \int_a^b x(t) f(t) dt, \quad x \in L_1[a, b],$$

где f — некоторая функция из $L_\infty[a, b]$, а $\|F\| = \|f\|_\infty$.

Согласно теореме Ф. Рисса, любой линейный ограниченный функционал F в гильбертовом пространстве $\mathcal{H}$ имеет вид $F(x) = \langle x, u \rangle$, $x \in \mathcal{H}$, где u — некоторый элемент из $\mathcal{H}$, причем $\|F\| = \|u\|$. Несложно установить, что $\mathcal{H}^*$ изометрично $\mathcal{H}$, при этом используют запись $\mathcal{H}^* = \mathcal{H}$.

В пространстве $\mathcal{U}^*$, сопряженном с нормированным пространством $\mathcal{U}$, рассматривают два вида сходимости: **сильную** — сходимость по норме в $\mathcal{U}^*$ (в этом случае сходимость последовательности $\{F_n\}$ к функционалу F означает, что

*См., например. Люстерник Л. А., Соболев В. И.

$\|F_n - F\| \rightarrow 0$ при $n \rightarrow \infty$) и *слабую* — *поточечную сходимость* (в этом случае сходимость последовательности $\{F_n\}$ к функционалу F означает, что $F_n(x) \rightarrow F(x)$ при $n \rightarrow \infty$ для любого элемента $x \in \mathcal{U}$).

Из сильной сходимости последовательности $\{F_n\} \subset \mathcal{U}^*$ вытекает ее слабая сходимость, причем к тому же пределу. Действительно, пусть $\|F_n - F\| \rightarrow 0$ при $n \rightarrow \infty$. Тогда для любого $x \in \mathcal{U}$ имеем $|F_n(x) - F(x)| \leq \|F_n - F\| \|x\|_{\mathcal{U}} \rightarrow 0$ при $n \rightarrow \infty$. Следовательно, $F_n(x) \rightarrow F(x)$ при $n \rightarrow \infty$.

В случае гильбертова пространства $\mathcal{H}$, учитывая, что $\mathcal{H} = \mathcal{H}^*$, а любой линейный ограниченный функционал имеет специальный вид (см. теорему 5.2), можно утверждать, что слабая сходимость последовательности $\{u_n\}$ к элементу $u \in \mathcal{H}$ равносильна условию: $\langle u_n, x \rangle \rightarrow \langle u, x \rangle$ при $n \rightarrow \infty$ для любого $x \in \mathcal{H}$.

Пример 5.7. Покажем, что последовательность, сходящаяся слабо, может не сходиться сильно. Рассмотрим гильбертово пространство $L_2[0, \pi]$ и в нем последовательность функций $f_n(t) = \sin nt$, $n \in \mathbb{N}$. Так как $\langle f_m, f_n \rangle = 0$ при $m \neq n$ и

$$\|f_n\|^2 = \int_0^\pi \sin^2 nt \, dt = \frac{\pi}{2}, \quad n \in \mathbb{N},$$

то $\|f_m - f_n\|^2 = \|f_m\|^2 + \|f_n\|^2 = \pi$. Значит, в силу определения 4.1 последовательность $\{f_n\}$ не является фундаментальной в $L_2[0, \pi]$ и не сходится сильно. В то же время $\{f_n\}$ сходится слабо к нулевой функции из $L_2[0, \pi]$. В самом деле, для любой функции $x \in L_2[0, \pi]$ имеем

$$\langle x, f_n \rangle = \int_0^\pi x(t) \sin nt \, dt = c_n,$$

где c_n — коэффициенты Фурье функции $x \in L_2[0, \pi]$ по ортогональной системе $\{f_n\}$. Так как ряд $\sum_{n=1}^\infty c_n^2$ сходится [IX], то

$c_n \rightarrow 0$ при $n \rightarrow \infty$. Это и означает, что последовательность $\{f_n\}$ сходится слабо в $L_2[0, \pi]$ к нулевой функции. #

Пространство, сопряженное с $\mathcal{U}^*$, называют **вторым сопряженным** с нормированным пространством $\mathcal{U}$ и обозначают $\mathcal{U}^{**}$. Таким образом, $\mathcal{U}^{**} = (\mathcal{U}^*)^*$. Выясним, какова связь между $\mathcal{U}$ и $\mathcal{U}^{**}$. Предварительно докажем теорему, которая является следствием *теоремы Хана — Банаха*.

Теорема 5.14. Пусть $\mathcal{U}$ — нормированное пространство и элемент $u_0 \in \mathcal{U}$ не является нулевым. Тогда существует линейный ограниченный функционал F , такой, что $F(u_0) = \|u_0\|_{\mathcal{U}}$, $\|F\| = 1$.

◀ Пусть $\mathcal{L} = \{tu_0\}$, $t \in \mathbb{R}$, т.е. $\mathcal{L}$ — одномерное подпространство в $\mathcal{U}$. Зададим на $\mathcal{L}$ линейный функционал $F(tu_0) = t\|u_0\|_{\mathcal{U}}$, $t \in \mathbb{R}$. Отметим, что $F(u_0) = \|u_0\|_{\mathcal{U}}$. Тогда для любого элемента $u \in \mathcal{L}$ имеем $F(u) = \|u\|_{\mathcal{U}}$. Следовательно, норма $\|F\|$ функционала F , определенного на одномерном нормированном пространстве $\mathcal{L}$, равна единице. По теореме Хана — Банаха функционал F можно продолжить до линейного ограниченного функционала, заданного на $\mathcal{U}$, с нормой, равной единице. ▶

Теорема 5.15. Пусть $\mathcal{U}$ — банахово пространство. Тогда $\mathcal{U}$ изометрично некоторому подпространству в $\mathcal{U}^{**}$.

◀ Любому элементу $u_0 \in \mathcal{U}$ поставим в соответствие функционал J_{u_0} , определенный на сопряженном с $\mathcal{U}$ пространстве $\mathcal{U}^*$ соотношением $J_{u_0}(F) = F(u_0)$, $F \in \mathcal{U}^*$.

Функционал J_{u_0} является линейным. Действительно, для любых $F_1, F_2 \in \mathcal{U}^*$ и $\alpha, \beta \in \mathbb{R}$ имеем

$$\begin{aligned} J_{u_0}(\alpha F_1 + \beta F_2) &= (\alpha F_1 + \beta F_2)(u_0) = \\ &= \alpha F_1(u_0) + \beta F_2(u_0) = \alpha J_{u_0}(F_1) + \beta J_{u_0}(F_2). \end{aligned}$$

Функционал J_{u_0} ограниченный, поскольку для любого элемента $F \in \mathcal{U}^*$ с нормой $\|F\| \leq 1$ справедливо неравенство $|J_{u_0}(F)| = |F(u_0)| \leq \|F\| \|u_0\|_{\mathcal{U}} \leq \|u_0\|_{\mathcal{U}}$. Из этих неравенств также заключаем, что $\|J_{u_0}\| \leq \|u_0\|_{\mathcal{U}}$.

Итак, определено отображение $\varphi: \mathcal{U} \rightarrow \mathcal{U}^{**}$ банахова пространства $\mathcal{U}$ на некоторое подмножество $\mathcal{U}^{**}$. Докажем, что это отображение линейное. В самом деле, для любых $u_0, v_0 \in \mathcal{U}$ и $\alpha, \beta \in \mathbb{R}$ имеем

$$\begin{aligned} J_{\alpha u_0 + \beta v_0}(F) &= F(\alpha u_0 + \beta v_0) = \alpha F(u_0) + \beta F(v_0) = \\ &= \alpha J_{u_0}(F) + \beta J_{v_0}(F), \quad F \in \mathcal{U}^*. \end{aligned}$$

Согласно теореме 5.14, из условия $J_{u_0} = 0 \in \mathcal{U}^*$ следует $u_0 = 0 \in \mathcal{U}$, и отображение φ является взаимно однозначным. Легко показать, что образом банахова пространства $\mathcal{U}$ при отображении φ будет некоторое подпространство $\mathcal{U}^{**}$.

Докажем, что отображение φ сохраняет нормы соответствующих элементов. Выше было показано, что $\|J_{u_0}\| \leq \|u_0\|_{\mathcal{U}}$. Согласно теореме 5.14, для любого ненулевого элемента $u_0 \in \mathcal{U}$ существует функционал $F_0 \in \mathcal{U}^*$, такой, что $\|F_0\| = 1$ и $F_0(u_0) = \|u_0\|_{\mathcal{U}}$. Тогда $J_{u_0}(F_0) = F_0(u_0) = \|u_0\|_{\mathcal{U}}$ и

$$\|J_{u_0}\| = \sup_{\|F\| \leq 1} |J_{u_0}(F)| \geq |J_{u_0}(F_0)| = \|u_0\|_{\mathcal{U}}.$$

Следовательно, $\|J_{u_0}\| = \|u_0\|_{\mathcal{U}}$. ►

Если $\mathcal{U}^{**} = \mathcal{U}$, то **банахово пространство $\mathcal{U}$ называют рефлексивным**. Таким является любое гильбертово пространство $\mathcal{H}$, поскольку $\mathcal{H}^{**} = (\mathcal{H}^*)^* = \mathcal{H}^* = \mathcal{H}$. Рефлексивными также будут и пространства $L_p[a, b]$, $1 < p < \infty$.

Перейдем к определению *сопряженного оператора*. Пусть A — линейный ограниченный оператор, переводящий нормированное пространство $\mathcal{U}$ в нормированное пространство $\mathcal{V}$. Поставим произвольному функционалу $J \in \mathcal{V}^*$ в соответствие функционал F , определенный в $\mathcal{U}$ по правилу

$$J \rightarrow F \iff F(u) = J(Au), \quad u \in \mathcal{U}. \quad (5.71)$$

Докажем, что функционал F , соответствующий функционалу $J \in \mathcal{V}^*$, является линейным ограниченным, т.е. $F \in \mathcal{U}^*$. Дей-

ствительно, для любых $u_1, u_2 \in \mathcal{U}$ и $\alpha, \beta \in \mathbb{R}$ имеем

$$\begin{aligned} F(\alpha u_1 + \beta u_2) &= J(A(\alpha u_1 + \beta u_2)) = J(\alpha A u_1 + \beta A u_2) = \\ &= \alpha J(A u_1) + \beta J(A u_2) = \alpha F(u_1) + \beta F(u_2). \end{aligned}$$

Следовательно, F — линейный функционал.

Для любого элемента $u \in \mathcal{U}$, удовлетворяющего условию $\|u\|_{\mathcal{U}} \leq 1$, имеем

$$|F(u)| = |J(Au)| \leq \|J\|_{\mathcal{V}^*} \|Au\|_{\mathcal{V}} \leq \|J\|_{\mathcal{V}^*} \|A\| \|u\|_{\mathcal{U}} \leq \|J\|_{\mathcal{V}^*} \|A\|,$$

где $\|\cdot\|_{\mathcal{U}}$ и $\|\cdot\|_{\mathcal{V}}$ — нормы в пространствах $\mathcal{U}$ и $\mathcal{V}$, а $\|A\|$ — норма оператора A в нормированном пространстве $\mathcal{L}(\mathcal{U}, \mathcal{V})$ (см. 4.5). Таким образом, $F \in \mathcal{U}^*$, причем $\|F\| \leq \|J\|_{\mathcal{V}^*} \|A\|$.

Итак, рассматриваемое соответствие между функционалами из $\mathcal{V}^*$ и $\mathcal{U}^*$ определяет отображение $A^*: \mathcal{V}^* \rightarrow \mathcal{U}^*$, которое называют **оператором, сопряженным** с линейным ограниченным оператором A . При этом используют обозначение $F = A^*J$, так что равенство в (5.71) принимает вид

$$(A^*J)(u) = J(Au), \quad u \in \mathcal{U}, \quad J \in \mathcal{V}^*. \quad (5.72)$$

Теорема 5.16. Оператор A^* , сопряженный с линейным ограниченным оператором $A: \mathcal{U} \rightarrow \mathcal{V}$, где $\mathcal{U}$ и $\mathcal{V}$ — нормированные пространства, является линейным ограниченным, причем $\|A^*\| = \|A\|$.

◀ Докажем линейность оператора A^* . Для любых $J_1, J_2 \in \mathcal{V}^*$, $\alpha, \beta \in \mathbb{R}$ и $u \in \mathcal{U}$ с учетом (5.72) имеем

$$\begin{aligned} (A^*(\alpha J_1 + \beta J_2))(u) &= (\alpha J_1 + \beta J_2)(Au) = \alpha J_1(Au) + \beta J_2(Au) = \\ &= \alpha(A^*J_1)(u) + \beta(A^*J_2)(u) = (\alpha A^*J_1 + \beta A^*J_2)(u). \end{aligned}$$

Теперь докажем ограниченность оператора A^* . Для всякого элемента $u \in \mathcal{U}$, $\|u\|_{\mathcal{U}} \leq 1$, учитывая (5.72), получаем

$$\begin{aligned} |(A^*J)(u)| &= |J(Au)| \leq \|J\|_{\mathcal{V}^*} \|Au\|_{\mathcal{V}} \leq \\ &\leq \|J\|_{\mathcal{V}^*} \|A\| \|u\|_{\mathcal{U}} \leq \|J\|_{\mathcal{V}^*} \|A\|. \end{aligned}$$

Следовательно,

$$\|A^*J\|_{\mathcal{U}^*} = \sup_{\|u\|_{\mathcal{U}} \leq 1} |(A^*J)(u)| \leq \|J\|_{\mathcal{V}^*} \|A\|.$$

Отсюда

$$\|A^*\| = \sup_{\|J\|_{\mathcal{V}^*} \leq 1} \|A^*J\|_{\mathcal{U}^*} \leq \|A\|.$$

Для произвольного элемента $u_0 \in \mathcal{U}$ с нормой $\|u_0\|_{\mathcal{U}} \leq 1$, согласно теореме 5.14, существует функционал $J_0 \in \mathcal{V}^*$, такой, что $\|J_0\|_{\mathcal{V}^*} = 1$ и $J_0(Au_0) = \|Au_0\|_{\mathcal{V}}$. Поэтому с учетом (5.72) имеем

$$\begin{aligned} \|Au_0\|_{\mathcal{V}} = J_0(Au_0) &= (A^*J_0)(u_0) \leq \|A^*J_0\|_{\mathcal{U}^*} \|u_0\|_{\mathcal{U}} \leq \\ &\leq \|A^*J_0\|_{\mathcal{U}^*} \leq \|A^*\| \|J_0\|_{\mathcal{V}^*} = \|A^*\|. \end{aligned}$$

Следовательно,

$$\|A\| = \sup_{\|u_0\|_{\mathcal{U}} \leq 1} \|Au_0\|_{\mathcal{V}} \leq \|A^*\|.$$

Отсюда, учитывая доказанное выше неравенство $\|A\| \geq \|A^*\|$, приходим к равенству $\|A^*\| = \|A\|$. ►

Пусть $\mathcal{H}$ — гильбертово пространство и A — линейный ограниченный оператор, действующий в $\mathcal{H}$. Так как $\mathcal{H}^* = \mathcal{H}$, то сопряженный с A оператор A^* действует также в $\mathcal{H}$. Произвольный функционал $J \in \mathcal{H}^*$, согласно теореме 5.2 Ф. Рисса, можно представить в виде $J(u) = \langle u, v \rangle$, где v — некоторый фиксированный элемент из $\mathcal{H}$. Оператор A^* , сопряженный с A , ставит функционалу J в соответствие линейный ограниченный функционал F по правилу

$$F = A^*J \iff F(u) = J(Au) = \langle Au, v \rangle, \quad u \in \mathcal{H}.$$

Однако, согласно теореме 5.2 Ф. Рисса, $F(u) = \langle u, v^* \rangle$, где $v^* \in \mathcal{H}$ — элемент, однозначно определяемый функционалом F .

Тогда имеем $\langle Au, v \rangle = \langle u, v^* \rangle$, $u \in \mathcal{H}$. Так как $v \in \mathcal{H}$ отождествлено с $J \in \mathcal{H}^*$, а $v^* \in \mathcal{H}$ — с $F \in \mathcal{H}^*$, и $F = A^*J$, то можно положить $v^* = A^*v$. Следовательно, получаем

$$\langle Au, v \rangle = \langle u, A^*v \rangle, \quad u, v \in \mathcal{H}. \quad (5.73)$$

Равенство (5.73) соответствует определению сопряженного оператора в евклидовом пространстве [IV] и однозначно определяет оператор A^* , сопряженный с A .

Пример 5.8. Рассмотрим в гильбертовом пространстве $L_2[0, 1]$ линейный ограниченный оператор

$$(Au)(t) = \int_0^1 K(t, s)u(s) ds, \quad u \in L_2[0, 1], \quad t \in [0, 1],$$

где $K(t, s)$ — функция, непрерывная в квадрате $[0, 1]^2$. Для любых функций $u, v \in L_2[0, 1]$, учитывая (5.73), имеем

$$\begin{aligned} \langle Au, v \rangle &= \int_0^1 v(t) \left(\int_0^1 K(t, s) u(s) ds \right) dt = \\ &= \int_0^1 \int_0^1 K(t, s) v(t) u(s) ds dt = \\ &= \int_0^1 u(s) \left(\int_0^1 K(t, s) v(t) dt \right) ds = \langle u, A^*v \rangle. \end{aligned}$$

Таким образом, сопряженный с A оператор A^* действует по правилу

$$(A^*u)(s) = \int_0^1 K(t, s)u(t) dt, \quad u \in L_2[0, 1], \quad s \in [0, 1]. \quad \#$$

Найдем оператор A^{**} , сопряженный с оператором A^* . Для любых $u, v \in \mathcal{H}$ с учетом (5.73) имеем

$$\langle u, A^{**}v \rangle = \langle A^*u, v \rangle = \langle v, A^*u \rangle = \langle Av, u \rangle = \langle u, Av \rangle.$$

Следовательно, $A^{**} = A$.

Линейный ограниченный оператор $A: \mathcal{H} \rightarrow \mathcal{H}$ называют **самосопряженным**, если $A^* = A$. Тогда для любых $u, v \in \mathcal{H}$, согласно (5.73), справедливо равенство $\langle Au, v \rangle = \langle u, Av \rangle$. Это соответствует определению симметрического оператора и определению самосопряженного оператора в евклидовом пространстве. Самосопряженными являются операторы многих уравнений математической физики. Например, самосопряженные операторы, действующие в комплексном гильбертовом пространстве, находят применение в квантовой механике.

Дополнение 5.2. Критерий базисности системы функций

Пусть B — банахово пространство, а B^* — его сопряженное пространство, т.е. пространство всех линейных непрерывных функционалов в B . Напомним, что система $\{x_n\} \subset B$ является *замкнутой* в B , если замыкание линейной оболочки этой системы совпадает с B [IX].

Систему $\{x_n\} \subset B$ называют **минимальной** в B , если для любого $k \in \mathbb{N}$ элемент $x_k \notin X_k$, где X_k — замыкание линейной оболочки элементов $\{x_n\}$ при $n \neq k$. *Систему* $\{x_n, y_n\}$ пар элементов $x_n \in B$ и $y_n \in B^*$ называют **биортогональной**, если $y_m(x_n) = 0$ при $n \neq m$, и **биортонормированной**, если

$$y_m(x_n) = \begin{cases} 1, & n = m; \\ 0, & n \neq m. \end{cases}$$

В конечномерном линейном нормированном пространстве элементы биортонормированной системы образуют биортогональные (или взаимные) базисы [IV].

Пусть $\{x_n, y_n\}$ — биортонормированная система. В этом случае $\{y_n\}$ называют **системой, сопряженной** с системой $\{x_n\}$. Если $\{x_n\}$ — замкнутая система в банаховом пространстве B , то сопряженная с ней система (если она существует) единственна. Для существования у данной последовательности $\{x_n\} \subset B$ последовательности $\{y_n\} \subset B^*$, образующей с $\{x_n\}$ биортонормированную систему, необходимо и достаточно, чтобы система $\{x_n\}$ была минимальной в B .

Напомним, что систему $\{x_n\} \subset B$ элементов банахова пространства B называют базисом, если для любого элемента $x \in B$ существует единственный ряд

$$x = \sum_{n=1}^{\infty} a_n x_n, \quad (5.74)$$

сходящийся к x по норме пространства B , где $a_n(x)$ — функционалы, определенные в B . Сформулируем критерий базисности системы функций в банаховом пространстве.

Теорема 5.17. Для того чтобы система $\{x_n\} \subset B$ была базисом в банаховом пространстве B , необходимо и достаточно, чтобы выполнялись три условия:

- 1) $\{x_n\}$ замкнута в B ;
- 2) $\{x_n\}$ минимальна в B ;
- 3) существует такое число $M > 0$, что для любых $x \in B$ и $N \in \mathbb{N}$ верно неравенство

$$\left\| \sum_{n=1}^N y_n(x) x_n \right\| \leq M \|x\|,$$

где $\{y_n\}$ — система, сопряженная с $\{x_n\}$.

Следствие 5.3. Если $\{x_n\}$ — базис в B , то функционалы $a_n(x)$ в (5.74) являются линейными ограниченными функционалами и определяются в B равенствами $a_n(x) = y_n(x)$, где $\{y_n\}$ — система, сопряженная с $\{x_n\}$.

Дополнение 5.3. Положительная определенность эллиптического оператора

В ограниченной области $\Omega \in \mathbb{R}^m$ с кусочно гладкой границей $\partial\Omega$ рассмотрим линейное дифференциальное уравнение с переменными коэффициентами

$$-\sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij}(x) \frac{\partial u(x)}{\partial x_j} \right) + c(x) u(x) = f(x), \quad x \in \Omega, \quad (5.75)$$

и граничными условиями

$$u(x) = 0, \quad x \in \partial\Omega_1; \quad (5.76)$$

$$\sum_{i=1}^m \sum_{j=1}^m a_{ij}(x) \frac{\partial u(x)}{\partial x_i} n_j(x) + \sigma(x) u(x) = 0, \quad x \in \partial\Omega_2, \quad (5.77)$$

на участках $\partial\Omega_1 \subset \partial\Omega$ и $\partial\Omega_2 = \partial\Omega \setminus \partial\Omega_1$. Здесь $u(x) \in C^2(\bar{\Omega})$ — искомая функция; $a_{ij}(x) = a_{ji}(x) \in C^1(\bar{\Omega})$, $c(x), f(x) \in C(\Omega)$ и $\sigma(x) \in C(\partial\Omega_2)$ — заданные функции; $n_j(x)$, $j = \overline{1, m}$, — направляющие косинусы единичного вектора $n \in \mathbb{R}^m$ внешней нормали к $\partial\Omega$. Далее ограничимся случаем $c(x) \geq 0$, $x \in \bar{\Omega}$, и $\sigma(x) \geq 0$, $x \in \partial\Omega_2$.

Рассмотрим гильбертово пространство $\mathcal{H} = L_2(\bar{\Omega})$ функций, суммируемых на $\bar{\Omega}$ с квадратом, в котором скалярное произведение и норма определены соотношениями

$$\langle u, v \rangle = \int_{\Omega} uv d\Omega \quad \text{и} \quad \|u\|^2 = \int_{\Omega} u^2 d\Omega. \quad (5.78)$$

В соответствии с (5.75), (5.76) и (5.77) при помощи равенств

$$Au = -\sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) + cu, \quad Nu = \sum_{i=1}^m \sum_{j=1}^m a_{ij} \frac{\partial u}{\partial x_i} n_j \quad (5.79)$$

введем операторы A и N , которые имеют области определения $D(A) = C^2(\Omega)$ и $D(N) = C^1(\partial\Omega_2)$ соответственно, являющиеся линейными многообразиями в $L_2(\bar{\Omega})$. Нетрудно проверить, что эти операторы являются линейными.

Убедимся в том, что A является симметрическим оператором. Используя (5.79), (5.78) и (2.99), получаем

$$\begin{aligned} \langle Au, v \rangle - \langle Av, u \rangle &= \\ &= - \int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m \left(v \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) - u \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial v}{\partial x_j} \right) \right) d\Omega = \\ &= - \int_{\partial\Omega} \sum_{i=1}^m \sum_{j=1}^m a_{ij} \left(v \frac{\partial u}{\partial x_j} - u \frac{\partial v}{\partial x_j} \right) n_i d(\partial\Omega) = 0, \end{aligned}$$

поскольку на $\partial\Omega_1$, согласно (5.76), $u = v = 0$, а на $\partial\Omega_2$, согласно (5.77),

$$\sum_{i=1}^m \sum_{j=1}^m a_{ij} \left(v \frac{\partial u}{\partial x_j} - u \frac{\partial v}{\partial x_j} \right) n_i = vNu - uNv = -v\sigma u + u\sigma v = 0.$$

Следовательно, $\langle Au, v \rangle = \langle Av, u \rangle$, т.е. оператор A является симметрическим.

В соответствии с (2.98) и (5.76)–(5.78) запишем

$$\begin{aligned} \langle Au, u \rangle &= - \int_{\Omega} u \sum_{i=1}^m \sum_{j=1}^m \left(\frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) - cu \right) d\Omega = \\ &= - \int_{\partial\Omega} u \sum_{i=1}^m \sum_{j=1}^m a_{ij} \frac{\partial u}{\partial x_j} n_i d(\partial\Omega) + \\ &\quad + \int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m \left(\frac{\partial u}{\partial x_i} a_{ij} \frac{\partial u}{\partial x_j} + cu^2 \right) d\Omega = \\ &= \int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m a_{ij} \frac{\partial u}{\partial x_i} \frac{\partial u}{\partial x_j} d\Omega + \int_{\Omega} cu^2 d\Omega + \int_{\partial\Omega_2} \sigma u^2 d(\partial\Omega). \quad (5.80) \end{aligned}$$

Если существует такое число $\mu_0 > 0$, что при любых $x \in \bar{\Omega}$ для произвольных наборов чисел $t_1, t_2, \dots, t_m \in \mathbb{R}$ выполняется неравенство

$$\sum_{i=1}^m \sum_{j=1}^m a_{ij}(x) t_i t_j \geq \mu_0 \sum_i t_i^2, \quad (5.81)$$

то оператор A называют **эллиптическим** на замыкании $\bar{\Omega}$. В этом случае (5.75) относят к уравнениям эллиптического типа. Отметим, что левая часть (5.81) является квадратичной формой, а коэффициенты $a_{ij} = a_{ji}$, $i, j = \overline{1, m}$, — элементами симметрической **матрицы** этой **квадратичной формы**. Тогда выполнение (5.81) означает, что эта квадратичная форма является положительно определенной.

Если существует такое $\mu_0 > 0$, что для любых $x \in \bar{\Omega}$

$$\sum_{i=1}^m \sum_{j=1}^m a_{ij}(x) t_i t_j \leq -\mu_0 \sum_i t_i^2,$$

то оператор A также называют эллиптическим. При этом квадратичная форма, стоящая в левой части этого неравенства, является отрицательно определенной. Но при изучении свойств оператора A при выбранном в (5.75) знаке перед символами сумм и условии $c(x) \geq 0$, $x \in \bar{\Omega}$, существенно выполнение условия эллиптичности оператора именно в виде (5.81).

Если оператор A эллиптический, то из (5.80) и (5.81) с учетом неотрицательности функций c и σ получаем

$$\begin{aligned} \langle Au, u \rangle \geq \mu_0 \int_{\Omega} \sum_{i=1}^m \left(\frac{\partial u}{\partial x_i} \right)^2 d\Omega + \\ + \int_{\Omega} cu^2 d\Omega + \int_{\partial\Omega_2} \sigma u^2 d(\partial\Omega) \geq 0. \end{aligned} \quad (5.82)$$

При $\langle Au, u \rangle = 0$ каждый из интегралов в (5.82) должен обратиться в нуль. Тогда из равенства нулю первого интеграла

следует, что $u \equiv C_0 = \text{const}$, но из граничного условия (5.76) на участке $\partial\Omega_1$ границы области Ω имеем $C_0 = 0$, так что $u \equiv 0$ на замыкании $\bar{\Omega}$. Поэтому, согласно определению 5.1, A является *положительным оператором*.

Но возможен случай, когда $\partial\Omega_1 = \emptyset$, т.е. $\partial\Omega_2 = \partial\Omega$. В этом случае равенство нулю второго и третьего интегралов в (5.82) даст

$$C_0^2 \int_{\Omega} c d\Omega = 0 \quad \text{и} \quad C_0^2 \int_{\partial\Omega} \sigma d(\partial\Omega) = 0.$$

Если одна из непрерывных и неотрицательных функций c или σ отлична от нуля хотя бы в одной точке области Ω или границы $\partial\Omega$ соответственно, то снова имеем $C_0 = 0$, а оператор A будет положительным. Однако это утверждение теряет силу, если $c \equiv 0$ в Ω и $\sigma \equiv 0$ на $\partial\Omega$. При этом решение задачи (5.75)–(5.77) может и не существовать или быть неединственным. Рассмотрим этот случай подробнее.

Интегрируя (5.75) при $c \equiv 0$, $\sigma \equiv 0$ по области Ω , получаем

$$-\int_{\Omega} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial}{\partial x_i} \left(a_{ij} \frac{\partial u}{\partial x_j} \right) d\Omega = \int_{\Omega} f d\Omega.$$

Преобразуя интеграл в левой части этого равенства по формуле Остроградского (1.26) и учитывая, что в данном случае $Nu = 0$ на $\partial\Omega$, устанавливаем необходимое условие

$$\int_{\Omega} f d\Omega = -\int_{\partial\Omega} \sum_{i=1}^m \sum_{j=1}^m a_{ij} \frac{\partial u}{\partial x_j} n_i d(\partial\Omega) = -\int_{\partial\Omega} Nu d(\partial\Omega) = 0 \quad (5.83)$$

существования решения задачи. Но даже если условие (5.83) выполнено и решение u существует, то функция, отличающаяся от u на произвольную константу, также является решением этой задачи. Подчиним решение u дополнительному условию,

аналогичному (5.83):

$$\int_{\Omega} u d\Omega = 0. \quad (5.84)$$

Тогда при $u = C_0$ имеем $C_0 = 0$, т.е. и в этом случае (при $c \equiv 0$ в Ω и $\sigma \equiv 0$ на $\partial\Omega$) оператор A будет положительным. Следовательно, если решение $u \in D(A)$ существует, то в силу теоремы 5.3 оно единственно в $D(A)$.

Выясним, при каких условиях оператор A , будучи эллиптическим, является *положительно определенным оператором*. Так как первый и третий интегралы в правой части (5.82) неотрицательны, то при $c(x) \geq c_0 > 0$, $x \in \bar{\Omega}$, с учетом (5.78) имеем

$$\langle Au, u \rangle \geq \int_{\Omega} cu^2 d\Omega \geq c_0 \int_{\Omega} u^2 d\Omega = c_0 \|u\|^2.$$

Таким образом, если $\sigma(x) \geq 0$ на $\partial\Omega_2 \subset \partial\Omega$ и $c(x) \geq c_0 > 0$ в Ω , то оператор A , согласно определению 5.1, является положительно определенным.

Если же $c_0 = 0$, то второй интеграл в правой части (5.82) неотрицателен, так что получаем

$$\langle Au, u \rangle \geq \mu_0 \int_{\Omega} \sum_{i=1}^m \left(\frac{\partial u}{\partial x_i} \right)^2 d\Omega + \int_{\partial\Omega_2} \sigma u^2 d(\partial\Omega). \quad (5.85)$$

Положим $u = vw$ и вычислим

$$\begin{aligned} \sum_{i=1}^m \left(\frac{\partial u}{\partial x_i} \right)^2 &= v^2 \sum_{i=1}^m \left(\frac{\partial w}{\partial x_i} \right)^2 + 2vw \sum_{i=1}^m \frac{\partial v}{\partial x_i} \frac{\partial w}{\partial x_i} + w^2 \sum_{i=1}^m \left(\frac{\partial v}{\partial x_i} \right)^2 = \\ &= v^2 \sum_{i=1}^m \left(\frac{\partial w}{\partial x_i} \right)^2 + \sum_{i=1}^m v \frac{\partial v}{\partial x_i} \frac{\partial w^2}{\partial x_i} + w^2 \sum_{i=1}^m \left(\frac{\partial}{\partial x_i} \left(v \frac{\partial v}{\partial x_i} \right) - v \frac{\partial^2 v}{\partial x_i^2} \right) = \\ &= v^2 \sum_{i=1}^m \left(\frac{\partial w}{\partial x_i} \right)^2 + \sum_{i=1}^m \frac{\partial}{\partial x_i} \left(w^2 v \frac{\partial v}{\partial x_i} \right) - w^2 v \sum_{i=1}^m \frac{\partial^2 v}{\partial x_i^2}. \end{aligned}$$

Отсюда, отбрасывая первую (неотрицательную) сумму в правой части равенства и интегрируя по области Ω , получаем

$$\int_{\Omega} \sum_{i=1}^m \left(\frac{\partial u}{\partial x_i} \right)^2 d\Omega \geq \int_{\Omega} \sum_{i=1}^m \frac{\partial}{\partial x_i} \left(w^2 v \frac{\partial v}{\partial x_i} \right) d\Omega - \int_{\Omega} w^2 v \sum_{i=1}^m \frac{\partial^2 v}{\partial x_i^2} d\Omega.$$

Положив в формуле Остроградского (1.25) $u_i = w^2 v \frac{\partial v}{\partial x_i}$, преобразуем первый интеграл в правой части этого неравенства:

$$\begin{aligned} \int_{\Omega} \sum_{i=1}^m \frac{\partial}{\partial x_i} \left(w^2 v \frac{\partial v}{\partial x_i} \right) d\Omega &= \\ &= \int_{\partial\Omega} \sum_{i=1}^m w^2 v \frac{\partial v}{\partial x_i} n_i d(\partial\Omega) = \int_{\partial\Omega} w^2 v \frac{\partial v}{\partial n} d(\partial\Omega), \end{aligned}$$

где n_i , $i = \overline{1, m}$, — направляющие косинусы единичного вектора $\mathbf{n}$ внешней нормали к границе области Ω , а $\frac{\partial v}{\partial n}$ — производная функции v по направлению этой нормали. Тогда вместо последнего неравенства будем иметь

$$\begin{aligned} - \int_{\Omega} w^2 v \sum_{i=1}^m \frac{\partial^2 v}{\partial x_i^2} d\Omega &\leq \int_{\Omega} \sum_{i=1}^m \left(\frac{\partial u}{\partial x_i} \right)^2 d\Omega - \int_{\partial\Omega} w^2 v \frac{\partial v}{\partial n} d(\partial\Omega) \leq \\ &\leq \int_{\Omega} \sum_{i=1}^m \left(\frac{\partial u}{\partial x_i} \right)^2 d\Omega + \left| \int_{\partial\Omega} w^2 v \frac{\partial v}{\partial n} d(\partial\Omega) \right|. \quad (5.86) \end{aligned}$$

Начало координат в $\mathbb{R}^m$ можно выбрать так, что все точки из $\overline{\Omega}$ будут внутренними точками декартова произведения m отрезков $[0, b_i]$, $i = \overline{1, m}$, т.е. область Ω вместе со своей границей $\partial\Omega$ будет полностью погружена в m -мерный прямоугольный параллелепипед со сторонами b_i , $i = \overline{1, m}$. Положим

$$v = \prod_{i=1}^m \sin \frac{\pi x_i}{b_i}$$

и вычислим

$$\sum_{i=1}^m \frac{\partial^2 v}{\partial x_i^2} = - \sum_{i=1}^m \frac{\pi^2}{b_i^2} \prod_{i=1}^m \sin \frac{\pi x_i}{b_i} = -qv, \quad q = \pi^2 \sum_{i=1}^m \frac{1}{b_i^2}.$$

Тогда для интеграла в левой части (5.86) получим

$$- \int_{\Omega} w^2 v \sum_{i=1}^m \frac{\partial^2 v}{\partial x_i^2} d\Omega = q \int_{\Omega} w^2 v^2 d\Omega = q \int_{\Omega} u^2 d\Omega. \quad (5.87)$$

Так как функция v в любой точке границы $\partial\Omega$ положительна, то для второго интеграла в правой части (5.86) имеем

$$\left| \int_{\partial\Omega} w^2 v \frac{\partial v}{\partial n} d(\partial\Omega) \right| = \left| \int_{\partial\Omega} \frac{u^2}{v} \frac{\partial v}{\partial n} d(\partial\Omega) \right| \leq \int_{\partial\Omega} \frac{u^2}{v} \left| \frac{\partial v}{\partial n} \right| d(\partial\Omega).$$

Ясно, что на границе $\partial\Omega$ величина $\frac{1}{v} \left| \frac{\partial v}{\partial n} \right|$ ограничена, т.е. $\frac{1}{v} \left| \frac{\partial v}{\partial n} \right| \leq C_1 = \text{const} > 0$, что приводит к оценке

$$\left| \int_{\partial\Omega} w^2 v \frac{\partial v}{\partial n} d(\partial\Omega) \right| \leq C_1 \int_{\partial\Omega} u^2 d(\partial\Omega). \quad (5.88)$$

Подставляя (5.87) и (5.88) в (5.86) и обозначая через C большее из чисел $1/q$ и C_1/q , получаем неравенство

$$\int_{\Omega} u^2 d\Omega \leq C \left(\int_{\Omega} \sum_{i=1}^m \left(\frac{\partial u}{\partial x_i} \right)^2 d\Omega + \int_{\partial\Omega} u^2 d(\partial\Omega) \right), \quad C > 0. \quad (5.89)$$

Так как функция $u \in D(A)$ удовлетворяет граничным условиям (5.76) и (5.77), то $\partial\Omega$ в (5.89) следует заменить на $\partial\Omega_2$.

Если функция σ в неравенстве (5.85) ограничена снизу значением $\sigma_0 > 0$, то, обозначая через μ меньшее из чисел μ_0 и σ_0 , из (5.85) находим

$$\langle Au, u \rangle \geq \mu \left(\int_{\Omega} \sum_{i=1}^m \left(\frac{\partial u}{\partial x_i} \right)^2 d\Omega + \int_{\partial\Omega_2} u^2 d(\partial\Omega) \right), \quad \mu > 0.$$

Сравнивая это неравенство с (5.89) при замене $\partial\Omega$ на $\partial\Omega_2$ и учитывая (5.78), в итоге получаем

$$\langle Au, u \rangle \geq \frac{\mu}{C} \int_{\Omega} u^2 d\Omega = \gamma^2 \|u\|^2, \quad \gamma^2 = \frac{\mu}{C}. \quad (5.90)$$

Таким образом, в случае $c_0 = 0$ оператор A является положительно определенным при условии, что $\sigma \geq \sigma_0 > 0$ на $\partial\Omega_2 \subset \partial\Omega$.

При $c \equiv 0$ в Ω и $\partial\Omega_2 = \emptyset$ из (5.85) и (5.89) следует, что оператор A остается положительно определенным. Можно показать*, что A является положительно определенным при $c \equiv 0$ в области Ω и $\sigma \equiv 0$ на части $\partial\Omega_2 \subset \partial\Omega$ границы этой области. Если же при этом $\partial\Omega_2 = \partial\Omega$, то оператор A сохраняет положительную определенность лишь на множестве $D_*(A) \subset \subset D(A)$ функций, удовлетворяющих дополнительному условию (5.84).

Итак, оператор A , определенный первым равенством (5.79), будет положительно определенным, если он является эллиптическим, т.е. для него выполнено (5.81), и при этом $c \geq c_0 > 0$ в Ω и $\sigma \geq 0$ на $\partial\Omega_2$. При $c \equiv 0$ в Ω оператор A , будучи эллиптическим, сохранит положительную определенность, если либо $\partial\Omega_2 = \emptyset$, либо $\sigma \geq \sigma_0 > 0$ на $\partial\Omega_2 \subset \partial\Omega$, либо $\sigma \equiv 0$ на $\partial\Omega_2 \subset \partial\Omega$ (т.е. при $\partial\Omega_1 \neq \emptyset$). Наконец, при $c \equiv 0$ в Ω и $\sigma \equiv 0$ на $\partial\Omega$ необходимо выполнить дополнительное условие (5.84).

Пример 5.9. Несложно проверить, что оператор Лапласа, представленный в виде

$$Au = -\nabla^2 u = -\sum_{i=1}^3 \frac{\partial^2 u}{\partial x_i^2},$$

является эллиптическим, поскольку для выполнения условия (5.81) достаточно принять $\mu_0 = 1$. Установлено [XV], что этот

*См., например: Мизлин С.Г., 1968 или Ректорис К

оператор является симметрическим и положительным при условии $u = 0$ на всей границе области. Изложенное выше показывает, что при этом условии оператор Лапласа, определенный на множестве функций, дважды непрерывно дифференцируемых в области Ω и удовлетворяющих однородным граничным условиям вида (5.76), (5.77), будет и положительно определенным при указанных выше ограничениях для σ на $\partial\Omega_2$ (при $c \equiv 0$ в Ω).

Эллиптическими являются и операторы дифференциальных уравнений, описывающих перенос физических субстанций в неоднородных и анизотропных средах и вытекающих из соответствующих законов сохранения (см. часть I). Например, в уравнении (2.96) нестационарной теплопроводности *тензор теплопроводности* $\tilde{\lambda}$ может быть представлен положительно определенной матрицей (2.95) третьего порядка. Квадратичная форма с такой матрицей в декартовой прямоугольной системе координат соответствует в общем случае *трехосному эллипсоиду*, а элементы λ_{ij} , $i, j = 1, 2, 3$, этой матрицы удовлетворяют неравенству (5.81). Таким образом, из вышеизложенного следует, что оператор A , определенный первым равенством в (5.79) при $a_{ij} = \lambda_{ij}$, является положительно определенным на множестве функций, дважды непрерывно дифференцируемых в области Ω и удовлетворяющих граничным условиям (5.76), (5.77) при указанных выше ограничениях для c в Ω и σ на $\partial\Omega_2$. Ясно, что при $a_{ij}(x) = \lambda(x) > 0$, $x \in \bar{\Omega}$, это свойство сохраняет на указанном множестве и оператор A , определенный равенством $Au = -\nabla(\lambda \nabla u)$.

Пример 5.10. Изложенный выше подход можно применить к исследованию *оператора Штурма — Лиувилля* A , который на множестве $D(A)$ дважды непрерывно дифференцируемых на отрезке $[0, 1]$ функций $u(x)$, обращающихся на концах этого отрезка в нуль (множество $D(A)$ — линейное многообразие в $L_2[0, 1]$), определен равенством

$$Au = -(p(x)u'(x))' + q(x)u(x), \quad x \in [0, 1]. \quad (5.91)$$

Используя интегрирование по частям и учитывая, что $u(0) = v(0) = u(1) = v(1) = 0$ для функций $u, v \in D(A)$, находим

$$\begin{aligned} \langle Au, v \rangle &= \int_0^1 (-(p(x) u'(x))' + q(x) u(x)) v(x) dx = \\ &= -p(x) u'(x) v(x) \Big|_0^1 + \int_0^1 p(x) u'(x) v'(x) dx + \int_0^1 q(x) u(x) v(x) dx = \\ &= p(x) v'(x) u(x) \Big|_0^1 - \int_0^1 (p(x) v'(x))' u(x) dx + \\ &\quad + \int_0^1 q(x) u(x) v(x) dx = \langle Av, u \rangle. \end{aligned}$$

Следовательно, оператор A является симметрическим. Далее вычислим

$$\begin{aligned} \langle Au, u \rangle &= \int_0^1 (-(p(x) u'(x))' + q(x) u(x)) u(x) dx = \\ &= -p(x) u'(x) u(x) \Big|_0^1 + \int_0^1 p(x) (u'(x))^2 dx + \int_0^1 q(x) u^2(x) dx = \\ &= \int_0^1 p(x) (u'(x))^2 dx + \int_0^1 q(x) u^2(x) dx. \quad (5.92) \end{aligned}$$

Примем, что $0 < p_0 \leq p(x)$ и $0 \leq q(x) \leq \bar{q}$ при любых $x \in [0, 1]$. Тогда вместо (5.92) получим

$$\langle Au, u \rangle \geq p_0 \int_0^1 (u'(x))^2 dx + \int_0^1 q(x) u^2(x) dx \geq 0. \quad (5.93)$$

При $\langle Au, u \rangle = 0$ каждый из интегралов в правой части (5.93) обращается в нуль. Из равенства нулю первого из них следует,

что $u(x) = C_0 = \text{const}$, $x \in [0, 1]$, а из условия $u(0) = u(1) = 0$ — что $C_0 = 0$. Таким образом, оператор A является положительным. Ясно, что при $q(x) \geq q_0 > 0$, $x \in (0, 1)$, оператор A будет положительно определенным, причем $\langle Au, u \rangle \geq q_0 \|u\|^2$.

Покажем, что оператор A может остаться положительно определенным при выполнении условий $p(x) \geq p_0 > 0$, $q(x) \geq q_0$, $x \in (0, 1)$, и в случае $q_0 < 0$, если значение $|q_0|$ не слишком велико. Положим $u(x) = v(x)w(x)$ и, опуская обозначение аргумента x , вычислим

$$(u')^2 = v^2(w')^2 + 2vv'ww' + (v')^2w^2 = v^2(w')^2 + (vv'w^2)' - vv''w^2.$$

Отбрасывая первое (неотрицательное) слагаемое в правой части этого равенства, получаем $(vv'w^2)' - vv''w^2 \leq (u')^2$ и, интегрируя по отрезку $[0, 1]$, находим

$$v(x)v'(x)w^2(x)\Big|_0^1 - \int_0^1 v(x)v''(x)w^2(x)dx \leq \int_0^1 (u'(x))^2 dx. \quad (5.94)$$

Выберем $v(x) = \sin \pi x$. Тогда $v''(x) = -\pi^2 \sin \pi x$ и

$$v(x)v''(x)w^2(x) = -\pi^2 v^2(x)w^2(x) = -\pi^2 u^2(x),$$

а вместо (5.94) будем иметь

$$\pi^2 \int_0^1 u^2(x) dx \leq \int_0^1 (u'(x))^2 dx. \quad (5.95)$$

Отметим, что увеличение множителя перед интегралом в левой части (5.95) невозможно, поскольку при $w(x) \equiv 1$, т.е. при $u(x) = v(x) = \sin \pi x$ (5.95) переходит в равенство.

Из (5.93) с учетом (5.95) следует

$$\langle Au, u \rangle \geq \pi^2 p_0 \int_0^1 u^2(x) dx + \int_0^1 q(x) u^2(x) dx \geq (\pi^2 p_0 + q_0) \int_0^1 u^2(x) dx.$$

Таким образом, при условии $\pi^2 p_0 + q_0 > 0$

$$\langle Au, u \rangle \geq (\pi^2 p_0 + q_0) \|u\|^2 \geq 0, \quad (5.96)$$

и оператор A является не только положительным, но и положительно определенным. Если в (5.91) $q(x) \equiv q_0 = 0$ и $p(x) \equiv p_0 = 1$, то имеем положительно определенный оператор $A = -\frac{d^2}{dx^2}$ с областью определения $D(A)$.

Несложно проверить, что рассмотренные в этом примере операторы сохраняют свойство положительной определенности на множестве $D(A)$ функций $u(x)$, удовлетворяющих не только условиям $u(0) = u(1) = 0$, но и условиям $u(0) = u'(1) = 0$ или $u'(0) = u(1) = 0$.

Вопросы и задачи

5.1. Доказать, что оператор дифференцирования $D = \frac{d}{dx}$, действующий из нормированного пространства $\mathcal{U}$ непрерывно дифференцируемых на отрезке $[a, b]$ функций u в нормированное пространство $\mathcal{V}$ непрерывных на $[a, b]$ функций v с нормами

$$\|u\|_{\mathcal{U}} = \sqrt{\int_a^b u^2(x) dx}, \quad u \in \mathcal{U}, \quad \text{и} \quad \|v\|_{\mathcal{V}} = \sqrt{\int_a^b v^2(x) dx}, \quad v \in \mathcal{V},$$

соответственно, не является ограниченным.

5.2. Доказать, что оператор $A: L_2[0, 1] \rightarrow L_2[0, 1]$, действующий по правилу

$$A\varphi = f \iff f(t) = \int_0^1 K(t, s)\varphi(s) ds, \quad t \in [0, 1],$$

где $K(t, s)$ — функция, непрерывная в квадрате $[0, 1]^2$, является симметрическим тогда и только тогда, когда $K(t, s) = K(s, t)$,

$t, s \in [0, 1]$. Выяснить, будет ли оператор A симметрическим, если:

- а) $K(t, s) = t^2 + s^2 + 1, \quad t, s \in [0, 1]$;
 б) $K(t, s) = t^3 + ts^2 + s^3 - s, \quad t, s \in [0, 1]$.

5.3. Найти собственные значения и собственные элементы оператора $A = \frac{d^2}{dx^2}$, определенного на множестве функций $u(x)$, непрерывно дифференцируемых на отрезке $[0, 1]$ и принимающих значения: а) $u(0) = u'(1) = 0$; б) $u'(0) = u(1) = 0$; в) $u'(0) = u'(1) = 0$; г) $u(0) = u(1) = 0, u'(0) = u'(1)$.

5.4. Доказать, что спектр вполне непрерывного симметрического оператора A целиком лежит на отрезке $[m, M]$, где

$$m = \inf_{\|u\|=1} \langle Au, u \rangle, \quad M = \sup_{\|u\|=1} \langle Au, u \rangle,$$

причем m, M — точки спектра.

5.5. Найти решение интегрального уравнения II рода

$$y(x) = \lambda \int_0^{\pi} y(s) \sin(x+s) ds + \sin x - \cos x.$$

5.6. Показать, что оператор A , заданный равенством (5.91), является положительно определенным на множестве дважды непрерывно дифференцируемых на отрезке $[0, 1]$ функций $u(x)$, удовлетворяющих условиям $u(0) = u'(1) = 0$.

5.7. Используя (5.95), показать, что оператор A (5.20), рассмотренный в примере 5.3, является положительно определенным на множестве $D(A)$ четырежды непрерывно дифференцируемых в интервале $(0, 1)$ функций, удовлетворяющих (5.19).

5.8. Каким вариантам граничных условий (помимо варианта (5.19)) должны удовлетворять функции из множества $D(A)$ четырежды непрерывно дифференцируемых в интервале $(0, 1)$ функций, чтобы оператор A (5.20), рассмотренный в примере 5.3, был положительным на $D(A)$?

6. ПРИБЛИЖЕННЫЕ АНАЛИТИЧЕСКИЕ МЕТОДЫ

В связи с развитием и совершенствованием вычислительной техники возросла роль численных методов приближенного решения задач математической физики. Но при этом не утратили своего значения приближенные аналитические методы, позволяющие получить в конечном виде соотношения между искомыми функциями и заданными параметрами рассматриваемой задачи.

Существует много подходов к построению приближенных аналитических методов решения задач математической физики. В этой главе рассмотрены общая схема такого построения и оценки возникающих при этом погрешностей, а также *проекционные методы* и методы, связанные с представлением искомого решения в виде разложения по малому параметру.

6.1. Общая схема построения приближенных методов

Рассмотренные в этой части свойства *функциональных пространств* и действующих в этих пространствах операторов позволяют наметить достаточно общую схему построения приближенных методов решения *операторного уравнения*

$$A(u) = f, \quad u \in D(A) \subset \mathcal{H}, \quad f \in \mathcal{H}, \quad (6.1)$$

к которому можно свести формулировку большинства задач математической физики. Для определенности будем предполагать, что существует единственное решение u^0 . Это *решение* может быть либо *классическим*, т.е. принадлежать области

$D(A)$ оператора A (необязательно линейного) и поэтому удовлетворять (6.1) непосредственно, либо *обобщенным* (в частности, *слабым*), т.е. $u^0 \notin D(A)$ (см. замечание 5.2).

Существо приближенного метода нахождения $u^0 \in D(A)$ обычно состоит в построении для (6.1) последовательности $\{u_n\}$ приближенных решений $u_n \in D(A)$, $n \in \mathbb{N}$, сходящейся по норме к u^0 . При практической реализации такого подхода приходится ограничиваться конечной последовательностью $\{u_n\}_N$, но тем не менее важно знать, что последовательность $\{u_n\}$ сходится к элементу u^0 . Кроме того, желательно получить оценки для величины $\alpha_n = \|u_n - u^0\|_{D(A)}$ (например, в виде неравенства $\alpha_n \leq C\beta_n$, где C — некоторая константа, а $\beta_n > 0$ — элементы последовательности, сходящейся к нулю). Если при нахождении u_n используют элемент u_{n-1} (при $n = 1$ этот элемент будет нулевым приближением), то говорят о *методе итераций*, или *методе последовательных приближений*. Наличие оценок может иногда при заданной погрешности приближенного решения u_N указать необходимое число итераций N .

Чтобы найти u_n , заменим (6.1) приближенным уравнением

$$A_n(x_n) = y_n, \quad x_n \in D(A_n), \quad y_n \in R(A_n), \quad (6.2)$$

где оператор A_n в некотором смысле аппроксимирует оператор A , а y_n — элемент f (во всяком случае для упрощения поиска u_n желательно, чтобы A_n был линейным оператором, у которого существует *обратный оператор* A_n^{-1}). При этом пространства $D(A_n)$ и $R(A_n)$ могут быть и конечномерными. В этом случае существование у линейного оператора A_n обратного означает, что матрица оператора A_n квадратная, а размерности $D(A_n)$ и $R(A_n)$ одинаковы.

Введем операторы X_n и Y_n , осуществляющие отображения $X_n: D(A) \rightarrow D(A_n)$ и $Y_n: R(A) \rightarrow R(A_n)$, причем $y_n = Y_n(f)$. Эти операторы могут, например, отображать бесконечномерное функциональное пространство в конечномерное или формировать либо *сетку конечных элементов*, либо *конечно-разностную сетку*. Пусть уравнение (6.2) имеет единственное ре-

шение $x_n^\circ \in D(A_n)$. Вообще говоря, $x_n^\circ \notin D(A)$, и поэтому x_n° нельзя отождествлять с $u_n \in D(A)$ и считать приближенным решением (6.1). Но при помощи оператора U_n можно перейти к этому решению: $u_n = U_n(x_n^\circ)$. В частности, U_n может быть оператором *интерполирования*, который по дискретному представлению элемента $x_n^\circ \in D(A_n)$ восстанавливает его образ $u_n \in D(A)$. На рис. 6.1 условно показана связь рассматриваемых пространств и их элементов.

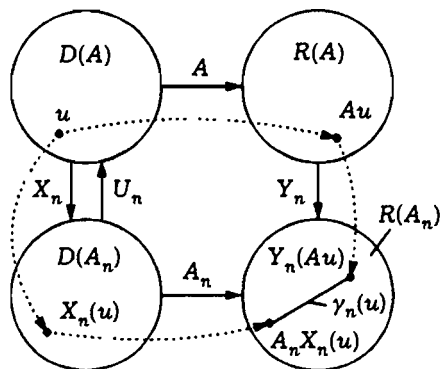


Рис. 6.1

Будем говорить, что последовательность $\{x_n^\circ\}$ сходится, если выполнено условие $\|x_n^\circ - X_n(u^\circ)\|_{D(A_n)} \rightarrow 0$ при $n \rightarrow \infty$, причем элементы $X_n(u^\circ)$ в каком-то смысле приближают элемент u° . Полагая далее A_n линейным оператором, имеющим обратный оператор A_n^{-1} , примем, что A_n аппроксимирует оператор A на элементе $u \in D(A)$, если

$$\gamma_n(u) = \|A_n X_n(u) - Y_n(Au)\|_{R(A_n)} \rightarrow 0 \text{ при } n \rightarrow \infty. \quad (6.3)$$

Значение $\gamma_n(u)$ — это расстояние в $R(A_n)$ между элементами $A_n X_n(u)$ и $Y_n(Au)$ (см. рис. 6.1).

Теорема 6.1. Если

$$v_n = \|A_n^{-1}\| \gamma_n(u^\circ) \rightarrow 0 \text{ при } n \rightarrow \infty, \quad (6.4)$$

то последовательность $\{x_n^\circ\}$ сходится, причем

$$\|x_n^\circ - X_n(u^\circ)\|_{D(A_n)} \leq v_n, \quad n \in \mathbb{N}. \quad (6.5)$$

◀ В силу равенств $A_n x_n^\circ = Y_n(f) = Y_n(Au^\circ)$ с учетом (6.3) имеем

$$\begin{aligned} \|x_n^\circ - X_n(u^\circ)\|_{D(A_n)} &= \|A_n^{-1} A_n (x_n^\circ - X_n(u^\circ))\|_{D(A_n)} \leq \\ &\leq \|A_n^{-1}\| \|A_n x_n^\circ - A_n (X_n(u^\circ))\| = \\ &= \|A_n^{-1}\| \|Y_n(Au^\circ) - A_n (X_n(u^\circ))\| = \|A_n^{-1}\| \gamma_n(u^\circ) = v_n, \end{aligned}$$

что доказывает теорему. ▶

В частности, если операторы A_n , $n \in \mathbb{N}$, аппроксимируют оператор A на элементе u° и операторы A_n^{-1} ограничены, то последовательность $\{x_n^\circ\}$ сходится. Но, согласно теореме 6.1, если выполнено условие (6.4), то последовательность может сходиться и в случае, когда либо отсутствует аппроксимация, либо оператор A_n^{-1} не является ограниченным.

Теорема 6.2. Если $v_n \rightarrow 0$ и $U_n(X_n(u^\circ)) \rightarrow u^\circ$ при $n \rightarrow \infty$, а оператор U_n является линейным и ограниченным, то последовательность $\{u_n\}$ элементов $u_n = U_n(x_n^\circ)$ сходится к искомому элементу u° , причем

$$\|u_n - u^\circ\|_{D(A)} \leq \|U_n\| v_n + \|U_n X_n(u^\circ) - u^\circ\|_{D(A)}, \quad n \in \mathbb{N}. \quad (6.6)$$

◀ Действительно, учитывая равенство $u_n = U_n(x_n^\circ)$ и неравенство треугольника, находим

$$\begin{aligned} \|u_n - u^\circ\|_{D(A)} &= \|U_n(x_n^\circ) - u^\circ\|_{D(A)} = \\ &= \|(U_n(x_n^\circ) - U_n X_n(u^\circ)) + (U_n X_n(u^\circ) - u^\circ)\|_{D(A)} \leq \\ &\leq \|U_n(x_n^\circ) - U_n X_n(u^\circ)\|_{D(A)} + \|U_n X_n(u^\circ) - u^\circ\|_{D(A)} \leq \\ &\leq \|U_n\| \|x_n^\circ - X_n(u^\circ)\|_{D(A_n)} + \|U_n X_n(u^\circ) - u^\circ\|_{D(A)}. \end{aligned}$$

Отсюда получаем (6.6), так как в случае $v_n \rightarrow 0$ при $n \rightarrow \infty$ справедливо утверждение теоремы 6.1 в виде (6.5). ►

Пусть оператор U_n , не являясь линейным, может быть представлен в виде $U_n(x) = V_n x + B(f)$, где $V_n: D(A_n) \rightarrow D(A)$ — линейный ограниченный оператор и $B: R(A) \rightarrow D(A)$ — некоторый оператор. Тогда нетрудно проверить, что теорема 6.2 сохраняет силу после замены в (6.6) $\|U_n\|$ на $\|V_n\|$. Практическое применение теорем 6.1 и 6.2 обычно затруднено, так как элемент u° неизвестен, но они оказываются полезными при построении приближенных методов.

Пример 6.1. В примере 4.17 установлено, что *интегральное уравнение II рода*

$$u(\xi) - \int_a^b K(\xi, t) u(t) dt = f(\xi), \quad (6.7)$$

где $\xi \in [a, b]$, $u(\xi)$, $f(\xi) \in C[a, b]$, $K(\xi, t)$ — *ядро* этого уравнения, непрерывное в квадрате $[a, b]^2$, имеет единственное решение $u^\circ(\xi)$ при задании действительной функции $f(\xi)$ и выполнении условия (4.43)

$$q = (b - a) \max_{\xi, t \in [a, b]} |K(\xi, t)| < 1. \quad (6.8)$$

Один из методов приближенного решения (6.7) состоит в разбиении отрезка $[a, b]$ на $n - 1$ частей точками ξ_i , $i = \overline{1, n}$, и приближенной замене определенного интеграла *квадратурной суммой*. Тогда вместо (6.7) получим систему линейных алгебраических уравнений (СЛАУ)

$$x_i - \sum_{j=1}^n D_j K(\xi_i, \xi_j) x_j = y_i, \quad i, j = \overline{1, n}, \quad y_i = f(\xi_i), \quad (6.9)$$

относительно значений $x_i = u(\xi_i)$ искомой функции $u(\xi)$ в узлах ξ_i , соответствующей *квадратурной формулы* (здесь D_j — *весовые коэффициенты* этой формулы).

В данном случае линейные операторы A и A_n определены левыми частями (6.7) и (6.9) соответственно. Из (6.7) следует, что $D(A) = R(A) = C[a, b]$. Операторы X_n и Y_n совпадают, т.е. $X_n = Y_n$, и отображают бесконечномерное множество $C[a, b]$ непрерывных на отрезке $[a, b]$ функций на подмножество $D(A_n) = R(A_n)$ n -мерного арифметического пространства $\mathbb{R}^n$ значений $x_i = X_n(u_i)$ и $y_i = Y_n(f_i)$ в узлах ξ_i квадратурной формулы, где $u_i = u(\xi_i)$, $f_i = f(\xi_i)$, $i = \overline{1, n}$.

Запишем СЛАУ (6.9) в виде

$$A_n x_n = y_n, \quad A_n = I - B_n,$$

где x_n и y_n — векторы с координатами x_i и y_i , $i = \overline{1, n}$, I — тождественный оператор, а B_n — оператор, матрица (B_{ij}) которого имеет элементы $B_{ij} = D_j K(\xi_i, \xi_j)$, $i, j = \overline{1, n}$. Норму этой матрицы, соответствующую кубической норме векторов x_n и y_n , найдем как наибольшую сумму абсолютных значений элементов B_{ij} в строке, т.е.

$$\|(B_{ij})\| = \max_{i=1, n} \sum_{j=1}^n |D_j K(\xi_i, \xi_j)| \leq \sum_{j=1}^n |D_j| \max_{i, j=1, n} |K(\xi_i, \xi_j)|.$$

Для большинства применяемых квадратурных формул весовые коэффициенты положительны, и их сумма равна $b - a$ [VI]. Поэтому с учетом (6.8) получим следующую оценку для нормы оператора B_n :

$$\|B_n\| = \|(B_{ij})\| \leq (b - a) \max_{\xi, t \in [a, b]} |K(\xi, t)| = q < 1.$$

Согласно теореме 4.18, заключаем, что оператор $A_n = I - B_n$ имеет обратный оператор A_n^{-1} , причем $\|A_n^{-1}\| \leq \frac{1}{1-q}$. Это означает, что существует вектор x_n^0 с координатами x_i^0 , $i = \overline{1, n}$, который является единственным решением СЛАУ (6.9).

Попробуем теперь применить теорему 6.1 для проверки сходимости последовательности $\{x_n^0\}$. Предварительно при

помощи (6.3) найдем выражение для $\gamma_n(u)$, для чего с учетом (6.7) и (6.9) запишем при $i = \overline{1, n}$

$$(Y_n(Au))_i = X_n(Au_i) = x_i - \int_a^b K(\xi_i, t) u(t) dt,$$

$$(A_n X_n(u))_i = x_i - \sum_{j=1}^n D_j K(\xi_i, \xi_j) x_j.$$

Следовательно,

$$(A_n X_n(u))_i - (Y_n(Au))_i = \int_a^b K(\xi_i, t) u(t) dt - \sum_{j=1}^n D_j K(\xi_i, \xi_j) u(\xi_j).$$

Отсюда в соответствии с (6.3) получаем

$$\begin{aligned} \gamma_n(u) = \max_{i=\overline{1, n}} \left| \int_a^b K(\xi_i, t) u(t) dt - \sum_{j=1}^n D_j K(\xi_i, \xi_j) u(\xi_j) \right| = \\ = \max_{i=\overline{1, n}} |R_i|, \quad (6.10) \end{aligned}$$

где R_i — погрешность используемой квадратурной формулы при фиксированном аргументе ξ_i ядра $K(\xi_i, t)$ в (6.10).

При равномерном разбиении отрезка $[a, b]$ на четное число $n - 1$ ($n \geq 3$) частичных отрезков разбиения длиной $h = \frac{b-a}{n-1}$ можно использовать квадратурную формулу парабол, для которой [VI]

$$\max_{i=\overline{1, n}} |R_i| \leq \frac{(b-a)^5}{180(n-1)^4} M_4, \quad (6.11)$$

где M_4 — наибольшее значение на $[a, b]$ абсолютной величины четвертой производной по t подынтегральной функции $K(\xi_i, t) u(t)$ в (6.10). В данном случае применение теоремы 6.1 возможно, если располагать оценкой для $\gamma_n(u^0)$. Покажем, что

эту оценку можно получить, если функции $K(\xi, t)$ и $f(\xi)$ четырежды непрерывно дифференцируемы по своим аргументам, т.е. покажем, как в таком случае можно найти значение M_4 в (6.11). Тогда интеграл в (6.7) можно рассматривать как зависящий от параметра ξ и дифференцировать его по этому параметру [VI].

Поскольку функция $u^\circ(\xi)$ удовлетворяет (6.7), т.е.

$$u^\circ(\xi) = \int_a^b K(\xi, t) u^\circ(t) dt + f(\xi), \quad (6.12)$$

дифференцированием этого выражения находим

$$\frac{d^m u^\circ(\xi)}{d\xi^m} = \int_a^b \frac{\partial^m K(\xi, t)}{\partial \xi^m} u^\circ(t) dt + \frac{d^m f(\xi)}{d\xi^m}, \quad m = \overline{1, 4}, \quad (6.13)$$

и убеждаемся, что функция $u^\circ(\xi)$ также четырежды непрерывно дифференцируема. Учитывая, что (см. пример 4.17)

$$\|u^\circ\|_{C[a,b]} = \max_{\xi \in [a,b]} |u^\circ(\xi)| \leq \frac{\|f\|_{C[a,b]}}{1-q} = \frac{1}{1-q} \max_{\xi \in [a,b]} |f(\xi)|,$$

и располагая оценками для наибольших абсолютных значений производных функций $K(\xi, t)$ и $f(\xi)$ по ξ , можно найти такие оценки и для $\frac{d^m u^\circ(\xi)}{d\xi^m}$ в (6.13), а затем получить оценку M_4 для производной

$$\frac{\partial^4 K(\xi, t) u^\circ(t)}{\partial t^4} = \sum_{m=1}^4 C_4^m \frac{\partial^m K(\xi, t)}{\partial t^m} \frac{\partial^{4-m} u^\circ(t)}{\partial t^{4-m}}$$

подынтегральной функции в (6.10).

В итоге из (6.10) и (6.11) получаем

$$\gamma_n(u^\circ) \leq \frac{(b-a)^5}{180(n-1)^4} M_4,$$

т.е. в соответствии с (6.3) оператор A_n аппроксимирует оператор A на элементе u° . Поскольку

$$v_n = \|A_n^{-1}\| \gamma_n(u^\circ) \leq \frac{M_4}{1-q} \frac{(b-a)^5}{180(n-1)^4},$$

то $v_n \rightarrow 0$ при $n \rightarrow \infty$, а значит, согласно теореме 6.1, последовательность $\{x_n^\circ\}$ сходится, причем из (6.5) следует, что координаты $(x_n^\circ)_i$ вектора x_n° при $n \rightarrow \infty$ стремятся к значениям $u^\circ(\xi_i)$ искомой функции $u^\circ(\xi)$ в соответствующих узлах формулы парабол.

В простейшем случае можно принять, что приближенное решение (6.7) является непрерывной действительной функцией $u_n(\xi) = U_n x_n^\circ$, совпадающей в узлах ξ_i со значениями $(x_n^\circ)_i$ и линейной между соседними узлами, т.е. воспользоваться *линейной интерполяцией*. Тогда для нормы определенного таким образом оператора U_n имеем $\|U_n\| = 1$, причем функция $U_n X_n(u^\circ)$ совпадает в узлах ξ_i с искомым решением $u^\circ(\xi)$, а между соседними узлами линейна. Поэтому в случае линейной интерполяции [II]

$$\|U_n X_n(u^\circ) - u^\circ\|_{C[a,b]} \leq M_2 \frac{h^2}{8} = M_2 \frac{(b-a)^2}{8(n-1)^2},$$

где

$$M_2 = \max_{\xi \in [a,b]} \left| \frac{d^2 u^\circ(\xi)}{d\xi^2} \right|.$$

Итак, и $v_n \rightarrow 0$, и $U_n X_n(u^\circ) \rightarrow u^\circ$ при $n \rightarrow \infty$. Следовательно, согласно теореме 6.2, последовательность $\{u_n(\xi)\}$ приближенных решений (6.7) сходится к искомой функции $u^\circ(\xi)$, причем в соответствии с (6.6)

$$\begin{aligned} \|u_n(\xi) - u^\circ(\xi)\|_{C[a,b]} &= \max_{\xi \in [a,b]} |u_n(\xi) - u^\circ(\xi)| \leq \\ &\leq \frac{M_4}{1-q} \frac{(b-a)^5}{180(n-1)^4} + M_2 \frac{(b-a)^2}{8(n-1)^2}. \end{aligned}$$

Более быструю сходимость последовательности $\{u_n(\xi)\}$ к искомому решению можно обеспечить, если оператор U_n задать соотношением

$$u_n(\xi) = \sum_{i=1}^n D_i K(\xi, \xi_i) (x_n^\circ)_i + f(\xi), \quad (6.14)$$

т.е. $u_n = U_n(x_n^\circ) = V_n x_n^\circ + I f$, где V_n — линейный ограниченный оператор, заданный выражением

$$V_n x_n^\circ = \sum_{i=1}^n D_i K(\xi, \xi_i) (x_n^\circ)_i.$$

В этом случае, учитывая (6.8), имеем

$$\|V x_n^\circ\| \leq \sum_{i=1}^n |D_i K(\xi, \xi_i)| |(x_n^\circ)_i| \leq q \|x_n^\circ\|,$$

где $\|x_n^\circ\| = \max_{i=1, n} |(x_n^\circ)_i|$ — кубическая норма вектора x_n° . Отсюда в соответствии с (4.26) следует, что $\|V_n\| \leq q < 1$. Теперь

$$U_n(X_n u^\circ) = \sum_{i=1}^n D_i K(\xi, \xi_i) u^\circ(\xi_i) + f(\xi) = \bar{u}_n(\xi),$$

так что, принимая во внимание (6.10), (6.11) и (6.12), получаем

$$\begin{aligned} \|U_n(X_n u^\circ) - u^\circ\|_{C[a,b]} &= \max_{\xi \in [a,b]} |\bar{u}_n(\xi) - u^\circ(\xi)| = \\ &= \max_{\xi \in [a,b]} \left| \sum_{i=1}^n D_i K(\xi, \xi_i) u^\circ(\xi_i) - \int_a^b K(\xi, t) u^\circ(t) dt \right| \leq M_4 \frac{(b-a)^5}{180(n-1)^4}. \end{aligned}$$

Таким образом, и в этом случае справедливо утверждение теоремы 6.2 о сходимости последовательности $\{u_n^\circ(\xi)\}$ к искомому

решению $u^0(\xi)$, но теперь в (6.6) $\|U_n\|$ следует заменить на $\|V_n\|$, что приводит к неравенству

$$\begin{aligned} \|u_n(\xi) - u^0(\xi)\|_{C[a,b]} &\leq \|V_n\| v_n + \|U_n(X_n u^0) - u^0\|_{C[a,b]} \leq \\ &\leq \frac{M_4 q}{1-q} \frac{(b-a)^5}{180(n-1)^4} + M_4 \frac{(b-a)^5}{180(n-1)^4} = \frac{M_4}{1-q} \frac{(b-a)^5}{180(n-1)^4}. \end{aligned}$$

Таким образом, при линейной интерполяции скорость сходимости последовательности $\{u_n^0(\xi)\}$ была пропорциональна $\frac{1}{(n-1)^2}$, а при использовании (6.14) стала пропорциональной $\frac{1}{(n-1)^4}$.

6.2. Погрешности приближенных методов

При поиске приближенного решения *операторного уравнения* вида (6.1) неизбежно возникновение погрешностей. Их источники связаны со следующими основными причинами: неточностями в задании оператора и правой части (6.1) при постановке задачи, погрешностями применяемого метода решения (6.1) и вычислительными погрешностями.

Рассмотрим сначала первую причину возникновения погрешностей, записав вместо (6.1) уравнение

$$(A + \Delta A)u = f + \Delta f, \quad u \in D(A) \subset \mathcal{H}, \quad f \in R(A) \subset \mathcal{H}, \quad (6.15)$$

где ΔA и $\Delta f \in R(\Delta A)$ — искажения в операторе и правой части (6.1) соответственно, причем $D(\Delta A) \supset D(A)$ и $R(\Delta A) \subset R(A)$. Предположим, что операторы A и ΔA являются линейными ограниченными и существует обратный оператор A^{-1} , а искажения в операторе A достаточно малы, т.е. $\|A^{-1}\Delta A\| \leq q < 1$. Тогда, согласно теореме 4.18, оператор $A + \Delta A$ имеет обратный оператор $S = (A + \Delta A)^{-1} = A^{-1}(I + A^{-1}\Delta A)^{-1}$, причем с учетом (4.35) и (4.37) $\|S\| \leq \frac{\|A^{-1}\|}{1-q}$. Это означает, что суще-

ствуется единственное решение u^* , удовлетворяющее (6.15):

$$(A + \Delta A)u^* = f + \Delta f. \quad (6.16)$$

Из (6.16) вычтем равенство $Au^0 = f$, где u^0 — решение (4.37) и запишем

$$A(u^* - u^0) = -\Delta Au^* + \Delta f.$$

Отсюда, учитывая неравенство треугольника и (4.26), получаем

$$\begin{aligned} \|u^* - u^0\| &\leq \|A^{-1}\Delta A\| \|u^*\| + \|A^{-1}\| \|\Delta f\| \leq \\ &\leq q \|u^*\| + \|A^{-1}\| \|\Delta f\|. \end{aligned} \quad (6.17)$$

Такую **оценку погрешности** $\|u^* - u^0\|$ называют **апостериорной** (по латыни а posteriori — из последующего), поскольку ее можно получить лишь после решения (6.15), уже располагая элементом u^* .

Из (6.15) с учетом оценки **нормы оператора** $S = (A + \Delta A)^{-1}$ имеем

$$u^* = S(f + \Delta f) \quad \text{и} \quad \|u^*\| \leq \frac{\|A^{-1}\|}{1 - q} \|f + \Delta f\|.$$

Тогда, подставляя $\|u^*\|$ в (6.17), находим **априорную оценку погрешности** (по латыни а priori — из предыдущего)

$$\|u^* - u^0\| \leq \frac{\|A^{-1}\|}{1 - q} (q \|f\| + \|\Delta f\|). \quad (6.18)$$

Для практического применения оценок (6.17) и (6.18) необходимы значения $\|A^{-1}\|$ и q , нахождение которых часто является непростой самостоятельной задачей. Из (6.16) и равенства $Au^0 = f$ получим

$$(A + \Delta A)(u^* - u^0) = -\Delta Au^0 + \Delta f.$$

Отсюда с учетом неравенства треугольника, (4.26) и оценки для нормы оператора S находим

$$\begin{aligned}\|u^* - u^0\| &\leq \|(A + \Delta A)^{-1} \Delta A\| \|u^0\| + \|(A + \Delta A)^{-1}\| \|\Delta f\| \leq \\ &\leq \frac{\|A^{-1}\|}{1 - q} (\|\Delta A\| \|u^0\| + \|\Delta f\|). \quad (6.19)\end{aligned}$$

Обозначим $\mu(A) = \|A\| \|A^{-1}\|$ (в конечномерном случае $\mu(A)$ называют *числом обусловленности матрицы* оператора A). Тогда, если $\|A^{-1}\| \|\Delta A\| < 1$, приняв $q = \|A^{-1}\| \|\Delta A\|$, из (6.19) получим

$$\|u^* - u^0\| \leq \mu(A) \frac{\frac{\|\Delta A\|}{\|A\|} \|u^0\| + \frac{\|\Delta f\|}{\|A\|}}{1 - \mu(A) \frac{\|\Delta A\|}{\|A\|}}.$$

Поскольку $\|f\| \leq \|A\| \cdot \|u^0\|$, приходим к априорной оценке для относительной погрешности

$$\frac{\|u^* - u^0\|}{\|u^0\|} \leq \frac{\mu(A)}{1 - \mu(A) \frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta f\|}{\|f\|} \right). \quad (6.20)$$

Отметим, что оценка (6.20) зависит от относительных погрешностей в операторе и правой части (6.15). Ясно, что при выполнении условия $\mu(A) \frac{\|\Delta A\|}{\|A\|} = q < 1$ и больших значениях $\mu(A)$ правая часть (6.20) может существенно превысить сумму $\frac{\|\Delta A\|}{\|A\|}$ и $\frac{\|\Delta f\|}{\|f\|}$.

Прежде чем рассматривать остальные причины возникновения погрешностей, предположим, что мы располагаем приближенным решением u_n уравнения (6.1), полученным в действительности как итог приближенного решения уравнения (6.15). Пусть при подстановке u_n в (6.15) возникает *невязка* $\delta = (A + \Delta A)u_n - f + \Delta f$, т.е. u_n удовлетворяет не (6.15), а уравнению $(A + \Delta A)u_n = f + \Delta f + \delta$. Поэтому, зная невязку δ ,

можно получить апостериорную оценку относительной погрешности приближенного решения u_n , если в (6.20) заменить $\|\Delta f\|$ на $\|\Delta f\| + \|\delta\|$:

$$\frac{\|u_n - u^0\|}{\|u^0\|} \leq \frac{\mu(A)}{1 - \mu(A) \frac{\|\Delta A\|}{\|A\|}} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta f\| + \|\delta\|}{\|f\|} \right). \quad (6.21)$$

Погрешность метода приближенного решения (6.1) путем замены этого уравнения на (6.2) можно оценить при помощи теоремы 6.1. Но в действительности при замене (6.1) на (6.2) возникают дополнительные погрешности, так что вместо (6.2) приходится решать „возмущенное“ уравнение

$$(A_n + \Delta A_n)x_n = y_n + \Delta y_n, \quad \Delta y_n \in R(A_n), \quad (6.22)$$

причем $D(\Delta A_n) \subset D(A_n)$ и $R(\Delta A_n) \supset R(A_n)$. Сравнивая (6.22) с (6.15), приходим к выводу, что по аналогии с (6.20) будем иметь априорную оценку относительной погрешности решения (6.2) в виде

$$\frac{\|x_n - x_n^0\|}{\|x_n^0\|} \leq \frac{\mu(A_n)}{1 - \mu(A_n) \frac{\|\Delta A_n\|}{\|A_n\|}} \left(\frac{\|\Delta A_n\|}{\|A_n\|} + \frac{\|\Delta y_n\|}{\|y_n\|} \right), \quad (6.23)$$

где x_n^0 и x_n – решения (6.2) и (6.22) соответственно, а $\mu(A_n) = \|A_n\| \|A_n^{-1}\|$.

В действительности за счет вычислительных погрешностей при решении (6.22) вместо x_n получим элемент $\tilde{x}_n \in D(A_n)$, подстановка которого в (6.22) приведет к невязке $\tilde{\delta}_n$. Замена в (6.23) $\|\Delta y_n\|$ на $\|\Delta y_n\| + \|\tilde{\delta}_n\|$ позволяет получить апостериорную оценку относительной погрешности $\frac{\|\tilde{x}_n - x_n^0\|}{\|x_n^0\|}$, вызванной совместно второй и третьей причинами.

Наконец, следует оценить погрешность Δu_n , возникающую при переходе от решения x_n^0 уравнения (6.2) к приближенному решению $u_n = U_n(x_n^0)$ уравнения (6.1). Учитывая возможное

искажение ΔU_n оператора U_n и погрешность $\Delta x_n^\circ = x_n - x_n^\circ$ решения (6.2), находим

$$u_n + \Delta u_n = U_n(x_n^\circ) + \Delta u_n = (U_n + \Delta U_n)(x_n^\circ + \Delta x_n^\circ),$$

или, считая операторы U_n и ΔU_n линейными,

$$\Delta u_n = (U_n + \Delta U_n)\Delta x_n^\circ + \Delta U_n x_n^\circ.$$

Отсюда с учетом неравенства треугольника и (4.26), получаем априорную оценку абсолютной погрешности

$$\begin{aligned} \|\Delta u_n\| &= \|(U_n + \Delta U_n)\Delta x_n^\circ + \Delta U_n x_n^\circ\| \leq \\ &\leq \left(1 + \frac{\|\Delta U_n\|}{\|U_n\|}\right) \|U_n\| \|\Delta x_n^\circ\| + \frac{\|\Delta U_n\|}{\|U_n\|} \|U_n\| \|x_n^\circ\| \leq \\ &\leq \left(\left(1 + \frac{\|\Delta U_n\|}{\|U_n\|}\right) \frac{\|\Delta x_n^\circ\|}{\|x_n^\circ\|} + \frac{\|\Delta U_n\|}{\|U_n\|}\right) \|U_n\| \|x_n^\circ\|. \end{aligned}$$

Чтобы учесть еще и вычислительную погрешность, достаточно к $\|\Delta x_n^\circ\|$ добавить норму невязки $\|\delta_n\|$, где

$$\delta_n = \tilde{u}_n - (U_n + \Delta U_n)(x_n^\circ + \Delta x_n^\circ),$$

а $\tilde{u}_n \in D(A)$ — элемент, вычисленный в итоге перехода от решения (6.2) к приближенному решению (6.1). Тогда придем к апостериорной оценке абсолютной погрешности

$$\|\tilde{u}_n - u_n\| \leq \left(\left(1 + \frac{\|\Delta U_n\|}{\|U_n\|}\right) \frac{\|\Delta x_n^\circ\| + \|\delta_n\|}{\|x_n^\circ\|} + \frac{\|\Delta U_n\|}{\|U_n\|}\right) \|U_n\| \|x_n^\circ\|.$$

6.3. Метод малого параметра

Пусть в банаховом пространстве B задано операторное уравнение $Bu = f$, $u, f \in B$, где B — линейный ограниченный оператор из банахова пространства $\mathcal{L}(B)$ линейных ограниченных операторов. Пусть оператор B можно представить в виде

$B = B_0 - \lambda A$, $\lambda \neq 0$, где B_0 имеет известный или достаточно просто получаемый обратный оператор $B_0^{-1} \in \mathcal{L}(B)$. Поскольку в $\mathcal{L}(B)$ определено умножение операторов, запишем

$$Bu = (B_0 - \lambda A)u = B_0(I - \lambda B_0^{-1}A)u = f, \quad f \in B, \quad (6.24)$$

где I — тождественный оператор.

Согласно теореме 4.18, оператор $I - \lambda B_0^{-1}A \in \mathcal{L}(B)$ при $|\lambda| \|B_0^{-1}A\| \leq q < 1$ имеет обратный оператор

$$S = (I - \lambda B_0^{-1}A)^{-1} = \sum_{k=0}^{\infty} \lambda^k (B_0^{-1}A)^k, \quad (6.25)$$

причем $(B_0^{-1}A)^0 = I$. Таким образом, выбирая параметр λ достаточно малым по абсолютному значению, можно обеспечить существование обратного оператора $S \in \mathcal{L}(B)$.

Используя (6.24) и (6.25), из равенств

$$BSB_0^{-1} = B_0(I - \lambda B_0^{-1}A)(I - \lambda B_0^{-1}A)^{-1}B_0^{-1} = I,$$

$$SB_0^{-1}B = (I - \lambda B_0^{-1}A)^{-1}B_0^{-1}B_0(I - \lambda B_0^{-1}A) = I$$

и теоремы 4.1 устанавливаем, что оператор SB_0^{-1} является обратным к B . Поэтому, учитывая (6.25), находим решение уравнения (6.24) в виде

$$u = SB_0^{-1}f = \sum_{k=0}^{\infty} \lambda^k (B_0^{-1}A)^k B_0^{-1}f = \sum_{k=0}^{\infty} \lambda^k \tilde{A}^k \tilde{u}, \quad (6.26)$$

где $\tilde{A} = B_0^{-1}A$ и $\tilde{u} = B_0^{-1}f$. Из условия существования обратного оператора S получаем, что ряд в (6.26) сходится, если $|\lambda| < \frac{1}{\|B_0^{-1}A\|}$.

Из (6.26) получим приближенное решение

$$u_N(\lambda) = \sum_{k=0}^N \lambda^k \tilde{A}^k \tilde{u}. \quad (6.27)$$

Вычитая (6.27) из (6.26) и оценивая полученную разность по норме, найдем оценку погрешности приближенного решения (6.27):

$$\begin{aligned} \|u - u_N\|_B &= \left\| \sum_{k=N+1}^{\infty} \lambda^k \tilde{A}^k \tilde{u} \right\|_B \leq \sum_{k=N+1}^{\infty} |\lambda|^k \|\tilde{A}^k \tilde{u}\|_B \leq \\ &\leq \sum_{k=N+1}^{\infty} |\lambda|^k \|\tilde{A}\|^k \|\tilde{u}\|_B = \frac{(|\lambda| \|\tilde{A}\|)^{N+1}}{1 - |\lambda| \|\tilde{A}\|} \|\tilde{u}\|_B. \end{aligned} \quad (6.28)$$

Пример 6.2. Пусть для интегрального уравнения II рода

$$u(\xi) - \lambda \int_a^b K(\xi, t) u(t) dt = f(\xi), \quad (6.29)$$

где $\lambda \neq 0$, $u, f \in C[a, b]$, ядро $K(\xi, t)$ является непрерывной в квадрате $[a, b]^2$ функцией и выполнено условие

$$q = |\lambda|(b-a) \max_{\xi, t \in [a, b]} |K(\xi, t)| < 1. \quad (6.30)$$

Представим (6.29) в виде операторного уравнения $(B_0 - \lambda A)u = f$, где $B_0 = I$, а A — оператор, который ставит в соответствие функции $u \in C[a, b]$ функцию $v \in C[a, b]$ по правилу

$$v(\xi) = \int_a^b K(\xi, t) u(t) dt.$$

Таким образом, $A: C[a, b] \rightarrow C[a, b]$.

С учетом (6.30) и примера 4.17 имеем

$$\begin{aligned} \|Au\|_{C[a, b]} &= \max_{\xi \in [a, b]} |v(\xi)| = \max_{\xi \in [a, b]} \left| \int_a^b K(\xi, t) u(t) dt \right| \leq \\ &\leq \int_a^b \max_{\xi \in [a, b]} |K(\xi, t) u(t)| dt \leq \frac{q \|u\|_{C[a, b]}}{|\lambda|}. \end{aligned}$$

Отсюда в соответствии с (4.26) следует, что $\|A\| \leq \frac{q}{|\lambda|}$, или $|\lambda| \|B_0^{-1} A\| \leq q < 1$, так как в данном случае $B_0 = I$. Таким образом, полагая в (6.26) $\tilde{A} = A$ и $\tilde{u} = B_0^{-1} f = I f = f$, для искомой функции получаем

$$u = \sum_{k=0}^{\infty} \lambda^k A^k f, \quad (6.31)$$

причем в данном случае $A^0 f = f$, $A f = \varphi_1$, где

$$\varphi_1(\xi) = \int_a^b K(\xi, t) f(t) dt.$$

Представим в явном виде оператор A^k , $k > 1$. Обозначим $K_1(\xi, t) = K(\xi, t)$. Тогда можно записать $A^2 f = \varphi_2$, где

$$\begin{aligned} \varphi_2(\xi) &= \int_a^b K(\xi, \tau) \varphi_1(\tau) d\tau = \int_a^b K(\xi, \tau) d\tau \int_a^b K_1(\tau, t) f(t) dt = \\ &= \int_a^b f(t) dt \int_a^b K(\xi, \tau) K_1(\tau, t) d\tau. \end{aligned}$$

Обозначая

$$K_2(\xi, t) = \int_a^b K(\xi, \tau) K_1(\tau, t) d\tau,$$

приходим к равенству

$$\varphi_2(\xi) = \int_a^b K_2(\xi, t) f(t) dt.$$

Используя метод математической индукции, нетрудно показать, что $A^k f = \varphi_k$, где

$$\varphi_k(\xi) = \int_a^b K_k(\xi, t) f(t) dt,$$

причем для нахождения ядра $K_k(\xi, t)$ имеем рекуррентную формулу

$$K_k(\xi, t) = \int_a^b K(\xi, \tau) K_{k-1}(\tau, t) d\tau, \quad k = 2, 3, \dots$$

Так как сходимость последовательности функций как элементов банахова пространства $C[a, b]$ эквивалентна сходимости этой функциональной последовательности на отрезке $[a, b]$, то (6.31) равносильно равенству

$$u(\xi) = f(\xi) + \sum_{k=1}^{\infty} \lambda^k \varphi_k(\xi), \quad \xi \in [a, b], \quad (6.32)$$

причем ряд в правой части этого равенства сходится к функции $u(\xi)$ равномерно на отрезке $[a, b]$. Рассмотрим некоторые частные случаи.

В уравнении (6.29) положим $a = 0$, $b = 1$ и $K(\xi, t) = e^{\xi-t}$. Так как $\max_{\xi, t \in [0, 1]} |K(\xi, t)| = e$, то условие (6.30) будет выполнено при $|\lambda| < 1/e$. В данном случае $K_1(\xi, t) = e^{\xi-t}$ и

$$K_2(\xi, t) = \int_0^1 K(\xi, \tau) K_1(\tau, t) d\tau = \int_0^1 e^{\xi-\tau} e^{\tau-t} d\tau = e^{\xi-t}.$$

Значит, для любого номера k имеем $K_k(\xi, t) = e^{\xi-t}$ и $A^k = A$. В соответствии с (6.32) и формулой для суммы членов геометри-

ческой прогрессии получаем

$$u(\xi) = f(\xi) + \sum_{k=1}^{\infty} \lambda^k \int_0^1 e^{\xi-t} f(t) dt = f(\xi) + \frac{\lambda e^{\xi}}{1-\lambda} \int_0^1 e^{-t} f(t) dt.$$

Пусть теперь в (6.29) $a = 0$, $b = 1$ и $K(\xi, t) = 2\xi - t$. Тогда условие (6.30) будет выполнено при $|\lambda| < \frac{1}{2}$. Найдём приближённое решение уравнения (6.29) для $f(\xi) \equiv 1$. В (6.27) ограничимся значением $N = 2$ и в соответствии с (6.32) вычислим

$$\varphi_1(\xi) = \int_0^1 (2\xi - t) dt = 2\xi - \frac{1}{2},$$

$$K_2(\xi, t) = \int_0^1 (2\xi - \tau)(2\tau - t) d\tau = 2\xi - 2\xi t - \frac{2}{3} + \frac{t}{2},$$

$$\varphi_2(\xi) = \int_0^1 \left(2\xi - 2\xi t - \frac{2}{3} + \frac{t}{2} \right) dt = 2\xi - \xi - \frac{2}{3} + \frac{1}{4} = \xi - \frac{5}{12}.$$

Тогда приближённое решение примет вид

$$u_2(\xi) = 1 + \left(2\xi - \frac{1}{2} \right) \lambda + \left(\xi - \frac{5}{12} \right) \lambda^2.$$

Нетрудно проверить, что точным решением уравнения (6.29) в рассматриваемом случае является функция

$$u(\xi) = 1 + \lambda \frac{2\xi - 1/2 - \lambda/6}{1 - \lambda/2 + \lambda^2/6} = \frac{1 + 2\xi\lambda - \lambda}{1 - \lambda/2 + \lambda^2/6},$$

которая также является линейной. При $\lambda = \frac{1}{2}$ различие приближённого и точного решений существенно: $u_2(\xi) = \frac{5}{4}\xi + \frac{31}{48} \approx 1,2500\xi + 0,6456$ и $u(\xi) = \frac{24}{17}\xi + \frac{12}{17} \approx 1,4118\xi + 0,7059$, но уже

при $\lambda = \frac{1}{4}$ имеем $u_2(\xi) = \frac{9}{16}\xi + \frac{163}{192} \approx 0,5625\xi + 0,8490$ и $u(\xi) = \frac{48}{83}\xi + \frac{72}{83} \approx 0,5783\xi + 0,8675$, т.е. коэффициенты отличаются не более чем на 2%. #

Рассмотренная сравнительно простая процедура построения приближенного решения носит название **метода малого параметра**. Однако если в операторном уравнении $Bu = f$ оператор B не является линейным ограниченным, то этот метод становится более громоздким. Модификация метода потребует введения ряда дополнительных понятий.

6.4. Общий случай метода малого параметра

Пусть Λ — некоторый промежуток числовой прямой, $\mathcal{U}$ — нормированное пространство, а $u: \Lambda \rightarrow \mathcal{U}$ — функция действительного переменного λ , имеющая область определения Λ и область значений в $\mathcal{U}$. В нормированном пространстве $\mathcal{U}$ с нормой $\|\cdot\|$ элемент $a \in \mathcal{U}$ называют пределом функции $u(\lambda)$ в точке $\lambda_0 \in \Lambda$, если $\|u(\lambda) - a\| \rightarrow 0$ при $\lambda \rightarrow \lambda_0$, и пишут

$$a = \lim_{\lambda \rightarrow \lambda_0} u(\lambda), \quad \text{или} \quad u(\lambda) \rightarrow a \text{ при } \lambda \rightarrow \lambda_0.$$

Функцию $u(\lambda)$ называют непрерывной в точке $\lambda_0 \in \Lambda$, если $\|u(\lambda) - u(\lambda_0)\| \rightarrow 0$ при $\lambda \rightarrow \lambda_0$. Отметим, что функция u является отображением метрических пространств. Поэтому введенные выше понятия предела и непрерывности функции в точке соответствуют понятиям предела и непрерывности отображения метрических пространств в точке [1].

Если существует предел

$$\lim_{\lambda \rightarrow \lambda_0} \frac{u(\lambda) - u(\lambda_0)}{\lambda - \lambda_0} = u'(\lambda_0) \in \mathcal{U},$$

то элемент $u'(\lambda_0)$ называют производной функции $u(\lambda)$ в точке λ_0 . Производные высших порядков определяют (как и для

действительной функции действительного переменного) последовательно. Если $u(\lambda)$ имеет в точке λ_0 производные любого порядка, то говорят, что функция $u(\lambda)$ бесконечно дифференцируема в точке λ_0 .

Пусть $\{u_n(\lambda)\}$ — последовательность функций $u_n: \Lambda \subset \mathbb{R} \rightarrow \mathcal{U}$. Тогда можно рассматривать функциональный ряд

$$\sum_{n=1}^{\infty} u_n(\lambda), \quad (6.33)$$

элементами которого являются функции действительного переменного $\lambda \in \Lambda$ с областью значений в нормированном пространстве $\mathcal{U}$. На такие ряды можно перенести все понятия, введенные для функциональных рядов в случае $\mathcal{U} = \mathbb{R}$ [IX].

Напомним, что при фиксированном $\lambda \in \Lambda$ ряд (6.33) как ряд элементов нормированного пространства $\mathcal{U}$ сходится к некоторому элементу $u(\lambda) \in \mathcal{U}$, если к этому элементу сходится последовательность $\{S_k(\lambda)\}$ частичных сумм

$$S_k(\lambda) = \sum_{n=1}^k u_n(\lambda)$$

этого ряда, т.е.

$$\lim_{k \rightarrow \infty} \left\| \sum_{n=1}^k u_n(\lambda) - u(\lambda) \right\| = 0.$$

Если при некотором $\lambda \in \Lambda$ сходится числовой ряд с элементами $\|u_n(\lambda)\|$, то ряд (6.33) называют абсолютно сходящимся в точке $\lambda \in \Lambda$. Отметим, что если $\mathcal{U}$ — банахово пространство, то абсолютно сходящийся ряд сходится к некоторому элементу $u(\lambda) \in \mathcal{U}$ [IX].

Множество всех точек $\lambda \in \Lambda$, для которых ряд (6.33) сходится, называют областью сходимости этого ряда, а множество всех λ , для которых он сходится абсолютно — областью

абсолютной сходимости. Говорят, что ряд (6.33) сходится к функции $u(\lambda)$ равномерно на множестве $\Lambda \subset \mathbb{R}$, если для любого $\epsilon > 0$ найдется номер N , зависящий только от ϵ , такой, что для всех $\lambda \in \Lambda$ и $k > N$ справедливо неравенство $\|S_k - u(\lambda)\| < \epsilon$. Для рядов вида (6.33) в банаховом пространстве $\mathcal{U}$ справедлива теорема, аналогичная признаку Вейерштрасса равномерной сходимости функционального ряда [IX]: если для всех $\lambda \in \Lambda$ справедливы неравенства $\|u_n(\lambda)\| \leq \alpha_n$, $n \in \mathbb{N}$, и $\sum_{n=1}^{\infty} \alpha_n < \infty$, то ряд (6.33) равномерно и абсолютно сходится на множестве Λ . В дальнейшем эту теорему также будем называть признаком Вейерштрасса равномерной сходимости ряда.

Ряд

$$\sum_{k=0}^{\infty} u_k(\lambda - \lambda_0)^k, \quad u_k \in \mathcal{U}, \quad \lambda, \lambda_0 \in \mathbb{R}, \quad (6.34)$$

элементов нормированного пространства $\mathcal{U}$ назовем степенным рядом с коэффициентами из нормированного пространства $\mathcal{U}$.

Определение 6.1. Функцию $u(\lambda) \in \mathcal{U}$ называют *аналитической* в точке $\lambda_0 \in \mathbb{R}$, если ее можно представить в некоторой окрестности этой точки сходящимся (по норме $\|\cdot\|$) рядом вида (6.34), т.е.

$$u(\lambda) = \sum_{n=0}^{\infty} u_n(\lambda - \lambda_0)^n. \quad (6.35)$$

Так как в (6.35) можно ввести новое переменное $\lambda - \lambda_0 = \mu$, то в дальнейшем без ограничения общности принимаем $\lambda_0 = 0$ и вместо (6.35) будем рассматривать ряд

$$u(\lambda) = \sum_{n=0}^{\infty} u_n \lambda^n. \quad (6.36)$$

Как и в случае действительных рядов, здесь справедлива теорема Абеля [IX], которую можно сформулировать следующим

образом: если степенной ряд (6.36) с коэффициентами из банахова пространства $\mathcal{U}$ сходится в некоторой точке $\lambda_0 \neq 0$, то он сходится абсолютно в каждой точке интервала $(-|\lambda_0|, |\lambda_0|)$ и сходится равномерно на любом отрезке $[-r, r]$, где $r < |\lambda_0|$; если же этот ряд расходится в некоторой точке λ_1 , то он расходится и в любой точке λ , для которой $|\lambda| > |\lambda_1|$. Отсюда следует, что для всякого степенного ряда (6.36) с коэффициентами из банахова пространства $\mathcal{U}$ существует такое число $R \geq 0$, что:

- 1) при $|\lambda| < R$ этот ряд сходится, причем абсолютно;
- 2) при $|\lambda| \leq r < R$ он сходится абсолютно и равномерно;
- 3) при $|\lambda| > R$ этот ряд расходится.

Число R при этом называют радиусом сходимости этого ряда, а промежуток $(-R, R)$ — интервалом сходимости. Из теоремы Абеля следует, что радиус сходимости можно определить соотношением

$$R = \sup_{\lambda \in \Omega} |\lambda| \geq 0.$$

При $R = 0$ ряд (6.36) сходится в единственной точке $\lambda = 0$, а при $R = \infty$ этот ряд сходится при любых значениях λ . Оценку снизу для радиуса сходимости устанавливает следующая лемма.

Лемма 6.1. Для радиуса сходимости степенного ряда (6.36) с коэффициентами из банахова пространства $\mathcal{U}$ справедлива оценка $R \geq \frac{1}{k}$, если существуют такие постоянные $M > 0$ и $k > 0$, что в (6.36), начиная с некоторого номера $n \in \mathbb{N}$, $\|u_n\| \leq M k^n$.

◀ Достаточно показать, что при $|\lambda| < \frac{1}{k}$ ряд (6.36) сходится. Пусть $|\lambda| k = q < 1$. Тогда, начиная с некоторого номера $n \in \mathbb{N}$, имеем

$$\|u_n \lambda^n\| = \|u_n\| |\lambda|^n \leq M (|\lambda| k)^n = M q^n.$$

Так как члены геометрической прогрессии $M q^n$ образуют при $q < 1$ сходящийся числовой ряд, то ряд (6.36) в силу признака Вейерштрасса сходится, причем абсолютно и равномерно. ►

Теорема 6.3. Сумма $u(\lambda)$ степенного ряда (6.36) с коэффициентами из банахова пространства $\mathcal{U}$ непрерывна при $|\lambda| < R$, где R — радиус сходимости этого ряда.

◀ Поскольку ряд (6.36) в интервале $(-R, R)$ сходится абсолютно, то ряд с действительными членами $\|u_n\| \lambda^n$ также сходится при $|\lambda| < R$, и его можно почленно дифференцировать в этом интервале [IX]. Ряд с членами $n\|u_n\| \lambda^{n-1}$ сходится при $|\lambda| < R$.

Пусть $\lambda_0 \in (-R, R)$. Тогда для некоторого числа $\rho \in (0, R)$ справедливо неравенство $|\lambda_0| < \rho$. Обозначим

$$C(\rho) = \sum_{n=1}^{\infty} n\|u_n\| \rho^{n-1}. \quad (6.37)$$

Выберем любое число λ , для которого $|\lambda| < \rho$. Поскольку $u(\lambda)$ — сумма ряда (6.36) для $\lambda \in (-R, R)$, то

$$\begin{aligned} u(\lambda) - u(\lambda_0) &= \sum_{n=1}^{\infty} u_n(\lambda^n - \lambda_0^n) = \\ &= \sum_{n=1}^{\infty} u_n(\lambda^{n-1} + \lambda^{n-2}\lambda_0 + \dots + \lambda_0^{n-1})(\lambda - \lambda_0). \end{aligned}$$

Отсюда, учитывая (6.37), получаем

$$\begin{aligned} \|u(\lambda) - u(\lambda_0)\| &\leq \\ &\leq \sum_{n=1}^{\infty} \|u_n\| (|\lambda|^{n-1} + |\lambda|^{n-2}|\lambda_0| + \dots + |\lambda_0|^{n-1}) |\lambda - \lambda_0| \leq C(\rho) |\lambda - \lambda_0|, \end{aligned}$$

и поэтому функция $u(\lambda)$ непрерывна при $|\lambda| < R$. ▶

Степенные ряды вида (6.36) с коэффициентами из банахова пространства $\mathcal{U}$, так же как и ряды с действительными членами, можно бесконечно много раз дифференцировать почленно в интервале сходимости. Справедлива следующая теорема*.

*См.: Треногин В.А.

Теорема 6.4. Сумма $u(\lambda)$ степенного ряда (6.36) с коэффициентами из банахова пространства $\mathcal{U}$, имеющего радиус сходимости R , является бесконечно дифференцируемой функцией в любой точке интервала $(-R, R)$, причем для любого $k \in \mathbb{N}$ имеем

$$u^{(k)}(\lambda) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)u_n \lambda^{n-k}, \quad \lambda \in (-R, R), \quad (6.38)$$

где $u^{(k)}(\lambda)$ — производная k -го порядка функции $u(\lambda)$. #

Пусть функция $u(\lambda)$ со значениями в банаховом пространстве $\mathcal{U}$ бесконечно дифференцируема в точке $\lambda = 0$. Тогда степенной ряд

$$u(\lambda) = \sum_{n=0}^{\infty} \frac{u^{(n)}(0)}{n!} \lambda^n, \quad (6.39)$$

где $u^{(n)}(0)$ — производная n -го порядка функции $u(\lambda)$ в точке $\lambda = 0$, причем $u^{(0)}(0) = u(0)$ и $0! = 1$, называют рядом Тейлора функции $u(\lambda)$.

Теорема 6.5. Степенной ряд (6.36) с коэффициентами из банахова пространства $\mathcal{U}$, сходящийся в интервале $(-R, R)$, является рядом Тейлора своей суммы $u(\lambda)$.

◀ Если при $\lambda \in (-R, R)$ справедливо (6.36), то в силу теоремы 6.4 в любой точке λ интервала $(-R, R)$ существуют производные $u^{(k)}(\lambda)$, $k \in \mathbb{R}$, функции $u(\lambda)$. С учетом (6.38) имеем $u^{(k)}(0) = k!u_k$. Следовательно, $u_k = \frac{u^{(k)}(0)}{k!}$ и

$$\sum_{n=0}^{\infty} u_n \lambda^n = \sum_{n=0}^{\infty} \frac{u^{(n)}(0)}{n!} \lambda^n. \quad \blacktriangleright$$

Следующая теорема, играющая важную роль применительно к *методу малого параметра*, является очевидным следствием теоремы 6.5.

Теорема 6.6. Если в нормированном пространстве $\mathcal{U}$ два степенных ряда

$$\sum_{n=0}^{\infty} u_n \lambda^n = \sum_{n=0}^{\infty} v_n \lambda^n, \quad u_n, v_n \in \mathcal{U}, \quad (6.40)$$

равны при $|\lambda| < R$ для некоторого числа $R > 0$, то равны и их коэффициенты при одинаковых степенях λ , т.е. $u_k = v_k$ для любого k . #

Пусть $\mathcal{L}(B, \mathcal{W})$ — банахово пространство линейных *ограниченных операторов*, действующих из банахова пространства B в банахово пространство $\mathcal{W}$. Отображение $B: \Lambda \subset \mathbb{R} \rightarrow \mathcal{L}(B, \mathcal{W})$ будем называть *оператор-функцией*. Отметим некоторые свойства оператор-функций, связанные с понятием аналитичности.

1. Если оператор-функции $B_1(\lambda)$, $B_2(\lambda)$, $\lambda \in \Lambda$, со значениями в $\mathcal{L}(B, \mathcal{W})$ являются аналитическими в точке $\lambda_0 \in \Lambda$, то оператор-функция $B_1(\lambda) + B_2(\lambda)$, $\lambda \in \Lambda$, также аналитична в этой точке.

2. Если оператор-функции $B_1(\lambda)$, $B_2(\lambda)$, $\lambda \in \Lambda$, со значениями в $\mathcal{L}(B, \mathcal{W})$ и $\mathcal{L}(\mathcal{W}, B)$ соответственно аналитичны в точке $\lambda_0 \in \Lambda$, то оператор-функция $B_1(\lambda) \circ B_2(\lambda)$ также аналитична в этой точке.

3. Если оператор-функция $B(\lambda)$, $\lambda \in \Lambda$, со значениями в $\mathcal{L}(B)$ аналитична в точке $\lambda_0 \in \Lambda$ и $\|B(\lambda)\| < 1$ в некоторой окрестности этой точки, то оператор-функция $(I - B(\lambda))^{-1}$, где I — *тождественный оператор* в B , аналитична в точке λ_0 .

4. Если оператор-функция $B: \Lambda \subset \mathbb{R} \rightarrow \mathcal{L}(B, \mathcal{W})$ и функция $u: \Lambda \subset \mathbb{R} \rightarrow B$ аналитичны в точке $\lambda_0 \in \Lambda$, то функция $B(\lambda)u(\lambda)$, $\lambda \in \Lambda$, аналитична в этой точке. Более того, если в некоторой окрестности точки λ_0 справедливы разложения

$$B(\lambda) = \sum_{k=0}^{\infty} B_k \lambda^k \quad \text{и} \quad u(\lambda) = \sum_{k=0}^{\infty} u_k \lambda^k,$$

то в этой окрестности справедливо разложение

$$B(\lambda)u(\lambda) = \sum_{k=0}^{\infty} (B_k u_0 + B_{k-1} u_1 + \dots + B_0 u_k) \lambda^k. \quad (6.41)$$

Пусть задано операторное уравнение

$$B(\lambda)u = f(\lambda), \quad u \in \mathcal{B}, \quad \lambda \in \mathbb{R}, \quad (6.42)$$

причем $B(\lambda) \in \mathcal{L}(\mathcal{B}, \mathcal{W})$, $f(\lambda) \in \mathcal{W}$ и $\mathcal{B}$, $\mathcal{W}$ — банаховы пространства. Покажем, что если $B(\lambda)$ и $u(\lambda)$ аналитичны в точке $\lambda = 0$, а оператор $B_0 = B(0)$ имеет ограниченный обратный оператор, то в некоторой окрестности точки $\lambda = 0$ у оператора $B(\lambda)$ существует ограниченный обратный $B^{-1}(\lambda)$, являющийся аналитической оператор-функцией. При этом уравнение (6.42) имеет единственное решение $u(\lambda) = B^{-1}(\lambda)f(\lambda)$, которое также является аналитической функцией в точке $\lambda = 0$.

Представим оператор $B(\lambda)$ в виде

$$\begin{aligned} B(\lambda) &= B_0 - (B_0 - B(\lambda)) = \\ &= B_0 \circ (I - B_0^{-1} \circ (B_0 - B(\lambda))) = B_0 \circ (I - A(\lambda)), \end{aligned} \quad (6.43)$$

где $A(\lambda) = B_0^{-1} \circ (B_0 - B(\lambda))$ — оператор, отображающий $\mathcal{B}$ в себя и являющийся ограниченным как композиция ограниченных операторов.

Из аналитичности оператора $B(\lambda)$ в точке $\lambda = 0$, согласно теореме 6.3, вытекает его непрерывность в этой точке. Поэтому существует такое число $r > 0$, что при $|\lambda| < r$

$$\|A(\lambda)\| \leq \|B_0^{-1}\|_{\mathcal{L}'} \|B(\lambda) - B_0\|_{\mathcal{L}} < 1. \quad (6.44)$$

Следовательно, согласно теореме 4.18, при $|\lambda| < r$ оператор $I - A(\lambda)$ имеет ограниченный обратный оператор $(I - A(\lambda))^{-1}$. Но тогда и у $B(\lambda)$ существует ограниченный обратный оператор $B^{-1}(\lambda) = (I - A(\lambda))^{-1} \circ B_0^{-1}$. Так как $A(\lambda)$ является аналитической в точке $\lambda = 0$ оператор-функцией, то $B^{-1}(\lambda)$ также аналитическая оператор-функция в этой точке.

Таким образом, уравнение (6.42) при $|\lambda| < r$ имеет единственное решение $u(\lambda) = B^{-1}(\lambda)f(\lambda)$. Если функция $f(\lambda)$ аналитична в точке $\lambda = 0$, то решение $u(\lambda)$ — аналитическая в точке $\lambda = 0$ функция, т.е. при $|\lambda| < R$ для некоторого $R > 0$ в соответствии с (6.36) имеем

$$u(\lambda) = \sum_{k=0}^{\infty} u_k \lambda^k, \quad u_k \in B. \quad (6.45)$$

Покажем, как можно найти коэффициенты u_k в (6.45), если известны представления оператор-функции $B(\lambda)$ и функции $f(\lambda)$ в виде степенных рядов с коэффициентами из банахова пространства. Пусть при $|\lambda| < \rho$

$$B(\lambda) = \sum_{k=0}^{\infty} B_k \lambda^k \quad \text{и} \quad f(\lambda) = \sum_{k=0}^{\infty} f_k \lambda^k. \quad (6.46)$$

Тогда $R = \min\{\rho, r\}$. Подставляя (6.45) в (6.42) и учитывая (6.46), приходим при $|\lambda| < R$ к тождеству

$$\left(\sum_{k=0}^{\infty} B_k \lambda^k \right) \sum_{k=0}^{\infty} u_k \lambda^k = \sum_{k=0}^{\infty} f_k \lambda^k,$$

или в соответствии с (6.41)

$$\sum_{k=0}^{\infty} (B_k u_0 + B_{k-1} u_1 + \dots + B_0 u_k) \lambda^k = \sum_{k=0}^{\infty} f_k \lambda^k.$$

Приравнявая, согласно теореме 6.6, коэффициенты при одинаковых степенях λ , получаем систему операторных уравнений

$$B_0 u_0 = f_0, \quad \sum_{k=0}^n B_k u_{n-k} = f_n, \quad n \in \mathbb{N}. \quad (6.47)$$

Решая эти уравнения последовательно, имеем

$$u_0 = B_0^{-1} f_0,$$

$$u_1 = B_0^{-1} \circ (f_1 - B_1 u_0),$$

$$u_2 = B_0^{-1} \circ (f_2 - B_1 u_1 - B_2 u_0),$$

.....

$$u_n = B_0^{-1} \circ (f_n - B_1 u_{n-1} - B_2 u_{n-2} - \dots - B_{n-1} u_1 - B_n u_0),$$

.....

Вычислив $n + 1$ первых элементов u_k , $k = \overline{0, n}$, найдем n -ю частичную сумму ряда (6.45), которую можно принять за приближенное решение

$$\tilde{u}(\lambda) = u_0 + \sum_{k=1}^n u_k \lambda^k, \quad |\lambda| < R = \min \{\rho, r\} \quad (6.48)$$

уравнения (6.42).

Пусть $B(\lambda)$ и $f(\lambda)$ в (6.42) таковы, что для коэффициентов B_k и f_k в разложениях (6.46), начиная с некоторого номера $k \geq 1$, справедливы оценки $\|f_k\|_W \leq M_1 \alpha^k$ и $\|B_k \circ B_0^{-1}\| \leq M \beta^{k-1}$. Тогда (6.42) имеет единственное решение в виде степенного ряда с радиусом сходимости $R \geq \min \left\{ \frac{1}{\alpha}, \frac{1}{(M + \beta)} \right\}$. Действительно, при $\alpha|\lambda| = q < 1$, начиная с некоторого номера k , имеем

$$\|f_k \lambda^k\|_W = \|f_k\|_W |\lambda|^k \leq M_1 \alpha^k |\lambda|^k = M q^k.$$

Отсюда, согласно лемме 6.1, следует оценка снизу $\rho \geq \frac{1}{\alpha}$ для радиуса сходимости степенного разложения $f(\lambda)$ в (6.46). Кроме того, если $|\lambda| < \frac{1}{(M + \beta)}$, а значит, и $|\lambda|\beta < 1$, то, используя (6.43), первое разложение (6.46) и формулу для суммы членов

убывающей прогрессии, получаем

$$\begin{aligned}\|A(\lambda)\| &= \|B_0^{-1} \circ (B(\lambda) - B_0)\| \leq \|B_0^{-1} \circ \left(\sum_{k=1}^{\infty} B_k \lambda^k\right)\| = \\ &= \left\|\sum_{k=1}^{\infty} B_0^{-1} \circ B_k \lambda^k\right\| \leq M|\lambda| \sum_{k=1}^{\infty} (|\lambda|\beta)^{k-1} = \\ &= \frac{M|\lambda|}{1 - |\lambda|\beta} < \frac{M|\lambda|}{1 - \frac{\beta}{M+\beta}} = |\lambda|(M + \beta) < 1.\end{aligned}$$

Таким образом, при $r = \frac{1}{(M+\beta)}$ выполнено условие (6.44), и (6.42) имеет единственное решение, которое при $|\lambda| < r$ можно представить в виде степенного ряда (6.45) с радиусом сходимости

$$R \geq \min \left\{ \frac{1}{\alpha}, \frac{1}{M+\beta} \right\}. \quad (6.49)$$

Пример 6.3. Линейное интегральное уравнение

$$u(\xi) - \frac{1}{2\pi} \int_{-\pi}^{\pi} u(t) \cos(\xi - t + \lambda \xi t) dt = f(\xi), \quad (6.50)$$

определяемое параметром $\lambda \neq 0$ и функцией $f \in C[-\pi, \pi]$, не зависящей от λ , рассмотрим как операторное уравнение с линейным оператором $B(\lambda): C[-\pi, \pi] \rightarrow C[-\pi, \pi]$, являющимся ограниченным для любого $\lambda \in \mathbb{R}$. Так как f не зависит от λ , то в данном случае в (6.47) $f_0(\xi) \equiv f(\xi)$ и $f_n(\xi) \equiv 0$, $n \in \mathbb{N}$.

Найдем разложение оператор-функции $B(\lambda)$ в степенной ряд. Поскольку

$$\begin{aligned}\frac{d^k}{d\lambda^k} \cos(\xi - t + \lambda \xi t) \Big|_{\lambda=0} &= \\ &= (\xi t)^k \cos\left(\xi - t + \lambda \xi t + k \frac{\pi}{2}\right) \Big|_{\lambda=0} = (\xi t)^k \cos\left(\xi - t + k \frac{\pi}{2}\right),\end{aligned}$$

то получим разложение в ряд Тейлора [IX]

$$\cos(\xi - t + \lambda \xi t) = \sum_{k=0}^{\infty} (\xi t)^k \cos\left(\xi - t + k \frac{\pi}{2}\right) \frac{\lambda^k}{k!}, \quad \lambda \in \mathbb{R}.$$

Следовательно, для любого элемента $u \in C[-\pi, \pi]$ имеем

$$B(\lambda) u(\xi) = u(\xi) - \frac{1}{2\pi} \int_{-\pi}^{\pi} u(t) \left(\sum_{k=0}^{\infty} (\xi t)^k \cos\left(\xi - t + k \frac{\pi}{2}\right) \frac{\lambda^k}{k!} \right) dt.$$

Поскольку при $t, \lambda \in [-\pi, \pi]$

$$\sum_{k=0}^{\infty} \frac{|u(t)(\xi t \lambda)^k \cos(\xi - t + k\pi/2)|}{k!} \leq \sum_{k=0}^{\infty} \frac{\|u\| \pi^{2k} |\lambda|^k}{k!} = \|u\| e^{\pi^2 |\lambda|},$$

то символы суммы и интеграла можно поменять местами [IX].

В итоге для любого $\lambda \in \mathbb{R}$ имеем

$$B(\lambda) u(\xi) = u(\xi) - \sum_{k=0}^{\infty} \frac{\lambda^k}{2\pi} \int_{-\pi}^{\pi} u(t) (\xi t)^k \frac{\cos(\xi - t + k\pi/2)}{k!} dt,$$

или $B(\lambda) = \sum_{k=0}^{\infty} B_k \lambda^k$, где

$$B_0 u(\xi) = u(\xi) - \frac{1}{2\pi} \int_{-\pi}^{\pi} u(t) \cos(\xi - t) dt,$$

$$B_k u(\xi) = -\frac{1}{2\pi} \int_{-\pi}^{\pi} u(t) (\xi t)^k \frac{\cos(\xi - t + k\pi/2)}{k!} dt, \quad k \in \mathbb{N}.$$

Будем искать решение уравнения (6.50) в виде

$$u(\xi) = \sum_{k=0}^{\infty} u_k(\xi) \lambda^k, \quad \xi \in [-\pi, \pi], \quad u_k \in C[-\pi, \pi]. \quad (6.51)$$

Первое уравнение в (6.47) примет вид

$$u_0(\xi) - \frac{1}{2\pi} \int_{-\pi}^{\pi} u_0(t) \cos(\xi - t) dt = f(\xi), \quad (6.52)$$

а для произвольного $n \in \mathbb{N}$

$$u_n(\xi) - \frac{1}{2\pi} \sum_{k=0}^n \int_{-\pi}^{\pi} u_{n-k}(t) (\xi t)^k \frac{\cos(\xi - t + k\pi/2)}{k!} dt = 0. \quad (6.53)$$

Так как $\cos(\xi - t) = \cos \xi \cos t + \sin \xi \sin t$, равенство (6.52) можно записать в виде

$$u_0(\xi) = f(\xi) + C_1 \cos \xi + C_2 \sin \xi, \quad (6.54)$$

где

$$C_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} u_0(t) \cos t dt, \quad C_2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} u_0(t) \sin t dt.$$

Умножим (6.54) поочередно на $\cos \xi$ и $\sin \xi$ и после интегрирования по ξ от $-\pi$ до π найдем

$$C_1 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos t dt, \quad C_2 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin t dt.$$

Таким образом, (6.54) эквивалентно уравнению

$$u_0(\xi) = f(\xi) + \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(\xi - t) dt, \quad (6.55)$$

т.е. (6.52) для любой функции $f \in C[-\pi, \pi]$ имеет и при том единственное решение в $C[-\pi, \pi]$. Поэтому существует оператор B_0^{-1} , обратный к оператору B_0 , причем для норм в $C[-\pi, \pi]$ имеем

$$\begin{aligned} \|u_0\| = \|B_0^{-1}f\| &\leq \|f\| + \frac{1}{\pi} \max_{\xi \in [-\pi, \pi]} \left| \int_{-\pi}^{\pi} f(t) \cos(\xi - t) dt \right| \leq \\ &\leq \|f\| + \frac{\|f\|}{\pi} \max_{\xi \in [-\pi, \pi]} \int_{-\pi}^{\pi} |\cos(t - \xi)| d(t - \xi) \leq \\ &\leq \|f\| + \frac{\|f\|}{\pi} \int_{-\pi}^{\pi} |\cos t| dt \leq \left(1 + \frac{4}{\pi}\right) \|f\|. \end{aligned}$$

Следовательно, $\|B_0^{-1}\| \leq 1 + \frac{4}{\pi}$.

При $n = 1$ в (6.53) имеем

$$u_1(\xi) - \frac{1}{2\pi} \int_{-\pi}^{\pi} u_1(t) \cos(\xi - t) dt - \frac{1}{2\pi} \int_{-\pi}^{\pi} u_0(t) \xi t \cos\left(\xi - t + \frac{\pi}{2}\right) dt = 0,$$

т.е. интегральное уравнение, аналогичное (6.52). Таким образом, функция $u_1(\xi)$ также представима в явном виде. Подобным образом можно найти решения уравнений, получаемых из (6.53) для последующих значений n .

Можно показать, что для нормы оператора B_k , $k \in \mathbb{N}$, справедливо соотношение $\|B_k\| \leq \frac{2}{\pi} \pi^{2k} = 2\pi(\pi^2)^{k-1}$, и поэтому $\|B_k B_0^{-1}\| \leq (2\pi + 8)(\pi^2)^{k-1}$, т.е. в (6.49) $M = 2\pi + 8$ и $\beta = \pi^2$, а так как в данном случае в (6.46) $f_k \equiv 0$, $k \in \mathbb{N}$, то в (6.49) $\alpha = 0$. Таким образом, существует и единственно решение уравнения (6.50) в виде степенного ряда (6.51), который в соответствии с (6.49) имеет радиус сходимости $R \geq \frac{1}{\pi^2 + 2\pi + 8} \approx 0,0414$.

Пример 6.4. В качестве иллюстрации применения метода малого параметра к решению краевой задачи рассмотрим нахождение приближенного установившегося распределения температуры $T(r, z)$ в тонкой оптической линзе (рис. 6.2). Неоднородное температурное поле, возникающее при воздействии на линзу различных энергетических факторов, вызывает термоупругие напряжения, которые влияют на качество изображения в оптической системе в связи с появлением термооптических aberrаций*.

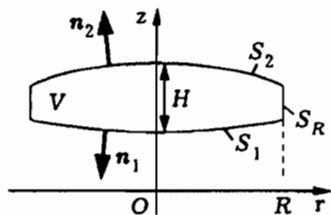


Рис. 6.2

Пусть объем V , занимаемый осесимметричной линзой толщиной H , ограничен участком S_R круговой цилиндрической поверхности радиуса R и двумя гладкими преломляющими поверхностями S_1 и S_2 , образованными вращением вокруг оси линзы кривых, заданных зависимостями $z = f_1(r)$ и $z = f_2(r)$ соответственно (см. рис. 6.2). При постоянных значениях коэффициента λ теплопроводности линзы и объемной мощности $I_V^{(q)}$ энерговыделения установившееся осесимметричное распределение температуры в линзе, описываемое функцией $T(r, z)$, зависит от двух координат: радиальной r и осевой z . Эта функция удовлетворяет уравнению Пуассона

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial T}{\partial r} \right) + \frac{\partial^2 T}{\partial z^2} = - \frac{I_V^{(q)}}{\lambda}. \quad (6.56)$$

Температуру на поверхности S_R примем равной T_R , т.е. граничное условие на этой поверхности имеет вид

$$T(R, z) = T_R. \quad (6.57)$$

На преломляющих поверхностях происходит теплообмен с окружающей средой, имеющей постоянную температуру T_* . Его

*См. Зино И.Е., Тропп Э.А.

интенсивность на поверхностях S_1 и S_2 характеризуют коэффициенты теплообмена α_1 и α_2 соответственно, входящие в граничные условия третьего рода [XII]

$$\lambda \frac{\partial T}{\partial n_1} \Big|_{S_1} = \alpha_1 (T_* - T) \Big|_{S_1}, \quad \lambda \frac{\partial T}{\partial n_2} \Big|_{S_2} = \alpha_2 (T_* - T) \Big|_{S_2}. \quad (6.58)$$

Производные функции $T(r, z)$ по направлениям n_1 и n_2 , определяемым единичными векторами $\mathbf{n}_1$ и $\mathbf{n}_2$ внешних нормалей в точках поверхностей S_1 и S_2 соответственно, можно выразить через производные по координатам r и z , используя направляющие косинусы этих нормалей относительно осей Ox и Oz :

$$\begin{aligned} \cos(\widehat{n_1, r}) &= \frac{f'_1(r)}{\sqrt{1 + f_1'^2(r)}}, & \cos(\widehat{n_1, z}) &= -\frac{1}{\sqrt{1 + f_1'^2(r)}}, \\ \cos(\widehat{n_2, r}) &= -\frac{f'_2(r)}{\sqrt{1 + f_2'^2(r)}}, & \cos(\widehat{n_2, z}) &= \frac{1}{\sqrt{1 + f_2'^2(r)}}. \end{aligned}$$

Введем безразмерные координаты $\rho = \frac{r}{R}$ и $\zeta = \frac{z}{H}$ и обозначим $\varepsilon = \frac{H}{R}$, $Q = I_V^{(q)} \frac{R^2}{\lambda}$. Тогда краевая задача (6.56)–(6.58) будет эквивалентна следующей краевой задаче для функции $T(\rho, \zeta)$. Уравнение (6.56) примет вид

$$\frac{\varepsilon^2}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial T}{\partial \rho} \right) + \frac{\partial^2 T}{\partial \zeta^2} = -\varepsilon^2 Q, \quad (6.59)$$

условие (6.57) — $T(1, \zeta) = T_R$, а вместо (6.58) получим

$$\begin{aligned} \left(\varepsilon^2 \tilde{f}'_1(\rho) \frac{\partial T}{\partial \rho} - \frac{\partial T}{\partial \zeta} \right) \Big|_{\zeta = \tilde{f}_1'^2(\rho)} &= \\ &= \text{Bi}_1 (T_* - T) \Big|_{\zeta = \tilde{f}_1'^2(\rho)} \sqrt{1 + \varepsilon^2 \tilde{f}_1'^2(\rho)} \end{aligned} \quad (6.60)$$

и

$$\begin{aligned} - \left(\varepsilon^2 \tilde{f}'_2(\rho) \frac{\partial T}{\partial \rho} - \frac{\partial T}{\partial \zeta} \right) \Big|_{\zeta = \tilde{f}_2'^2(\rho)} &= \\ &= \text{Bi}_2 (T_* - T) \Big|_{\zeta = \tilde{f}_2'^2(\rho)} \sqrt{1 + \varepsilon^2 \tilde{f}_2'^2(\rho)}, \end{aligned} \quad (6.61)$$

где $\tilde{f}_1(\rho) = \frac{f_1(R\rho)}{H}$, $\tilde{f}_2(\rho) = \frac{f_2(R\rho)}{H}$ и $Bi_1 = \frac{\alpha_1 H}{\lambda}$, $Bi_2 = \frac{\alpha_2 H}{\lambda}$. Отметим, что безразмерный параметр $Bi = \frac{\alpha H}{\lambda}$ является отношением термического сопротивления $\frac{H}{\lambda}$ пластины толщиной H , выполненной из материала линзы, к термическому сопротивлению $\frac{1}{\alpha}$ теплообмена с окружающей средой*.

В данном случае (6.59) соответствует операторному уравнению $B(\varepsilon)T = f(\varepsilon)$, где $f(\varepsilon) = -\varepsilon^2 Q$ и $B: C^2(V) \rightarrow C(V)$, причем $B = (B_0 + \varepsilon^2 B_2) \in \mathcal{L}(C^2(V), C(V))$, где

$$B_0 = \frac{\partial^2}{\partial \zeta^2}, \quad B_2 = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial}{\partial \rho} \right).$$

Для тонкой линзы параметр $\varepsilon^2 = \left(\frac{H}{R}\right)^2 \ll 1$, т.е. является малым. Поэтому в (6.60), (6.61) примем

$$\sqrt{1 + \varepsilon^2 \tilde{f}_1'^2(\rho)} \approx 1 + \frac{1}{2} \varepsilon^2 \tilde{f}_1'^2(\rho) - \frac{1}{8} \varepsilon^4 \tilde{f}_1'^4(\rho), \quad (6.62)$$

$$\sqrt{1 + \varepsilon^2 \tilde{f}_2'^2(\rho)} \approx 1 + \frac{1}{2} \varepsilon^2 \tilde{f}_2'^2(\rho) - \frac{1}{8} \varepsilon^4 \tilde{f}_2'^4(\rho). \quad (6.63)$$

Решим сначала задачу (6.59)–(6.61), не принимая во внимание граничное условие на поверхности S_R . Распределение температуры с учетом вида функций $B(\varepsilon)$ и $f(\varepsilon)$ будем искать в виде

$$\tilde{T}(\rho, \zeta) = \varepsilon^2 \tilde{T}_0(\rho, \zeta) + \varepsilon^4 \tilde{T}_1(\rho, \zeta) + \dots, \quad (6.64)$$

считая, что числа Био Bi_1 и Bi_2 имеют порядок единицы. Подставляя (6.64) в (6.59) и учитывая приближенные равенства (6.62) и (6.63) в уравнениях (6.60) и (6.61), после приравнивания нулю выражений при каждой степени параметра ε получим

*Этот параметр называют числом Био в честь французского физика Ж. Б. Био (1774–1862).

последовательно решаемую совокупность краевых задач относительно функций $\tilde{T}_0(\rho, \zeta)$, $\tilde{T}_1(\rho, \zeta)$ и т.д. Так, для нахождения первой из этих функций имеем уравнение

$$\frac{\partial^2 \tilde{T}_0}{\partial \zeta^2} = -Q \quad (6.65)$$

с граничными условиями

$$-\frac{\partial \tilde{T}_0}{\partial \zeta} = \text{Bi}_1(T_* - \tilde{T}_0), \quad \frac{\partial \tilde{T}_0}{\partial \zeta} = \text{Bi}_2(T_* - \tilde{T}_0) \quad (6.66)$$

при $\zeta = \tilde{f}_1(\rho)$ и $\zeta = \tilde{f}_2(\rho)$ соответственно. Функция $\tilde{T}_1(\rho, \zeta)$ должна удовлетворять уравнению

$$\frac{\partial^2 \tilde{T}_1}{\partial \zeta^2} = -\frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial \tilde{T}_0}{\partial \rho} \right)$$

и граничным условиям

$$-\frac{\partial \tilde{T}_1}{\partial \zeta} + \text{Bi}_1 \tilde{T}_1 = -\tilde{f}'_1(\rho) \frac{\partial \tilde{T}_0}{\partial \rho} - \frac{\text{Bi}_1}{2} \tilde{f}_1'^2(\rho) \tilde{T}_0,$$

$$\frac{\partial \tilde{T}_1}{\partial \zeta} + \text{Bi}_2 \tilde{T}_1 = \tilde{f}'_2(\rho) \frac{\partial \tilde{T}_0}{\partial \rho} - \frac{\text{Bi}_2}{2} \tilde{f}_2'^2(\rho) \tilde{T}_0$$

при $\zeta = \tilde{f}_1(\rho)$ и $\zeta = \tilde{f}_2(\rho)$ соответственно.

Дважды интегрируя уравнение (6.65), получаем

$$\tilde{T}_0(\rho, \zeta) = -\frac{Q}{2} \zeta^2 + C_1(\rho) \zeta + C_0(\rho). \quad (6.67)$$

Подстановка этого представления в граничные условия (6.66) приводит к системе алгебраических уравнений относительно функций $C_1(\rho)$ и $C_2(\rho)$. Решая систему, находим

$$C_1(\rho) = Q \frac{\tilde{f}_1 \text{Bi}_2 + \tilde{f}_2 \text{Bi}_1 + \frac{1}{2} \text{Bi}_1 \text{Bi}_2 (\tilde{f}_2^2 - \tilde{f}_1^2)}{\text{Bi}_1 + \text{Bi}_2 + \text{Bi}_1 \text{Bi}_2 (\tilde{f}_2 - \tilde{f}_1)}$$

и

$$C_0(\rho) = Q \tilde{f}_1 \tilde{f}_2 \frac{\frac{1}{2} \text{Bi}_1 \text{Bi}_2 (\tilde{f}_2 - \tilde{f}_1) - \text{Bi}_1 - \text{Bi}_2}{\text{Bi}_1 + \text{Bi}_2 + \text{Bi}_1 \text{Bi}_2 (\tilde{f}_2 - \tilde{f}_1)} + \\ + Q \frac{\tilde{f}_2 - \tilde{f}_1 + \frac{1}{2} (\text{Bi}_1 \tilde{f}_1^2 + \text{Bi}_2 \tilde{f}_2^2)}{\text{Bi}_1 + \text{Bi}_2 + \text{Bi}_1 \text{Bi}_2 (\tilde{f}_2 - \tilde{f}_1)}.$$

В этих формулах у функций $\tilde{f}_1$ и $\tilde{f}_2$ опущен аргумент ρ . Далее можно перейти к нахождению функции $\tilde{T}_1(\rho, \zeta)$ и т.д.

Подчеркнем, что функции $\tilde{T}_0(\rho, \zeta)$, $\tilde{T}_1(\rho, \zeta)$, ... найдены без учета граничного условия на поверхности S_R . Поэтому и решение $\tilde{T}(\rho, \zeta)$ в виде (6.64) не обязательно удовлетворяет этому условию.

Найдем теперь решение $\bar{T}(\rho, \zeta)$ краевой задачи для однородного уравнения, соответствующего (6.59), и однородных граничных условий, соответствующих (6.60), (6.61), но потребуем, чтобы было выполнено условие $\tilde{T}(1, \zeta) + \bar{T}(1, \zeta) = T_R$ на поверхности S_R . При решении этой задачи используем безразмерные координаты $\xi = \frac{1-\rho}{\varepsilon}$ и ζ . Тогда решение $u(\xi, \zeta) = \bar{T}(1 - \varepsilon\xi, \zeta)$ на поверхности S_R должно удовлетворять равенству

$$u(0, \zeta) + \tilde{T}(1, \zeta) = T_R. \quad (6.68)$$

Решение $u(\xi, \zeta)$ будем опять искать в виде

$$\tilde{u}(\xi, \zeta) = \varepsilon^2 \tilde{u}_0(\xi, \zeta) + \varepsilon^4 \tilde{u}_1(\xi, \zeta) + \dots, \quad (6.69)$$

причем $u_n(\xi, \zeta) \rightarrow 0$ при $\xi \rightarrow \infty$ для $n = 0, 1, 2, \dots$, и в соответствии с (6.64) и (6.68) $u_0(0, \zeta) = T_R - \tilde{T}_0(1, \zeta)$, $u_1(0, \zeta) = -\tilde{T}_1(1, \zeta)$ и т.д. Кроме того, функция $u(\xi, \zeta)$ должна удовлетворять однородному уравнению

$$\frac{\partial^2 u}{\partial \xi^2} + \frac{\varepsilon}{1 - \varepsilon\xi} \frac{\partial u}{\partial \xi} + \frac{\partial^2 u}{\partial \zeta^2} = 0, \quad (6.70)$$

которое соответствует неоднородному уравнению (6.59) после замены $\xi = \frac{1-\rho}{\varepsilon}$, и однородным граничным условиям

$$\bar{f}'_1(\xi) \frac{\partial u}{\partial \xi} - \frac{\partial u}{\partial \zeta} + \text{Bi}_1 u \sqrt{1 + \varepsilon^2 \bar{f}_1'^2(\xi)} = 0, \quad (6.71)$$

$$-\bar{f}'_2(\xi) \frac{\partial u}{\partial \xi} + \frac{\partial u}{\partial \zeta} + \text{Bi}_2 u \sqrt{1 + \varepsilon^2 \bar{f}_2'^2(\xi)} = 0 \quad (6.72)$$

при $\zeta = \bar{f}_1(\xi) = \tilde{f}_1(1 - \varepsilon\xi)$ и $\zeta = \bar{f}_2(\xi) = \tilde{f}_2(1 - \varepsilon\xi)$ соответственно, которые следуют из условий (6.60), (6.61) после указанной замены.

Если подставить (6.69) в (6.70), то, приравнявая нулю выражения при каждой степени параметра ε , для нахождения функции $u_0(\xi, \zeta)$ получим уравнение Лапласа

$$\frac{\partial^2 u_0}{\partial \xi^2} + \frac{\partial^2 u_0}{\partial \zeta^2} = 0. \quad (6.73)$$

Введение безразмерной координаты ξ привело к растяжению области, прилегающей к поверхности S_R , в направлении изменения этой координаты. Поэтому эту область можно приближенно представить полубесконечной полосой, заданной неравенствами

$$0 \leq \xi < \infty, \quad \bar{f}_1(0) \leq \zeta \leq \bar{f}_2(0),$$

и (6.73) рассматривать в такой полосе. Тогда, подставляя разложение (6.69) в уравнения (6.71) и (6.72), а затем приравнявая нулю выражение при ε^2 , можно положить $\bar{f}'_1(\xi) = \bar{f}'_2(\xi) = 0$ и для функции $u_0(\xi, \zeta)$ получить однородные граничные условия

$$-\frac{\partial u_0}{\partial \zeta} + \text{Bi}_1 u_0 = 0, \quad \frac{\partial u_0}{\partial \zeta} + \text{Bi}_2 u_0 = 0 \quad (6.74)$$

при $\zeta = \bar{f}_1(0)$ и $\zeta = \bar{f}_2(0)$ соответственно. Краевая задача для (6.73) с граничными условиями (6.74), $u_0(0, \zeta) = T_R - \tilde{T}_0(1, \zeta)$

и $u_0(\xi, \zeta) \rightarrow 0$ при $\xi \rightarrow \infty$ может быть решена методом Фурье (разделения переменных) [XII] или при помощи интегрального преобразования либо на отрезке $[\bar{f}_1(0), \bar{f}_2(0)]$, либо в полуограниченном промежутке $[0, \infty]$ [XI]. Формулировки краевых задач для нахождения функций $u_n(\xi, \zeta)$, $n \in \mathbb{N}$, входящих в правую часть (6.69), являются более громоздкими, но решение этих задач можно получить теми же методами.

В итоге, согласно принципу суперпозиции решений [XII], функция

$$T(\rho, \zeta) = \tilde{T}(\rho, \zeta) + \bar{T}(\rho, \zeta) = \tilde{T}(\rho, \zeta) + u\left(\frac{1-\rho}{\varepsilon}, \zeta\right)$$

будет решением исходной краевой задачи (6.56)–(6.58).

6.5. Метод ортогональных проекций

Пусть искомая функция u удовлетворяет операторному уравнению

$$Au = f, \quad (6.75)$$

где A — линейный непрерывный оператор, области определения $D(A)$ и значений $R(A)$ которого являются всюду плотными подмножествами гильбертова пространства $\mathcal{H}$ со скалярным произведением $\langle \cdot, \cdot \rangle$. Если перенести заданный элемент f в левую часть этого уравнения, то получим равенство

$$Au - f = 0,$$

где 0 — нулевой элемент в $\mathcal{H}$. Выберем в $\mathcal{H}$ счетный базис $\{v_k\}$. Тогда это равенство, согласно теореме 5.1, можно заменить системой эквивалентных равенств

$$\langle Au - f, v_k \rangle = 0, \quad k \in \mathbb{N}. \quad (6.76)$$

Пусть последовательность $\{u_n\}$ образует в $D(A)$ счетный базис. Тогда в соответствии с (4.52) искомый элемент $u \in D(A)$

можно единственным образом представить в виде

$$u = \sum_{n=1}^{\infty} a_n u_n, \quad a_n \in \mathbb{R}, \quad u_n \in D(A). \quad (6.77)$$

Подставляя (6.77) в (6.76) и учитывая свойства скалярного произведения и оператора A , получаем

$$\begin{aligned} \langle Au, v_k \rangle &= \left\langle A \sum_{n=1}^{\infty} a_n u_n, v_k \right\rangle = \left\langle A \lim_{N \rightarrow \infty} \sum_{n=1}^N a_n u_n, v_k \right\rangle = \\ &= \left\langle \lim_{N \rightarrow \infty} \sum_{n=1}^N a_n Au_n, v_k \right\rangle = \lim_{N \rightarrow \infty} \sum_{n=1}^N a_n \langle Au_n, v_k \rangle = \\ &= \sum_{n=1}^{\infty} a_n \langle Au_n, v_k \rangle = \langle f, v_k \rangle, \quad k \in \mathbb{N}. \end{aligned}$$

Таким образом, система равенств (6.76) равносильна системе

$$\sum_{n=1}^{\infty} a_n \langle Au_n, v_k \rangle = \langle f, v_k \rangle, \quad k \in \mathbb{N}, \quad (6.78)$$

представляющей собой бесконечную систему линейных алгебраических уравнений относительно координат a_n элемента u в базисе $\{u_n\}$. Если в (6.77) ограничиться первыми N элементами u_n счетного базиса, т.е. приближенно принять

$$u \approx \tilde{u}_N = \sum_{n=1}^N a_n u_n, \quad u_n \in D(A), \quad (6.79)$$

и в (6.78) ограничиться первыми N равенствами, в которых суммирование выполняется от 1 до N , то получим конечную систему N линейных алгебраических уравнений (СЛАУ)

$$\sum_{n=1}^N a_n \langle Au_n, v_k \rangle = \langle f, v_k \rangle, \quad k = \overline{1, N}, \quad (6.80)$$

относительно N первых координат a_n . Элементами матрицы такой системы будут скалярные произведения $\langle Au_n, v_k \rangle$.

Решения СЛАУ (6.80) при различных N естественно рассматривать как приближения неизвестного решения уравнения (6.75). Если, в частности, $\{v_k\}$ является ортонормированным базисом в $\mathcal{H}$, то решению СЛАУ (6.80) соответствует решение $\tilde{u}_N$ вида (6.79) уравнения $A_N u = \tilde{f}_N$, где A_N — оператор, действующий из N -мерного подпространства $D_N(A)$, совпадающего с *линейной оболочкой* системы $\{u_n\}_N$, в N -мерное подпространство $R_N(A)$, совпадающее с *линейной оболочкой* системы $\{v_k\}_N$. При этом $A_N u$ совпадает с N -й частичной суммой разложения элемента Au по базису $\{v_k\}$, а $\tilde{f}_N$ является N -й частичной суммой ряда $f = \sum_{k=1}^{\infty} f_k v_k$, где $f_k = \langle f, v_k \rangle$.

Однако в случае произвольного базиса $\{v_k\}$, даже если определитель матрицы СЛАУ (6.80) отличен от нуля, нельзя утверждать, что последовательность $\{\tilde{u}_N\}$ приближенных решений вида (6.79) при $N \rightarrow \infty$ сходится к *классическому* u_0 или *слабому* u_* *решениям* уравнения (6.75). При некоторых ограничениях ответ на вопрос о сходимости $\{\tilde{u}_N\}$ к u_0 или к u_* удастся решить теоретическим путем (см. Д.6.1). Однако на практике часто приходится ограничиваться лишь сравнением между собой нескольких приближенных решений.

Описанная схема построения приближенного решения уравнения (6.75), приводящая к СЛАУ (6.80), лежит в основе большой группы приближенных аналитических методов, которые обычно объединяют под общим названием **метод ортогональных проекций**. Такое название объясняется тем, что равенства (6.80) представляют собой условия ортогональности элемента $A\tilde{u}_N - f$ всем базисным элементам v_k N -мерного подпространства $R_N(A)$. При этом **функции** u_n , $n = \overline{1, N}$, которые используют для представления приближенного решения (6.79), называют **базисными**, а **функции** v_k в СЛАУ (6.80) — **проекционными**.

Метод ортогональных проекций является частным случаем *проекционного метода*, в котором условия равенства нулю проекций **невязки** $A\tilde{u}_N - f$ **операторного уравнения** на элементы v_k , $k = \overline{1, N}$, базиса N -мерного подпространства $\mathcal{H}_N \subset R(A)$ имеют более общий вид (см. Д.6.1). Если $R(A) \subset L_2(\mathbb{R})$, то чаще метод ортогональных проекций называют **методом взвешенных невязок**. В этом случае счетный базис в $R(A)$ будет образован последовательностью $\{v_k\}$ действительных функций v_k , $k \in \mathbb{N}$, так что (6.80) можно рассматривать как N условий равенства нулю определенного в $L_2(\mathbb{R})$ скалярного произведения невязки операторного уравнения и **весовых функций** v_k , $k = \overline{1, N}$. Иногда процедуру применения (6.80) называют **методом моментов**, или методом Галеркина — Петрова*.

Особенности каждого конкретного метода в группе методов, определяемых условиями (6.80), зависят от выбора счетных базисов в $D(A)$ и $R(A)$. Такие методы рассмотрены ниже.

6.6. Коллокации в подобластях и в точках

Рассмотрим один из наиболее простых приемов нахождения приближенного решения *операторного уравнения* $Au = f$ (6.75). Пусть *области определения* $D(A)$ и *значений* $R(A)$ оператора A , входящего в (6.75), являются *всюду плотными подмножествами гильбертова пространства* $L_2(V)$ *функций, суммируемых* (интегрируемых по Лебегу) *с квадратом* в области $V \subset \mathbb{R}^m$.

Разобьем V на N подобластей V_k , $k = \overline{1, N}$, так, чтобы выполнялось равенство

$$\bigcup_{k=1}^N V_k = V_0,$$

*Б.Г. Галёркин (1871–1945) — российский инженер и ученый в области механики. Г.И. Петров (1912–1987) — российский ученый в области механики.

где $V_k \cap V_l = \emptyset$ при $k \neq l$, причем мера Лебега множества $V \setminus V_0$ равна нулю. Для любого $k = \overline{1, N}$ примем

$$v_k(x) = \begin{cases} 1, & x \in V_k; \\ 0, & x \in V \setminus V_k. \end{cases} \quad (6.81)$$

Используя схему построения *метода ортогональных проекций*, приближенное решение уравнения $Au = f$ будем искать в виде (6.79)

$$\tilde{u}_N = \sum_{n=1}^N a_n u_n,$$

где u_n , $n = \overline{1, N}$, — первые N элементов *счетного базиса* в $D(A)$, а N коэффициентов a_n являются решением системы линейных алгебраических уравнений (СЛАУ) вида (6.80)

$$\sum_{n=1}^N a_n \int_{V_k} A u_n dV = \int_{V_k} f dV, \quad k = \overline{1, N}, \quad (6.82)$$

поскольку в $L_2(V)$ с учетом (6.81)

$$\begin{aligned} \langle A u_n, v_k \rangle &= \int_V A u_n v_k dV = \int_{V_k} A u_n dV, \\ \langle f, v_k \rangle &= \int_V f v_k dV = \int_{V_k} f dV. \end{aligned}$$

Описанную процедуру поиска приближенного решения уравнения $Au = f$ называют **методом коллокации в подобластях** V_k , $k = \overline{1, N}$ (по латыни collocatio — размещение, расстановка), или **методом разделения области** V .

Пример 6.5. Особенности применения метода коллокации в подобластях рассмотрим на примере хорошо известной

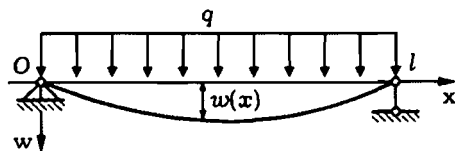


Рис. 6.3

задачи из курса сопротивления материалов, имеющей точное решение. Пусть шарнирно закрепленная на концах балка имеет постоянную по ее длине l жесткость EI_* на изгиб (E — модуль упругости материала балки, I_* — геометрический момент инерции ее поперечного сечения) и нагружена равномерно распределенной по длине поперечной нагрузкой $q = \text{const}$ (рис. 6.3). Зависимость поперечного прогиба $w(x)$ балки от продольной координаты x удовлетворяет обыкновенному дифференциальному уравнению (ОДУ) четвертого порядка

$$EI_* \frac{d^4 w(x)}{dx^4} = q, \quad x \in (0, l), \quad (6.83)$$

и граничным условиям

$$w(0) = w(l) = \frac{d^2 w(x)}{dx^2} \Big|_{x=0} = \frac{d^2 w(x)}{dx^2} \Big|_{x=l} = 0.$$

Перейдем в (6.83) и в граничных условиях к безразмерным величинам, обозначив $\xi = \frac{x}{l}$ и $u(\xi) = \frac{EI_* w(\xi)}{ql^4}$:

$$\begin{cases} \frac{d^4 u(\xi)}{d\xi^4} = 1, & \xi \in (0, 1), \\ u(0) = u(1) = \frac{d^2 u(\xi)}{d\xi^2} \Big|_{\xi=0} = \frac{d^2 u(\xi)}{d\xi^2} \Big|_{\xi=1} = 0. \end{cases} \quad (6.84)$$

Решение краевой задачи (6.84) можно получить последовательным интегрированием и нахождением произвольных постоянных из граничных условий.

Дважды интегрируя дифференциальное уравнение в задаче (6.84), получаем

$$\frac{d^2 u(\xi)}{d\xi^2} = \frac{1}{2}\xi^2 + C_1\xi + C_2$$

и из условия $\left. \frac{d^2 u(\xi)}{d\xi^2} \right|_{\xi=0} = 0$ находим $C_2 = 0$, а затем из условия $\left. \frac{d^2 u(\xi)}{d\xi^2} \right|_{\xi=1} = 0$ вычисляем $C_1 = -\frac{1}{2}$, так что можно записать промежуточный результат в виде ОДУ второго порядка

$$\frac{d^2 u(\xi)}{d\xi^2} = \frac{\xi^2}{2} - \frac{\xi}{2}, \quad \xi \in (0, 1), \quad (6.85)$$

с граничными условиями $u(0) = u(1) = 0$. В свою очередь двукратное интегрирование (6.85) дает

$$u(\xi) = \frac{\xi^4}{24} - \frac{\xi^3}{12} + C_3\xi + C_4.$$

Из условия $u(0) = 0$ следует, что $C_4 = 0$, а затем из условия $u(1) = 0$ вычисляем $C_3 = \frac{1}{24}$. В итоге получаем

$$u(\xi) = \frac{\xi^4}{24} - \frac{\xi^3}{12} + \frac{\xi}{24}. \quad (6.86)$$

Найдем теперь приближенное решение краевой задачи (6.84) методом коллокации и сравним с точным решением. Точность приближения будем оценивать по отлнчию приближенного решения от точного в точке максимального прогиба балки. Для точного решения из необходимого условия экстремума функции $u(\xi)$

$$\frac{du(\xi)}{d\xi} = \frac{\xi^3}{6} - \frac{\xi^2}{4} + \frac{1}{24} = 0$$

находим стационарную точку $\xi = \frac{1}{2}$, в которой эта функция имеет максимум, так как из (6.85) следует, что $\frac{d^2 u(\xi)}{d\xi^2} < 0$ при $\xi \in (0, 1)$. В этой точке максимальное значение безразмерного прогиба балки в соответствии с (6.86) равно

$$u^* = \frac{5}{384} \approx 0,01302. \quad (6.87)$$

Оценим теперь это значение при помощи метода коллокации.

В (6.84) операторное уравнение содержит линейный оператор $A = \frac{d^4}{d\xi^4}$, определенный на всюду плотном линейном многообразии $X \subset L_2[0, 1]$ четырежды непрерывно дифференцируемых на отрезке $[0, 1]$ функций. В качестве счетного базиса в этом подпространстве выберем систему функций $\sin n\pi\xi \in D(A)$, $n \in \mathbb{N}$, удовлетворяющих однородным граничным условиям краевой задачи (6.84).

В методе коллокации сначала ограничимся случаем $N = 1$. Тогда в представлении (6.79) имеем $u(\xi) \approx u_1(\xi) = a_1 \sin \pi\xi$, а в (6.82) $V_1 = V = [0, 1]$ и $dV = d\xi$, так что из (6.82) получим

$$a_1 \int_0^1 \frac{d^4 \sin \pi\xi}{d\xi^4} d\xi = \int_0^1 d\xi,$$

или

$$a_1 \frac{d^3 \sin \pi\xi}{d\xi^3} \Big|_0^1 = -a_1 \pi^3 \cos \pi\xi \Big|_0^1 = 1.$$

Отсюда $a_1 = \frac{1}{2\pi^3} \approx 0,01613$, т.е. в этом приближении максимальный безразмерный прогиб $u_1^* = \frac{1}{2\pi^3}$ балки примерно на четверть превосходит его точное значение u^* в (6.87).

При $N = 2$ имеем $u_2(\xi) = a_1 \sin \pi\xi + a_2 \sin 2\pi\xi$. Точкой $\xi = 1/2$ разобьем отрезок $[0, 1]$ на два промежутка $V_1 = [0, 1/2]$ и

$V_2 = (1/2, 1]$ и в соответствии с (6.82) запишем СЛАУ относительно коэффициентов a_1 и a_2 :

$$a_1 \int_0^{1/2} \frac{d^4 \sin \pi \xi}{d\xi^4} d\xi + a_2 \int_0^{1/2} \frac{d^4 \sin 2\pi \xi}{d\xi^4} d\xi = \int_0^{1/2} d\xi,$$

$$a_1 \int_{1/2}^1 \frac{d^4 \sin \pi \xi}{d\xi^4} d\xi + a_2 \int_{1/2}^1 \frac{d^4 \sin 2\pi \xi}{d\xi^4} d\xi = \int_{1/2}^1 d\xi.$$

Учитывая, что

$$\frac{d^3 \sin \pi \xi}{d\xi^3} = -\pi^3 \cos \pi \xi, \quad \frac{d^3 \sin 2\pi \xi}{d\xi^3} = -8\pi^3 \cos 2\pi \xi,$$

определим коэффициенты записанной СЛАУ:

$$\int_0^{1/2} \frac{d^4 \sin \pi \xi}{d\xi^4} d\xi = \pi^3, \quad \int_0^{1/2} \frac{d^4 \sin 2\pi \xi}{d\xi^4} d\xi = 16\pi^3,$$

$$\int_{1/2}^1 \frac{d^4 \sin \pi \xi}{d\xi^4} d\xi = \pi^3, \quad \int_{1/2}^1 \frac{d^4 \sin 2\pi \xi}{d\xi^4} d\xi = -16\pi^3.$$

Так как правые части обоих уравнений равны $1/2$, окончательно получаем СЛАУ

$$a_1 + 16a_2 = \frac{1}{2\pi^3}, \quad a_1 - 16a_2 = \frac{1}{2\pi^3},$$

из которой следует, что $a_2 = 0$ и снова $a_1 = \frac{1}{2\pi^3}$. Итак, при равномерном разбиении отрезка на две части ($N = 2$) получили то же решение, что и при $N = 1$. Такой результат не является случайным. Дело в том, что функция $\sin 2\pi \xi$ нарушает симметрию прогиба симметрично нагруженной балки (см. рис. 6.3).

При $N = 3$ используем равномерное разбиение отрезка $[0, 1]$ на три промежутка $V_1 = [0, 1/3]$, $V_2 = (1/3, 2/3)$ и $V_3 = [2/3, 1]$. В этом случае

$$u_3(\xi) = a_1 \sin \pi \xi + a_2 \sin 2\pi \xi + a_3 \sin 3\pi \xi, \quad (6.88)$$

и система (6.82) принимает вид

$$a_1 \int_0^{1/3} \frac{d^4 \sin \pi \xi}{d\xi^4} d\xi + a_2 \int_0^{1/3} \frac{d^4 \sin 2\pi \xi}{d\xi^4} d\xi + a_3 \int_0^{1/3} \frac{d^4 \sin 3\pi \xi}{d\xi^4} d\xi = \int_0^{1/3} d\xi,$$

$$a_1 \int_{1/3}^{2/3} \frac{d^4 \sin \pi \xi}{d\xi^4} d\xi + a_2 \int_{1/3}^{2/3} \frac{d^4 \sin 2\pi \xi}{d\xi^4} d\xi + a_3 \int_{1/3}^{2/3} \frac{d^4 \sin 3\pi \xi}{d\xi^4} d\xi = \int_{1/3}^{2/3} d\xi,$$

$$a_1 \int_{2/3}^1 \frac{d^4 \sin \pi \xi}{d\xi^4} d\xi + a_2 \int_{2/3}^1 \frac{d^4 \sin 2\pi \xi}{d\xi^4} d\xi + a_3 \int_{2/3}^1 \frac{d^4 \sin 3\pi \xi}{d\xi^4} d\xi = \int_{2/3}^1 d\xi.$$

Вычисляя интегралы, получаем СЛАУ

$$\begin{cases} a_1 + 24a_2 + 108a_3 = \frac{2}{3\pi^3}, \\ a_1 - 54a_3 = \frac{1}{3}, \\ a_1 - 24a_2 + 108a_3 = \frac{2}{3\pi^3}, \end{cases}$$

решением которой являются $a_1 = \frac{4}{9\pi^3}$, $a_2 = 0$, $a_3 = \frac{1}{486\pi^3}$. При $\xi = 1/2$ вычислим максимальный безразмерный прогиб балки

$$u_3^* = a_1 - a_3 = \frac{4}{9\pi^3} - \frac{1}{486\pi^3} = \frac{215}{486\pi^3} \approx 0,01427,$$

превышающий примерно на 10 % его точное значение u^* .

Теперь применим метод коллокации в подобластях к решению ОДУ (6.85) второго порядка. При $N = 1$ имеем $\tilde{u}_1(\xi) = \tilde{a}_1 \sin \pi \xi$ и в соответствии с (6.82) запишем

$$\tilde{a}_1 \int_0^1 \frac{d^2 \sin \pi \xi}{d\xi^2} d\xi = \frac{1}{2} \int_0^1 (\xi^2 - \xi) d\xi, \quad \text{или} \quad \tilde{a}_1 \frac{d \sin \pi \xi}{d\xi} \Big|_0^1 = -\frac{1}{12}.$$

Отсюда $-2\pi \tilde{a}_1 = -\frac{1}{12}$, или $\tilde{a}_1 = \frac{1}{24\pi} \approx 0,01326$. В этом приближении максимальный безразмерный прогиб $\tilde{u}_1^* \approx 0,01326$ балки лишь на 2% выше его точного значения u^* . В случае $N = 2$ возникает ситуация, аналогичная рассмотренной выше.

При $N = 3$ приближенное решение имеет вид

$$\tilde{u}_3(\xi) = \tilde{a}_1 \sin \pi \xi + \tilde{a}_2 \sin 2\pi \xi + \tilde{a}_3 \sin 3\pi \xi. \quad (6.89)$$

Тогда, разбивая отрезок $[0, 1]$ равномерно на три промежутка и используя (6.82), приходим к СЛАУ

$$\begin{cases} \tilde{a}_1 + 6\tilde{a}_2 + 12\tilde{a}_3 = \frac{7}{162\pi}, \\ \tilde{a}_1 - 6\tilde{a}_3 = \frac{13}{324\pi}, \\ \tilde{a}_1 - 6\tilde{a}_2 + 12\tilde{a}_3 = \frac{7}{162\pi}. \end{cases}$$

Отсюда сначала находим $\tilde{a}_2 = 0$, а затем $\tilde{a}_1 = \frac{10}{243\pi}$ и $\tilde{a}_3 = \frac{1}{5832\pi}$. Максимальный безразмерный прогиб балки

$$\tilde{u}_3^* = \tilde{a}_1 - \tilde{a}_3 = \frac{10}{243\pi} - \frac{1}{5832\pi} = \frac{239}{5832\pi} \approx 0,01305,$$

что всего на 0,23% отличается от его точного значения u^* .

Таким образом, применение метода коллокации по подобластям к решению одной и той же задачи о прогибе балки, но описываемой ОДУ различных порядков, дает более точный результат в случае уравнения более низкого порядка. #

Перейдем к рассмотрению другого пути построения приближенного решения операторного уравнения $Au = f$. Коэффициенты a_n в представлении (6.79) будем искать из условия равенства нулевому элементу 0 невязки $Au - f$ этого операторного уравнения в заданной системе N точек $x_k \in V$, $k = \overline{1, N}$. Тогда вместо (6.82) получим СЛАУ

$$\sum_{n=1}^N a_n Au_n(x_k) = f(x_k), \quad k = \overline{1, N}. \quad (6.90)$$

Такую процедуру поиска приближенного решения уравнения $Au = f$ называют **методом коллокации в точках**, а точки $x_k \in V$ — **точками коллокации**.

Пример 6.6. Применим метод коллокации в точках к решению краевой задачи (6.84). При $N = 1$, приняв как и в примере 6.5 $u_1(\xi) = a_1 \sin \pi \xi$, для точки $\xi_1 = 1/2$ в середине отрезка $[0, 1]$ вместо (6.90) запишем

$$a_1 \frac{d^4 \sin \pi \xi}{d\xi^4} \Big|_{\xi=1/2} = 1,$$

или $a_1 \pi^4 = 1$, т.е. $a_1 = \frac{1}{\pi^4} \approx 0,01027$. В этом случае максимальный безразмерный прогиб $u_1^* = a_1$ балки примерно на 20 % ниже его точного значения u^* в (6.87). Отметим, что результат весьма чувствителен к выбору точки коллокации. Так, выбрав $\xi_1 = \frac{1}{4}$, получим $u_1^* = \frac{\sqrt{2}}{\pi^4} \approx 0,01452$, а при $\xi_1 = \frac{1}{6}$ получим $u_1^* = \frac{2}{\pi^4} \approx 0,02053$. Ясно, что существует точка $\xi_1 \in (0, 1)$, обеспечивающая совпадение значения u_1^* с u^* , однако ее можно найти, лишь зная заранее значение u^* . Но и в этом случае приближенное решение $u_1(\xi)$ не будет совпадать на отрезке $[0, 1]$ с точным решением (6.86).

Нетрудно проверить, что при $N = 2$ и симметрично расположенных на отрезке $[0, 1]$ точках коллокации $\xi_1 = t$ и $\xi_2 = 1 - t$ в приближенном решении $u_2(\xi) = a_1 \sin \pi \xi + a_2 \sin 2\pi \xi$ всегда $a_2 = 0$

и $a_1 = \frac{1}{\pi^4 \sin \pi t}$. Так, при $\xi_1 = \frac{1}{3}$ и $\xi_2 = \frac{2}{3}$ имеем $a_1 = \frac{2}{\sqrt{3}\pi^4} \approx 0,01185$.

В случае $N = 3$ функция $u_3(\xi)$ имеет вид (6.88) и при выборе равномерно расположенных на отрезке $[0, 1]$ точек коллокации $\xi_1 = \frac{1}{4}$, $\xi_2 = \frac{1}{2}$ и $\xi_3 = \frac{3}{4}$, используя (6.90), приходим к СЛАУ

$$\begin{cases} \frac{a_1}{\sqrt{2}} + 16a_2 + \frac{81a_3}{\sqrt{2}} = \frac{1}{\pi^4}, \\ a_1 - 81a_3 = \frac{1}{\pi^4}, \\ \frac{a_1}{\sqrt{2}} - 16a_2 + \frac{81a_3}{\sqrt{2}} = \frac{1}{\pi^4}, \end{cases}$$

из которой находим $a_1 = \frac{\sqrt{2}+1}{2\pi^4}$, $a_2 = 0$ и $a_3 = \frac{\sqrt{2}-1}{162\pi^4}$. Приближенное значение максимального безразмерного прогиба балки равно

$$u_3^* = a_1 - a_3 = \frac{\sqrt{2}+1}{2\pi^4} - \frac{\sqrt{2}-1}{162\pi^4} = \frac{40\sqrt{2}+41}{81\pi^4} \approx 0,01237,$$

что на 5 % меньше его точного значения u^* в (6.87).

Применим теперь метод коллокации в точках к решению ОДУ (6.85) второго порядка. Используя при $N = 1$ приближенное решение в виде $\tilde{u}_1(\xi) = \tilde{a}_1 \sin \pi \xi$ и точку коллокации $\xi_1 = 1/2$, в соответствии с (6.90) запишем

$$\tilde{a}_1 \frac{d^2 \sin \pi \xi}{d\xi^2} \Big|_{\xi=1/2} = \frac{1}{2} (\xi^2 - \xi) \Big|_{\xi=1/2}.$$

Отсюда $-\pi^2 \tilde{a}_1 = -\frac{1}{8}$, или $\tilde{a}_1 = \frac{1}{8\pi^2} \approx 0,01267$. В этом приближении максимальный безразмерный прогиб $\tilde{u}_1^* \approx 0,01267$ балки лишь на 2,7 % ниже его точного значения u^* в (6.87). При $N = 2$ ситуация аналогична рассмотренной выше. В частности, при равномерно расположенных на отрезке $[0, 1]$ точках коллокации $\xi_1 = 1/3$ и $\xi_2 = 2/3$ имеем $\tilde{a}_2 = 0$ и $\tilde{a}_1 = \tilde{u}_2^* = \frac{2}{9\sqrt{3}\pi^2} \approx 0,01300$.

Используя при $N = 3$ (6.90) и (6.89) и принимая $\xi_1 = 1/4$, $\xi_2 = 1/2$ и $\xi_3 = 3/4$, получаем СЛАУ

$$\begin{cases} \frac{\tilde{a}_1}{\sqrt{2}} + 4\tilde{a}_2 + \frac{9\tilde{a}_3}{\sqrt{2}} = \frac{3}{32\pi^2}, \\ \tilde{a}_1 - 9\tilde{a}_3 = \frac{1}{8\pi^2}, \\ \frac{\tilde{a}_1}{\sqrt{2}} - 4\tilde{a}_2 + \frac{9\tilde{a}_3}{\sqrt{2}} = \frac{3}{32\pi^2}. \end{cases}$$

Отсюда находим сначала $\tilde{a}_2 = 0$, а затем $\tilde{a}_1 = \frac{3\sqrt{2}+4}{64\pi^4}$ и $\tilde{a}_3 = \frac{3\sqrt{2}-4}{576\pi^2}$. Вычисляя максимальный безразмерный прогиб балки

$$\tilde{u}_3^* = \tilde{a}_1 - \tilde{a}_3 = \frac{3\sqrt{2}+4}{64\pi^2} - \frac{3\sqrt{2}-4}{576\pi^2} = \frac{3\sqrt{2}+5}{72\pi^2} \approx 0,01301,$$

устанавливаем, что он менее чем на 0,1 % отличается от точного значения u^* в (6.87).

Итак, применение метода коллокации по точкам дало лучшие результаты (как и в примере 6.5) при описании прогиба балки при помощи ОДУ более низкого порядка.

Пример 6.7. Рассмотрим применение метода коллокации для решения уравнений в частных производных. Остановимся на сравнительно простом примере линейной одномерной задачи нестационарной теплопроводности в пластине толщиной h . Пусть зависимость температуры $T(t, x)$ от времени t и координаты x удовлетворяет уравнению теплопроводности

$$\frac{\partial T}{\partial t} = \tilde{a} \frac{\partial^2 T}{\partial x^2}, \quad t > 0, \quad x \in [0, h], \quad (6.91)$$

с начальным условием $T(0, x) = T_0$ при $t = 0$ и граничными условиями

$$\left. \frac{\partial T}{\partial x} \right|_{x=0} = 0, \quad T(t, h) = T^*, \quad (6.92)$$

где $\tilde{a}$ — коэффициент температуропроводности материала пластины, а T^* — заданная температура поверхности пластины при $x = h$. Отметим, что первое условие в (6.92) по физическому смыслу означает идеальную тепловую изоляцию поверхности пластины при $x = 0$.

Приближенное решение задачи будем искать в виде

$$\tilde{T}(t, x) = T^* - D(t) \left(1 - \left(\frac{x}{h}\right)^2\right), \quad (6.93)$$

удовлетворяющем (6.92). Неизвестную функцию $D(t)$ сначала найдем приближенно методом коллокации в подобластях, положив $N = 1$. Для этого после подстановки (6.93) в (6.91) проведем осреднение по всей толщине пластины:

$$\begin{aligned} \int_0^h \left(\frac{\partial \tilde{T}}{\partial t} - \tilde{a} \frac{\partial^2 \tilde{T}}{\partial x^2} \right) dx &= - \int_0^h \left(1 - \frac{x^2}{h^2} \right) \frac{dD(t)}{dt} dx + \\ &+ \int_0^h \tilde{a} D(t) \frac{d}{dx} \left(-\frac{2x}{h^2} \right) dx = -\frac{2h}{3} \frac{dD(t)}{dt} - \frac{2\tilde{a}}{h} D(t) = 0. \end{aligned}$$

В результате приходим к однородному ОДУ

$$\frac{dD(t)}{dt} + \frac{3\tilde{a}}{h^2} D(t) = 0, \quad (6.94)$$

решение которого имеет вид

$$D(t) = D(0) \exp\left(-3\frac{\tilde{a}t}{h^2}\right) = D(0) \exp(-3Fo), \quad (6.95)$$

где $Fo = \frac{\tilde{a}t}{h^2}$ — число Фурье, имеющее смысл безразмерного времени, а $D(0)$ — пока неизвестное начальное значение функции $D(t)$.

Отметим, что в данном случае осреднение (6.91) по толщине пластины соответствует выполнению закона сохранения

тепловой энергии для этой пластины. Поэтому применительно к задачам теплопроводности метод коллокации в подобластях при $N = 1$ иногда называют интегральным методом теплового баланса.

Для нахождения значения $D(0)$ в (6.95) можно потребовать, чтобы распределение $\tilde{T}(t, x)$ (6.93) температуры при $t = 0$ удовлетворяло в среднем по толщине пластины начальному условию:

$$\frac{1}{h} \int_0^h \tilde{T}(0, x) dx = T^* - \frac{D(0)}{h} \int_0^h \left(1 - \frac{x^2}{h^2}\right) dx = T^* - \frac{2}{3} D(0) = T_0.$$

Отсюда $D(0) = \frac{3}{2}(T^* - T_0)$ и в итоге, подставляя (6.95) в (6.93), получаем приближенное решение

$$\tilde{T}(t, x) = T^* - \frac{3}{2}(T^* - T_0) \left(1 - \frac{x^2}{h^2}\right) \exp(-3Fo).$$

Введем новую функцию

$$\tilde{\Theta}(Fo, \xi) = \frac{T^* - \tilde{T}(t, x)}{T^* - T_0},$$

где $\xi = \frac{x}{h}$, и представим полученное приближенное решение в безразмерной форме:

$$\tilde{\Theta}_1(Fo, \xi) = 3 \frac{1 - \xi^2}{2} \exp(-3Fo), \quad \xi \in [0, 1]. \quad (6.96)$$

Известно точное решение $T(t, x)$ рассматриваемой задачи. Это решение в безразмерной форме имеет вид*

$$\begin{aligned} \Theta(Fo, \xi) &= \frac{T^* - T(t, x)}{T^* - T_0} = \\ &= \frac{4}{\pi} \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{2m-1} \cos \frac{2m-1}{2} \pi \xi \exp\left(-\frac{(2m-1)^2 \pi^2}{4} Fo\right). \end{aligned} \quad (6.97)$$

*См., например: Зарубин В.С., 1983.

Таблица 6.1

Fo	Точное решение $\Theta(Fo, \xi)$ (6.97)		Приближенное решение $\tilde{\Theta}_1(Fo, \xi)$				
			(6.96)	(6.98), $\xi_1 = \xi_1^*$		(6.100), (6.101)	(6.125)
	$\xi = 0$	$\xi = \xi_1^*$	$\xi = 0$	$\xi = 0$	$\xi = \xi_1^*$	$\xi = 0$	$\xi = 0$
0,00	1,000	1,000	1,500	1,234	1,000	1,000	1,250
0,02	1,000	0,995	1,413	1,175	0,952	1,000	1,189
0,04	0,999	0,954	1,330	1,118	0,906	1,000	1,131
0,06	0,992	0,897	1,253	1,063	0,862	1,000	1,076
0,08	0,975	0,842	1,180	1,013	0,821	1,000	1,023
0,10	0,949	0,792	1,111	0,964	0,781	0,951	0,974
0,20	0,772	0,605	0,823	0,753	0,610	0,705	0,758
0,30	0,609	0,471	0,610	0,588	0,477	0,522	0,590
0,40	0,474	0,368	0,452	0,460	0,373	0,387	0,460
0,50	0,371	0,287	0,337	0,359	0,291	0,288	0,358
0,60	0,290	0,225	0,248	0,281	0,228	0,212	0,279
0,80	0,177	0,137	0,136	0,171	0,139	0,116	0,169
1,00	0,108	0,084	0,075	0,105	0,085	0,064	0,103
1,50	0,032	0,024	0,017	0,031	0,025	0,015	0,029
2,00	0,009	0,007	0,004	0,009	0,007	0,003	0,008

В табл. 6.1 приведены результаты расчета (с точностью до трех знаков после запятой) значений функций $\tilde{\Theta}_1(Fo, \xi)$ и $\Theta(Fo, \xi)$ при $\xi = 0$ по формулам (6.96) и (6.97) для различных значений Fo . Сравнение этих результатов показывает, что приближение (6.96) является довольно грубым, особенно при малых значениях Fo (при $Fo \leq 0,1$ результаты расчета по формуле (6.96) противоречат физическому смыслу задачи).

Погрешность приближенного решения можно уменьшить вдвое, если значение $D(0)$ в (6.95) найти, исходя из условия, что среднеквадратичное отклонение

$$\Delta R = \frac{1}{h} \int_0^h (\tilde{T}(0, x) - T_0)^2 dx$$

распределения температуры в момент времени $t = 0$ от заданного начального значения T_0 достигает минимума. Вычислим

это отклонение, используя представление (6.93) распределения температуры:

$$\begin{aligned}\Delta R &= \frac{1}{h} \int_0^h \left(T^* - T_0 - D(0) \left(1 - \frac{x^2}{h^2} \right) \right)^2 dx = \\ &= (T^* - T_0)^2 - 4 \frac{T^* - T_0}{3} D(0) + \frac{8}{15} D^2(0).\end{aligned}$$

Из условия $\frac{d\Delta R}{dD(0)} = 0$ получим $D(0) = \frac{5}{4}(T^* - T_0)$. Однако и в этом случае большое отличие коэффициента 3 при числе Фурье в отрицательном показателе экспоненты приближенного решения (6.96) от соответствующего коэффициента $\frac{\pi^2}{4} \approx 2,4674$ в первом слагаемом точного решения (6.97) приводит к значительной погрешности.

Теперь для нахождения функции $D(t)$ в приближенном представлении $\tilde{T}(t, x)$ (6.93) распределения температуры используем метод коллокации в точках. Потребуем, чтобы оно удовлетворяло уравнению (6.91) и начальному условию в одной точке коллокации с координатой $x_1 \in [0, h)$. Тогда, подставляя (6.93) в (6.91) и полагая $x = x_1$, получаем однородное ОДУ

$$\left(1 - \frac{x_1^2}{h^2} \right) \frac{dD(t)}{dt} + \frac{2\tilde{a}}{h^2} D(t) = 0.$$

Отсюда

$$D(t) = D(0) \exp \left(-2 \frac{\tilde{a}t}{h^2} \frac{1}{1 - \frac{x_1^2}{h^2}} \right) = D(0) \exp \left(-2 \frac{Fo}{1 - \frac{x_1^2}{h^2}} \right).$$

Начальное значение $D(0)$ находим, используя (6.93) при $t = 0$, из условия $\tilde{T}(0, x_1) = T_0$:

$$T_0 = T^* - D(0) \left(1 - \frac{x_1^2}{h^2} \right), \quad \text{или} \quad D(0) = \frac{T^* - T_0}{1 - \left(\frac{x_1}{h} \right)^2}.$$

В итоге получаем

$$\tilde{\Theta}_2(Fo, \xi) = \frac{1 - \xi^2}{1 - \xi_1^2} \exp\left(-2 \frac{Fo}{1 - \xi_1^2}\right), \quad \xi_1 = \frac{x_1}{h}. \quad (6.98)$$

Ясно, что лучшие результаты (6.98) даст при расчете изменения температуры во времени в точке коллокации, т.е. при $\xi = \xi_1$. Если значение ξ_1 выбрать из условия равенства показателей экспонент в (6.98) и в первом слагаемом точного решения (6.97), то получим $\xi_1^* = \sqrt{1 - \frac{8}{\pi^2}} \approx 0,4352$, что приводит к удовлетворительным результатам расчета по этим формулам при $Fo > 0,1$ не только в точке $\xi = \xi_1^*$, но и на поверхности пластины при $\xi = 0$ (см. табл. 6.1).

Для уменьшения погрешности приближенного решения при малых значениях числа Фурье можно добавить в правую часть (6.93) дополнительные слагаемые с зависящими от времени коэффициентами, для нахождения которых потребуется использовать методы коллокации в подобластях или в точках при $N > 1$. В этом случае придется решать систему из N ОДУ. Если начальное распределение температуры в пластине однородно, то понизить погрешность приближенного решения при малых значениях числа Фурье можно также путем условного разбиения процесса нестационарной теплопроводности на две стадии (рис. 6.4). Первая стадия соответствует распространению возмущения температурного поля от поверхности при $x = h$, вызванного скачкообразным изменением температуры от начального значения T_0 до значения T^* , а вторая стадия — изменению температуры по всей толщине пластины одновременно.

Величина $h_*(t)$ на рис. 6.4 характеризует глубину проникновения темпера-

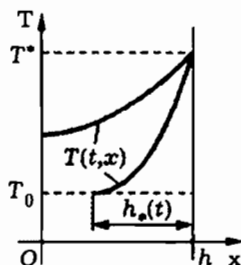


Рис. 6.4

турного возмущения и изменяется на первой стадии от 0 до h . В наиболее простом виде распределение температуры на этой стадии можно приближенно представить в виде

$$T_1(t, x) = T_0 + (T^* - T_0) \left(1 - \frac{h-x}{h_*(t)}\right)^2, \quad h - h_*(t) \leq x \leq h. \quad (6.99)$$

Подставляя (6.99) в (6.91) и интегрируя по переменному x в пределах от $h - h_*(t)$ до h , в соответствии с интегральным методом теплового баланса получаем

$$\begin{aligned} \int_{h-h_*(t)}^h \left(\frac{\partial T_1}{\partial t} - \tilde{a} \frac{\partial^2 T_1}{\partial x^2} \right) dx &= (T^* - T_0) \int_{h-h_*(t)}^h \frac{\partial}{\partial t} \left(1 - \frac{h-x}{h_*(t)}\right)^2 dx - \\ &- \tilde{a} (T^* - T_0) \int_{h-h_*(t)}^h \frac{\partial^2}{\partial x^2} \left(1 - \frac{h-x}{h_*(t)}\right)^2 dx = \\ &= (T^* - T_0) \left(\frac{1}{3} \frac{dh_*(t)}{dt} - \frac{2\tilde{a}}{h_*(t)} \right) = 0, \end{aligned}$$

или ОДУ $h_*(t) dh_*(t) = 6\tilde{a} dt$ с начальным условием $h_*(0) = 0$. После подстановки решения $h_*(t) = \sqrt{12at}$ этого ОДУ в (6.99) при $1 - \sqrt{12Fo} \leq \xi \leq 1$ имеем

$$\Theta_1(Fo, \xi) = \frac{T^* - T_1(t, x)}{T^* - T_0} = 1 - \left(1 - \frac{1 - \xi}{\sqrt{12Fo}}\right)^2. \quad (6.100)$$

Температурное возмущение достигает противоположной поверхности пластины ($\xi = 0$) за время t_1 , которое соответствует значению $Fo_1 = \frac{at_1}{h^2} = \frac{1}{12}$.

Примем, что распределение температуры на второй стадии процесса имеет вид (6.93). Поэтому функция $D(t)$ в (6.93) снова будет удовлетворять однородному ОДУ (6.94). Однако значение $D(0)$ в решении (6.95) этого ОДУ теперь следует искать из условия совпадения распределений температуры в

конце первой (при $Fo = \frac{1}{12}$) и начале второй стадий процесса теплопроводности, что дает $D(0) = T^* - T_0$. В итоге при $Fo \geq \frac{1}{12}$ получим

$$\tilde{\Theta}_3(Fo, \xi) = \frac{T^* - \tilde{T}(t, x)}{T^* - T_0} = (1 - \xi^2) \exp\left(-3\left(Fo - \frac{1}{12}\right)\right). \quad (6.101)$$

Эту формулу можно записать также в виде

$$\tilde{\Theta}_3(Fo, \xi) = \frac{T^* - \tilde{T}(t, x)}{T^* - T_0} = A'(1 - \xi^2) \exp(-3Fo), \quad (6.102)$$

где $A' = \exp(0,25) \approx 1,2840$ довольно близко к значению коэффициента $A_1 = \frac{4}{\pi} \approx 1,2732$ при первом члене точного решения (6.97). Результаты расчетов по (6.100) и (6.101) при $\xi = 0$ представлены выше (см. табл. 6.1) и достаточно хорошо согласуются с точным решением.

6.7. Метод наименьших квадратов

Рассмотрим еще один метод приближенного решения *операторного уравнения* $Au = f$, где A — линейный непрерывный оператор, области определения $D(A)$ и значений $R(A)$ которого являются всюду плотными подмножествами гильбертова пространства $\mathcal{H}$. Пусть система $\{u_n\} \subset D(A)$ образует счетный базис в $D(A)$. Приближенное решение

$$\tilde{u}_N = \sum_{n=1}^N a_n u_n, \quad a_n \in \mathbb{R}, \quad (6.103)$$

уравнения $Au = f$ будем искать из условия минимума нормы $\|A\tilde{u}_N - f\|$ невязки этого операторного уравнения, в котором коэффициенты a_n принимают любые действительные значения.

Из этого условия получаем

$$\min_{\tilde{\mathbf{u}}_N \in D(A)} \|\mathbf{A}\tilde{\mathbf{u}}_N - \mathbf{f}\|^2 = \min_{a_n \in \mathbb{R}} \left\| \sum_{n=1}^N a_n \mathbf{A}\mathbf{u}_n - \mathbf{f} \right\|^2. \quad (6.104)$$

Для произвольных значений $a_n \in \mathbb{R}$, $n = \overline{1, N}$, имеем

$$\begin{aligned} \left\| \sum_{n=1}^N a_n \mathbf{A}\mathbf{u}_n - \mathbf{f} \right\|_{\mathcal{H}}^2 &= \left\langle \sum_{n=1}^N a_n \mathbf{A}\mathbf{u}_n - \mathbf{f}, \sum_{k=1}^N a_k \mathbf{A}\mathbf{u}_k - \mathbf{f} \right\rangle = \\ &= \sum_{n=1}^N a_n \sum_{k=1}^N a_k \langle \mathbf{A}\mathbf{u}_n, \mathbf{A}\mathbf{u}_k \rangle - 2 \sum_{k=1}^N a_k \langle \mathbf{f}, \mathbf{A}\mathbf{u}_k \rangle + \langle \mathbf{f}, \mathbf{f} \rangle. \end{aligned} \quad (6.105)$$

Таким образом, нужно найти минимум неотрицательной функции N переменных a_n , $n = \overline{1, N}$. Необходимыми условиями достижения этого минимума [V] будут равенства нулю частных производных правой части (6.105) по переменным a_k , $k = \overline{1, N}$, что приводит к системе линейных алгебраических уравнений (СЛАУ)

$$\sum_{n=1}^N a_n \langle \mathbf{A}\mathbf{u}_n, \mathbf{A}\mathbf{u}_k \rangle = \langle \mathbf{f}, \mathbf{A}\mathbf{u}_k \rangle, \quad k = \overline{1, N}, \quad (6.106)$$

относительно коэффициентов a_n , совпадающей с (6.80) при $\mathbf{v}_k = \mathbf{A}\mathbf{u}_k \in R(A) \subset \mathcal{H}$.

Если однородное операторное уравнение $\mathbf{A}\mathbf{u} = \mathbf{0}$, где $\mathbf{0}$ нулевой элемент в $\mathcal{H}$, имеет лишь нулевое решение $\mathbf{u} = \mathbf{0}$, т.е. ядро линейного оператора $\mathbf{A}$ состоит только из нулевого элемента, то система $\{\mathbf{A}\mathbf{u}_n\}_N$ является линейно независимой в $\mathcal{H}$ при любом N , так как последовательность $\{\mathbf{u}_n\}$ образует в $D(A)$ счетный базис. В противном случае из элементов системы $\{\mathbf{A}\mathbf{u}_n\}_N$ можно было бы составить равную $\mathbf{0}$ нетривиальную линейную комбинацию

$$\sum_{n=1}^N b_n \mathbf{A}\mathbf{u}_n = \mathbf{A} \sum_{n=1}^N b_n \mathbf{u}_n = \mathbf{0}, \quad \sum_{n=1}^N |b_n| > 0,$$

но это означало бы, что существует ненулевое решение уравнения $Au = 0$, поскольку функции u_n счетного базиса линейно независимы и их нетривиальная линейная комбинация отлична от 0.

Скалярные произведения в левой части (6.106) являются элементами матрицы Грама системы функций Au_n , $n = \overline{1, N}$. Но для линейно независимой системы функций определитель этой матрицы отличен от нуля [III]. Поэтому СЛАУ (6.106) при любом N имеет единственное решение относительно неизвестных a_n , $n = \overline{1, N}$. Описанный способ нахождения коэффициентов a_n в (6.103) называют **методом наименьших квадратов** приближенного решения операторного уравнения $Au = f$. При этом (6.104) можно трактовать как условие наилучшей по норме аппроксимации элемента f линейной комбинацией элементов Au_n , $n = \overline{1, N}$. Если система $\{Au_n\}_N$ является ортогональной в $\mathcal{H}$, т.е. $\langle Au_n, Au_k \rangle = 0$ при $n \neq k$, то СЛАУ (6.106) имеет единственное решение $a_n = \frac{\langle f, Au_n \rangle}{\|Au_n\|^2}$, $n = \overline{1, N}$.

Пусть для оператора A область определения $D(A) \subset \mathcal{H} = L_2(V)$. Тогда вместо (6.106) можно написать

$$\sum_{n=1}^N a_n \int_V (Au_n) Au_k dV = \int_V f(Au_k) dV, \quad k = \overline{1, N}. \quad (6.107)$$

Если система $\{Au_n\}_N$ ортогональна в $L_2(V)$, то из (6.107) сразу получаем

$$a_n = \frac{1}{\|Au_n\|^2} \int_V f(Au_n) dV, \quad n = \overline{1, N}. \quad (6.108)$$

Пример 6.8. Применим метод наименьших квадратов к приближенному решению задачи (6.84). Это решение будем искать (как и в примере 6.5) на всюду плотном линейном многообразии $X \subset L_2[0, 1]$ четырежды непрерывно дифференцируемых на отрезке $[0, 1]$ функций, удовлетворяющих граничным

условиям (6.84). В качестве счетного базиса в X рассмотрим систему функций $u_n(\xi) = \sin n\pi\xi$, $n \in \mathbb{N}$. Отметим, что $u_n(\xi) \in D(A)$ образуют ортогональную в $L_2[0, 1]$ систему функций, так как при $n \neq k$

$$\langle u_n, u_k \rangle = \int_0^1 \sin n\pi\xi \sin k\pi\xi d\xi = 0.$$

Функции $Au_n(\xi) = (n\pi)^4 \sin n\pi\xi$, $n \in \mathbb{N}$, также ортогональны на отрезке $[0, 1]$. Кроме того, учитывая, что правая часть f дифференциального уравнения в задаче (6.84) равна единице, получим

$$\langle f, Au_n \rangle = \int_0^1 Au_n(\xi) d\xi = \left(\frac{n}{\pi}\right)^4 \int_0^1 \sin n\pi\xi d\xi = \left(\frac{n}{\pi}\right)^3 (1 - (-1)^n),$$

$$\|Au_n\|^2 = \int_0^1 (Au_n(\xi))^2 d\xi = \left(\frac{n}{\pi}\right)^8 \int_0^1 \sin^2 n\pi\xi d\xi = \frac{(n\pi)^8}{2}.$$

Отсюда следует, согласно (6.108), что $a_n = 0$ для всех четных $n = 2m$, $m \in \mathbb{N}$, а для всех нечетных $n = 2m - 1$ находим $a_n = a_{2m-1} = \frac{4}{(2m-1)^5 \pi^5}$. Так, $a_1 = \frac{4}{\pi^5} \approx 0,0130709$ и $a_3 = \frac{4}{243\pi^5} \approx 0,0000537$. Уже в первом приближении максимальный безразмерный прогиб $u_1^* = a_1 = \frac{4}{\pi^5}$ балки (см. рис. 6.3) менее чем на 0,4 % превышает его точное значение (6.87), равное $u^* = \frac{5}{384} \approx 0,0130208$, а в следующем приближении значение $u_3^* = a_1 - a_3 = \frac{968}{243\pi^5} \approx 0,0130162$ отличается от точного менее чем на 0,04 %. Нетрудно проверить, что решение уравнения (6.85) второго порядка с дифференциальным оператором $\tilde{A} = \frac{d^2}{d\xi^2}$ приведет к тому же результату.

Итак, приближенное решение задачи (6.84) можно представить в виде

$$\tilde{u}_N(\xi) = \sum_{m=1}^N \frac{4 \sin(2m-1)\pi\xi}{((2m-1)\pi)^5}. \quad (6.109)$$

Ясно, что получающийся из (6.109) при $N \rightarrow \infty$ ряд

$$\sum_{m=1}^{\infty} \frac{4 \sin(2m-1)\pi\xi}{((2m-1)\pi)^5} \quad (6.110)$$

сходится на отрезке $[0, 1]$, причем равномерно, так как ряд $\frac{4}{\pi^5} \sum_{m=1}^{\infty} \frac{1}{(2m-1)^5}$ является *мажорирующим* для ряда (6.110).

Выясним, к какой функции $v(\xi)$ сходится ряд (6.110). После почленного четырехкратного дифференцирования (6.110) получим ряд

$$\frac{4}{\pi} \sum_{m=1}^{\infty} \frac{\sin(2m-1)\pi\xi}{(2m-1)}, \quad (6.111)$$

который сходится к единице поточечно в интервале $(0, 1)$ и является рядом Фурье функции $\varphi(x) = 1$ в этом интервале [IX]. Поэтому функция $\varphi(x)$ является в этом интервале четвертой производной суммы ряда (6.110), т.е. функции $v(\xi)$. Таким образом, $\frac{d^4 v(\xi)}{d\xi^4} = 1$, $\xi \in (0, 1)$, и, следовательно, функция $v(\xi)$ удовлетворяет уравнению в (6.84). Но она удовлетворяет и всем граничным условиям задачи (6.84) при $\xi = 0$ и $\xi = 1$, что следует непосредственно из (6.110), а также из ряда, полученного двукратным дифференцированием (6.110). Отсюда заключаем, что функция $v(\xi)$ как сумма ряда (6.110) совпадает с точным решением (6.86) задачи (6.84), а $\tilde{u}_N(\xi)$ сходится к точному решению $u(\xi)$ при $N \rightarrow \infty$ равномерно на отрезке $[0, 1]$, а следовательно, и в $L_2[0, 1]$.

Оценим скорость сходимости по норме $\|u - \tilde{u}_N\|$. Так как N -й остаток функционального ряда (6.110)

$$u(\xi) - \tilde{u}_N(\xi) = \sum_{m=N+1}^{\infty} \frac{4 \sin(2m-1)\pi\xi}{((2m-1)\pi)^5}$$

сходится на отрезке $[0, 1]$ абсолютно и равномерно [IX], то сходится и квадрат этого остатка. Поэтому в силу ортогональности на отрезке $[0, 1]$ функций $u_n(\xi) = \sin n\pi\xi$ имеем

$$\begin{aligned} \|u - \tilde{u}_N\|^2 &= \int_0^1 \left(\sum_{m=N+1}^{\infty} \frac{4 \sin(2m-1)\pi\xi}{((2m-1)\pi)^5} \right)^2 d\xi = \\ &= \frac{8}{\pi^{10}} \sum_{m=N+1}^{\infty} \frac{1}{(2m-1)^{10}}. \end{aligned}$$

Если в дифференциальное уравнение $\frac{d^4 u(\xi)}{d\xi^4} = 1$ задачи (6.84) подставить приближенное решение $\tilde{u}_N$ из (6.109), то получим невязку

$$1 - A\tilde{u}_N = 1 - \frac{d^4 \tilde{u}_N(\xi)}{d\xi^4} = 1 - \sum_{m=1}^N \frac{4 \sin(2m-1)\pi\xi}{(2m-1)\pi}.$$

Тогда, принимая во внимание ортогональность на отрезке $[0, 1]$ функций $u_n(\xi) = \sin n\pi\xi$, находим [IX]

$$\begin{aligned} \|1 - A\tilde{u}_N\|^2 &= \int_0^1 \left(1 - \sum_{m=1}^N \frac{4 \sin(2m-1)\pi\xi}{(2m-1)\pi} \right)^2 d\xi = \\ &= 1 - 8 \sum_{m=1}^N \int_0^1 \frac{\sin(2m-1)\pi\xi}{(2m-1)\pi} d\xi + \int_0^1 \left(\sum_{m=1}^N \frac{4 \sin(2m-1)\pi\xi}{(2m-1)\pi} \right)^2 d\xi = \\ &= 1 - \sum_{m=1}^N \frac{16}{(2m-1)^2 \pi^2} + \sum_{m=1}^N \frac{8}{(2m-1)^2 \pi^2} = 1 - \sum_{m=1}^N \frac{8}{(2m-1)^2 \pi^2}. \end{aligned}$$

Учитывая, что*

$$\sum_{m=1}^{\infty} \frac{1}{(2m-1)^2} = \frac{\pi^2}{8},$$

в итоге получаем

$$\|1 - A\tilde{u}_N\|^2 = \frac{8}{\pi^2} \sum_{m=N+1}^{\infty} \frac{1}{(2m-1)^2}.$$

Таким образом, для данной задачи скорость сходимости приближенного решения к точному при $N \rightarrow \infty$ существенно выше стремления к нулю нормы невязки. #

Ясно, что пример 6.8 носит иллюстративный характер, поскольку задачу (6.84) легко решить точно последовательным интегрированием (см. пример 6.5). В общем случае установить сходимость к точному решению приближенного решения, полученного методом наименьших квадратов, довольно сложно.

Теорема 6.7. Пусть A — линейный непрерывный оператор в гильбертовом пространстве $\mathcal{H}$, область определения $D(A)$ и область значений $R(A)$ которого всюду плотны в $\mathcal{H}$, $\{u_n\}$ — счетный базис в $D(A)$. Тогда построенное по методу наименьших квадратов приближенное решение (6.103) сходится в $\mathcal{H}$ при $N \rightarrow \infty$ к классическому решению операторного уравнения $Au = f$, $f \in R(A)$, если система $\{Au_n\}$ является счетным базисом в $R(A)$ и для некоторого $\tau > 0$ выполнено условие

$$\|Au\| \geq \tau \|u\|, \quad u \in D(A). \quad (6.112)$$

◀ При выполнении условия (6.112), согласно теореме 4.16, существует ограниченный обратный оператор A^{-1} , т.е. операторное уравнение $Au = f$, $f \in R(A)$, имеет единственное решение $u_0 = A^{-1}f$. В силу свойств счетного базиса $\{Au_n\}$ система $\{Au_n\}_N$ является линейно независимой при любом $N \in \mathbb{N}$,

*См.: Градштейн И.С., Рыжик И.М.

т.е. СЛАУ (6.106) при любом N имеет единственное решение относительно коэффициентов a_n , $n = \overline{1, N}$. Кроме того, для произвольного $f \in R(A)$ при заданном $\varepsilon > 0$ можно найти такое $N_* \in \mathbb{N}$ и такие коэффициенты $\alpha_n \in \mathbb{R}$, $n = \overline{1, N_*}$, что справедливо неравенство

$$\left\| f - \sum_{n=1}^{N_*} \alpha_n A u_n \right\| < \tau \varepsilon. \quad (6.113)$$

Если в (6.113) α_n заменить соответственно на коэффициенты a_n , $n = \overline{1, N_*}$, найденные из условия (6.104), то неравенство останется верным, поскольку в этом случае его левая часть достигает минимума. Но и при $N > N_*$ будет выполнено неравенство

$$\left\| f - \sum_{n=1}^N a_n A u_n \right\| < \tau \varepsilon, \quad (6.114)$$

если коэффициенты a_n , $n = \overline{1, N}$, найдены из условия (6.104). Действительно, при $a_n = \alpha_n$, $n = \overline{1, N_*}$, и $a_n = 0$, $n = \overline{N_* + 1, N}$, левые части (6.113) и (6.114) совпадут, а при нахождении коэффициентов a_n из условия (6.104) левая часть (6.114) может только уменьшиться.

Таким образом, с учетом (6.79), (6.114) и равенства $f = A u_0$ получаем

$$\|A u_0 - A \tilde{u}_N\| = \|A(u_0 - \tilde{u}_N)\| < \tau \varepsilon.$$

Отсюда, учитывая (6.112), имеем $\|u_0 - \tilde{u}_N\| < \varepsilon$. Следовательно, $\tilde{u}_N \rightarrow u_0$ при $N \rightarrow \infty$. ►

Отметим, что при выполнении условий теоремы 6.7 для построенного по методу наименьших квадратов приближенного решения $\tilde{u}_N$ (6.103) имеем $A \tilde{u}_N \rightarrow f$ при $N \rightarrow \infty$, т.е. стремится к нулю норма $\|A \tilde{u}_N - f\|$ невязки операторного уравнения, возникающей при подстановке $\tilde{u}_N$ в уравнение $A u = f$. Это

позволяет при помощи вытекающего из (6.112) неравенства

$$\|\tilde{u}_N - u_0\| \leq \frac{1}{\tau} \|A\tilde{u}_N - f\| \quad (6.115)$$

оценивать по норме погрешность приближенного решения. В заключение заметим, что (6.112) верно для *положительно определенного оператора* A . В самом деле, в этом случае в соответствии с (5.2) и (5.11) имеем $\|Au\| \|u\| \geq (Au, u) \geq \gamma^2 \|u\|^2$, откуда при $\gamma^2 = \tau$ следует (6.112).

6.8. Методы Бубнова — Галеркина и Ритца

Рассмотрим частный случай *метода ортогональных проекций* приближенного решения *операторного уравнения* $Au = f$. Пусть области определения $D(A)$ и значений $R(A)$ оператора A являются *всюду плотными подмножествами* *сепарабельного гильбертова пространства* $\mathcal{H}$, а система $\{u_k\}$ — *счетным базисом* в $D(A)$ и $R(A)$. Приближенное решение $\tilde{u}_N$ вида (6.79) уравнения $Au = f$, согласно методу ортогональных проекций для случая, когда *проекционные функции* совпадают с *базисными*, т.е. $v_k = u_k$, $k = \overline{1, N}$, находим, решая получаемую из (6.80) систему линейных алгебраических уравнений (СЛАУ)

$$\sum_{n=1}^N a_n \langle Au_n, u_k \rangle = \langle f, u_k \rangle, \quad k = \overline{1, N}, \quad (6.116)$$

относительно коэффициентов a_n .

Такую процедуру нахождения приближенного решения операторного уравнения $Au = f$ называют *методом Бубнова* — Галеркина*. Поскольку метод ортогональных проекций является, в свою очередь, одним из вариантов *проекционного метода*, то условия существования решения СЛАУ (6.116) и сходимости приближенного решения следуют из теорем 6.11,

*И.Г. Бубнов (1872–1919) — русский инженер.

6.12 и замечаний 6.3, 6.4 (см. Д.6.1). Отметим, что метод Бубнова — Галеркина можно использовать, не накладывая на оператор A существенных ограничений: он может не быть положительно определенным оператором, симметрическим и даже линейным.

Рассмотрим более подробно практически важный случай, когда оператор A является положительно определенным. В этом случае элементы u_n счетного базиса $\{u_n\}$, используемого при построении приближенного решения $\tilde{u}_N$ вида (6.103), могут и не принадлежать области $D(A)$ определения оператора A , т.е. $u_n \notin D(A)$, $n = \overline{1, N}$. Это существенно расширяет возможности метода Бубнова — Галеркина по сравнению с методами коллокации и наименьших квадратов.

Пусть $\mathcal{H}_A$ — энергетическое пространство, которое является пополнением нормированного пространства $D(A)$ по энергетической норме

$$\|u\|_A = \sqrt{\langle u, u \rangle_A},$$

индуцированной энергетическим скалярным произведением

$$\langle u, v \rangle_A = \langle Au, v \rangle.$$

Теперь приближенное решение операторного уравнения $Au = f$ с положительно определенным оператором A вместо (6.79) можно искать в виде

$$\tilde{u}_N = \sum_{n=1}^N a_n u_n, \quad u_n \in \mathcal{H}_A, \quad (6.117)$$

где функции u_n , $n \in \mathbb{N}$, образуют счетный базис в энергетическом пространстве $\mathcal{H}_A$. При этом коэффициенты a_n , $n = \overline{1, N}$, находят из решения СЛАУ (6.116), которая принимает вид

$$\sum_{n=1}^N a_n \langle u_n, u_k \rangle_A = \langle f, u_k \rangle, \quad k = \overline{1, N}. \quad (6.118)$$

Элементы $\langle u_n, u_k \rangle_A$, $k, n = \overline{1, N}$, в (6.118) составляют *матрицу Грама* системы функций $\{u_n\}_N$, которая в силу свойств счетного базиса $\{u_n\}$ является линейно независимой в энергетическом пространстве $\mathcal{H}_A$. Следовательно, определитель этой матрицы отличен от нуля, а СЛАУ (6.118) имеет единственное решение относительно коэффициентов a_n , $n = \overline{1, N}$.

Рассмотрим еще один метод приближенного решения уравнения $Au = f$ с положительно определенным оператором A , в конечном итоге опять приводящий к решению СЛАУ (6.118). *Обобщенное решение* u_* этого уравнения единственно и минимизирует *квадратичный функционал* $J[u] = \langle Au, u \rangle - 2\langle f, u \rangle$, который называют также *функционалом энергии*. Если элементы *минимизирующей последовательности* $\{\tilde{u}_N\}$ искать в виде (6.117), то при нахождении коэффициентов a_n , $n = \overline{1, N}$, придем к совпадающей с (6.118) СЛАУ. Такую процедуру нахождения коэффициентов принято называть *методом Ритца**.

Таким образом, методы Ритца и Бубнова — Галеркина приближенного решения уравнения $Au = f$ в случае положительно определенного оператора A приводят к решению одной и той же СЛАУ (6.118). Однако метод Бубнова — Галеркина в отличие от метода Ритца можно применять и для решения уравнения $Au = f$ с произвольным оператором A .

Пусть $\{u_n\}$ — базис в $\mathcal{H}_A$, ортонормированный относительно энергетического скалярного произведения. В этом случае приближенное решение (6.117) с коэффициентами $a_n = \langle f, w_n \rangle_A$, полученными методом Ритца, будет частичной суммой ряда Фурье [XV]

$$u_* = \sum_{n=1}^{\infty} \alpha_n u_n, \quad \alpha_n = \langle f, w_n \rangle_A, \quad u_n \in \mathcal{H}_A, \quad (6.119)$$

обобщенного решения u_* уравнения $Au = f$. Это означает, что в случае ортонормированного в $\mathcal{H}_A$ базиса $\{u_n\}$ приближенное

*В. Ритц (1878–1909) — немецкий физик и математик.

решение (6.117) уравнения $Au = f$ с положительно определенным оператором A сходится в $\mathcal{H}_A$ к обобщенному решению (6.119) этого уравнения по энергетической норме, причем в силу свойств ряда Фурье [IX]

$$\|u_* - \tilde{u}_N\|_A^2 = \sum_{n=N+1}^{\infty} \alpha_n^2. \quad (6.120)$$

Но для положительно определенного оператора A обобщенное решение u_* уравнения $Au = f$ совпадает со *слабым решением* этого уравнения, удовлетворяющим равенству $\langle u_*, v \rangle_A = \langle f, v \rangle$ для любого $v \in \mathcal{H}_A$. Это следует из совпадения СЛАУ для нахождения коэффициентов приближенного решения методами Ритца и Бубнова — Галеркина. Поэтому приближенное решение $\tilde{u}_N$ сходится при $N \rightarrow \infty$ и к слабому решению $u_* \in \mathcal{H}_A$ уравнения $Au = f$ по энергетической норме. Поскольку для энергетической нормы $\|u\|_A \geq \gamma \|u\|$, то $\tilde{u}_N$ сходится к u_* и по норме $\|\cdot\|$ в гильбертовом пространстве.

Замечание 6.1. Выбор в (6.117) и (6.118) в качестве базисных функций (и совпадающих с ними проекционных) собственных элементов положительно определенного оператора A обеспечивает сходимость не только приближенного решения к слабому решению уравнения $Au = f$, но и сходимость к нулю нормы *невязки* этого *операторного уравнения*. Действительно, пусть u_n , $n \in \mathbb{N}$, — собственные элементы оператора A , соответствующие собственным значениям $\lambda_n > 0$, образующие в $\mathcal{H}$ ортонормированный базис. Если несколько различных линейно независимых собственных элементов соответствуют одному и тому же собственному значению, то положим, что в этом случае существует столько же равных между собой собственных значений, т.е. различными значениями индекса n обозначим различные собственные элементы. Умножая скалярно равенство $Au_n = \lambda_n u_n$ на u_k и учитывая, что $\langle Au_n, u_k \rangle = \langle u_n, u_k \rangle_A$, запишем

$$\langle Au_n, u_k \rangle = \lambda_n \langle u_n, u_k \rangle = \lambda_n \delta_{nk} = \langle u_n, u_k \rangle_A,$$

где $\delta_{nk} = 1$ при $n = k$ и $\delta_{nk} = 0$ при $n \neq k$. Отсюда следует, что $\langle u_n, u_n \rangle_A = \lambda_n$ и $\langle u_n, u_k \rangle_A = 0$ при $k \neq n$, т.е. в соответствии с (6.118) $a_n = \frac{\langle f, u_n \rangle}{\lambda_n}$, так что вместо (6.117) будем иметь

$$\tilde{u}_N = \sum_{n=1}^N \frac{\langle f, u_n \rangle}{\lambda_n} u_n. \quad (6.121)$$

Подвергая (6.121) действию оператора A , получаем

$$A\tilde{u}_N = \sum_{n=1}^N \frac{\langle f, u_n \rangle}{\lambda_n} A u_n = \sum_{n=1}^N \langle f, u_n \rangle u_n. \quad (6.122)$$

Правая часть (6.122) является конечной суммой разложения элемента f в ряд Фурье по ортонормированной системе $\{u_n\}$ функций u_n , $n \in \mathbb{N}$. Поэтому $A\tilde{u}_N \rightarrow f$ при $N \rightarrow \infty$.

Пример 6.9. Используем метод Бубнова — Галеркина для приближенного решения краевой задачи (6.84), уравнение которой содержит положительно определенный оператор $A = \frac{d^4}{dx^4}$. Решение будем искать во всюду плотном подмножестве $X \subset C L_2[0, 1]$ четырежды непрерывно дифференцируемых на отрезке $[0, 1]$ функций, которое является *линейным многообразием*. Как и в примере 6.5, в качестве счетного базиса в X выберем систему функций $u_n(x) = \sin n\pi x \in D(A)$, $n \in \mathbb{N}$, удовлетворяющих граничным условиям в (6.84).

Так как функции $u_n(x) \in D(A)$ ортогональны на отрезке $[0, 1]$ (см. пример 6.8), то при $n \neq k$

$$\langle A u_n, u_k \rangle = \int_0^1 (A u_n(x)) u_k(x) dx = (n\pi)^4 \int_0^1 \sin n\pi x \sin k\pi x dx = 0.$$

В случае $n = k$ находим

$$\langle A u_n, u_n \rangle = \int_0^1 (A u_n(x)) u_n(x) dx = (n\pi)^4 \int_0^1 \sin^2 n\pi x dx = \frac{(n\pi)^4}{2}.$$

Учитывая, что в задаче (6.84) правая часть f дифференциального уравнения равна единице, для нечетных $k = 2m - 1$ имеем

$$\langle f, u_k \rangle = \int_0^1 \sin k\pi\xi d\xi = \frac{2}{k\pi}$$

и в соответствии с (6.116) для нечетных $n = 2m - 1$ получаем $a_n = a_{2m-1} = \frac{4}{(2m-1)^5\pi^5}$ и $a_n = a_{2m} = 0$ для всех четных $n = 2m$. Это означает, что для рассматриваемой задачи метод Бубнова — Галеркина и метод наименьших квадратов приводят к одинаковому результату. Несложно проверить, что к тому же результату приведет и решение уравнения (6.85) с граничными условиями $u(0) = u(1) = 0$. #

Метод Бубнова — Галеркина широко применяют и для приближенного решения тех операторных уравнений, в которых оператор не является положительно определенным.

Пример 6.10. Применим метод Бубнова — Галеркина как вариант метода взвешенных невязок к приближенному решению задачи, рассмотренной в примере 6.7. Приближенное решение сначала будем искать в виде (6.93), а в качестве *весовой функции* возьмем входящую в (6.93) функцию $v_1(x) = 1 - (x/h)^2$. Эта функция при использовании метода Бубнова — Галеркина выполняет одновременно роль и базисной и проекционной функций.

Подставив (6.93) в (6.91), приравняем нулю взвешенную невязку

$$\begin{aligned} \int_0^h \left(\frac{\partial \tilde{T}}{\partial t} - a \frac{\partial^2 \tilde{T}}{\partial x^2} \right) v_1(x) dx &= - \int_0^h \left(1 - \frac{x^2}{h^2} \right)^2 \frac{dD(t)}{dt} dx - \\ &- \int_0^h a D(t) \left(1 - \frac{x^2}{h^2} \right) \frac{2}{h^2} dx = - \frac{8h}{15} \frac{dD(t)}{dt} - \frac{4a}{3h} D(t) = 0. \end{aligned}$$

Отсюда следует обыкновенное дифференциальное уравнение

$$\frac{D(t)}{dt} + \frac{5a}{2h^2} D(t) = 0, \quad (6.123)$$

решение которого представим в виде

$$D(t) = D(0) \exp\left(-\frac{5at}{2h^2}\right) = D(0) \exp(-2,5Fo), \quad (6.124)$$

где $Fo = \frac{at}{h^2}$ — число Фурье (безразмерное время), а $D(0)$ — начальное значение функции $D(t)$, которое найдем из равенства нулю взвешенной невязки в начальном условии задачи:

$$\begin{aligned} \int_0^h (\tilde{T}(0, x) - T_0) v_1(x) dx = \\ = \int_0^h \left(T^* - D(0) \left(1 - \frac{x^2}{h^2} \right) - T_0 \right) \left(1 - \frac{x^2}{h^2} \right) dx = 0. \end{aligned}$$

После вычислений получим $D(0) = \frac{5}{4}(T^* - T_0)$ и, подставляя (6.124) в (6.93), запишем

$$\tilde{\Theta}(Fo, \xi) = \frac{T^* - \tilde{T}(t, x)}{T^* - T_0} = 5 \frac{1 - \xi^2}{4} \exp(-2,5Fo), \quad \xi = \frac{x}{h}. \quad (6.125)$$

Результаты расчета по (6.125) при $\xi = 0$ приведены в табл. 6.1.

6.9. Задачи на собственные значения

Рассмотрим *однородное операторное уравнение* вида (5.41)

$$Au - \lambda Bu = 0, \quad u, 0 \in D(A) \subset \mathcal{H}, \quad (6.126)$$

где A и B — линейные операторы, действующие в *гильбертовом пространстве* $\mathcal{H}$, причем $D(A) \subset D(B)$ и $D(A)$ — *всюду*

плотное в $\mathcal{H}$ подмножество. Напомним, что число λ называют собственным значением этого операторного уравнения, если (6.126) при этом λ имеет решение, отличное от тривиального решения $u = 0$. Выше (см. 5.5) показано, что если A — симметрический оператор, а B — положительно определенный, то при решении задачи на собственные значения в (6.126) без потери общности можно принять оператор A также положительно определенным. Как и выше (см. 5.5) предположим, что все собственные значения операторного уравнения (6.126) простые, образуют счетное множество и их можно представить в виде элементов возрастающей последовательности $\{\lambda_m\}$. Тогда каждому из этих собственных значений λ_m отвечают собственный элемент w_n , $\|w_n\| = 1$, и одномерное собственное подпространство S_m , совпадающее с линейной оболочкой этого элемента.

В прикладных исследованиях задачи на собственные значения уравнения (6.126) возникают, например, при анализе колебаний различных динамических систем. В этом случае собственные значения пропорциональны квадрату частот собственных колебаний. При исследовании условий перехода системы из одного состояния в другое собственные значения характеризуют уровень внешних воздействий на систему, при котором такой переход возможен. В частности, собственные значения могут иметь смысл критических нагрузок, вызывающих потерю устойчивости равновесия или движения системы.

Приближенное решение задачи на собственные значения уравнения (6.126) можно найти с помощью метода Бубнова — Галеркина. Искомый собственный элемент приближенно представим в виде

$$\tilde{u}_N = \sum_{n=1}^N a_n u_n, \quad a_n \in \mathbb{R}, \quad u_n \in D(A), \quad N \in \mathbb{N}, \quad (6.127)$$

где $u_n \in D(A)$ — элементы последовательности $\{u_n\}$, образующей в $D(A)$ счетный базис. Подставляя (6.127) в (6.126) и

используя (6.76) при $v_k = u_k$, $k = \overline{1, N}$, приходим к однородной системе линейных относительно коэффициентов a_n алгебраических уравнений (СЛАУ)

$$\sum_{n=1}^N a_n \langle Au_n - \lambda Bu_n, u_k \rangle = 0, \quad k = \overline{1, N}, \quad (6.128)$$

но содержащей неизвестное число λ .

Эта СЛАУ имеет нетривиальное решение относительно коэффициентов a_n при условии $\det G_N = 0$, где G_N — симметрическая матрица порядка N с элементами $\langle Au_n - \lambda Bu_n, u_k \rangle$, $k, n = \overline{1, N}$. Каждое значение $\lambda_m^{(N)}$, $m = \overline{1, N}$, удовлетворяющее уравнению $\det G_N = 0$, примем в качестве приближения к собственному значению λ_m уравнения (6.126) при условии, что $\lambda_m^{(N)}$ (как и λ_m) занумерованы в порядке возрастания m . Ясно, что при фиксированном N можно найти приближенные значения лишь для N собственных значений уравнения (6.126).

Замечание 6.2. Можно показать*, что при любом $m \leq N$ верно неравенство $\lambda_m^{(N)} \geq \lambda_m$ и при возрастании N значения $\lambda_m^{(N)}$ для фиксированных m не возрастают, причем $\lambda_m^{(N)} \rightarrow \lambda_m$ и $\|w_m^{(N)} - w_m\|_{\mathcal{H}} \rightarrow 0$ при $N \rightarrow \infty$, где в соответствии с (6.127)

$$w_m^{(N)} = \sum_{n=1}^N a_n^{(m)} u_n, \quad u_n \in D(A),$$

и $a_n^{(m)}$ — координаты единичного вектора $w_m^{(N)}$, удовлетворяющего СЛАУ (6.128) при $\lambda = \lambda_m^{(N)}$. Отметим, что в частном случае $B = I$, где I — тождественный оператор в $\mathcal{H}$, числа $\lambda_m^{(N)}$, $m = \overline{1, N}$, являются собственными значениями симметрической матрицы $\tilde{A}$ с элементами $\langle Au_n, u_k \rangle$, $k, n = \overline{1, N}$, и служат

*См. Гавурин М.К.

приближением к собственным значениям симметрического оператора A [XV]. При этом $w_n^{(N)}$ будет единичным собственным вектором матрицы $\tilde{A}$, отвечающим ее собственному значению $\lambda_m^{(N)}$.

Таким образом, из решения уравнения $\det G = 0$ можно получить значения $\lambda_m^{(N)}$, $m = \overline{1, N}$, дающие оценки сверху для собственных значений λ_m уравнения (6.126). В ряде прикладных задач не менее важным является получение для λ_m оценок снизу (в частности, для выяснения возможной погрешности приближенного решения вариационной задачи [XV]). Такую оценку для λ_1 несложно получить, если известно или легко вычисляемо собственное значение λ'_1 операторного уравнения $A'u = \lambda' B'u$, где A' — симметрический оператор с областью определения $D(A') \subset D(A)$, B' — положительно определенный оператор, для которого $D(B') \supset D(B)$, причем

$$\frac{\langle A'u, u \rangle}{\langle B'u, u \rangle} \leq \frac{\langle Au, u \rangle}{\langle Bu, u \rangle}, \quad u \in D(A'). \quad (6.129)$$

В этом случае, согласно теореме 5.6 и замечанию 5.5, имеем $\lambda'_1 \leq \lambda_1$. Ясно, что аналогичный прием можно использовать с целью получения для λ_1 оценки сверху. Для этого необходимо располагать собственным значением λ''_1 операторного уравнения $A''u = \lambda'' B''u$, где A'' — симметрический оператор, для которого $D(A'') \supset D(A)$, а B'' — положительно определенный оператор с областью определения $D(B'') \subset D(B)$, причем $D(A'') \subset D(B'')$ и

$$\frac{\langle A''u, u \rangle}{\langle B''u, u \rangle} \geq \frac{\langle Au, u \rangle}{\langle Bu, u \rangle}, \quad u \in D(A). \quad (6.130)$$

Тогда, согласно замечанию 5.5 и теореме 5.6, получим, что $\lambda''_1 \geq \lambda_1$.

Пример 6.11. Из условия равновесия шарнирно опертого упругого стержня длиной l , сжатого вдоль оси Ox силой

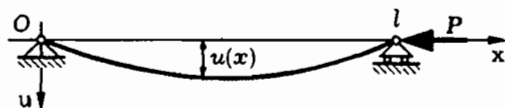


Рис. 6.5

P (рис. 6.5), следует обыкновенное дифференциальное уравнение (ОДУ)

$$EI_*(x) \frac{d^2 u(x)}{dx^2} + Pu(x) = 0, \quad x \in (0, l), \quad (6.131)$$

с граничными условиями $u(0) = u(l) = 0$, где E — модуль упругости материала стержня; $I_*(x) = \frac{\pi}{4} r^4(x)$ — момент инерции кругового поперечного сечения стержня с зависящим от координаты x радиусом $r(x)$; $u(x)$ — отклонение точек оси стержня от оси Ox (при отсутствии силы P ось стержня совпадает с осью Ox).

Примем $r(x) = a\sqrt{1 + \frac{x}{l}}$ и, обозначив $\xi = \frac{x}{l}$ и $\lambda = \frac{Pl^2}{\pi Ea^4}$, от (6.131) перейдем к краевой задаче

$$\begin{cases} -\frac{d^2 u(\xi)}{d\xi^2} - \lambda \frac{4u(\xi)}{(1+\xi)^2} = 0, & \xi \in (0, 1); \\ u(0) = u(1) = 0 \end{cases} \quad (6.132)$$

для однородного ОДУ с однородными граничными условиями. Пусть $A = -\frac{d^2}{d\xi^2}$ и $Bu = \frac{4u}{(1+\xi)^2} > 0$. Оператор A является положительно определенным на линейном многообразии $D(A)$ дважды непрерывно дифференцируемых на отрезке $[0, 1]$ функций, обращающихся на концах этого отрезка в нуль (см. пример 5.10). Оператор B также положительно определенный, но на множестве $D(B)$ функций, непрерывных на отрезке $[0, 1]$, поскольку $\langle Bu, u \rangle = 4 \frac{\|u\|^2}{(1+\xi)^2} > \|u\|^2$. Таким образом, $D(A) \subset D(B)$, и задача (6.132) является задачей на собственные значения операторного уравнения (6.126).

При заданных граничных условиях наряду с тривиальным решением $u(\xi) = 0$, $\xi \in [0, 1]$, краевая задача (6.132) может иметь решения $u_n(\xi) \neq 0$, соответствующие собственным значениям λ_n . В данном случае важно оценить наименьшее собственное значение λ_1 , пропорциональное наименьшей сжимающей силе, вызывающей потерю устойчивости прямолинейной формы равновесия стержня.

Для оценки значения λ_1 сверху используем функцию $\tilde{u}(\xi) = \xi(1 - \xi)$, дважды непрерывно дифференцируемую на отрезке $[0, 1]$ и удовлетворяющую граничным условиям $u(0) = u(1) = 0$, и вычислим

$$\langle A\tilde{u}, \tilde{u} \rangle = - \int_0^1 \left(\frac{d^2 \tilde{u}(\xi)}{d\xi^2} \right) \tilde{u}(\xi) d\xi = 2 \int_0^1 \xi(1 - \xi) d\xi = \frac{1}{3},$$

$$\langle B\tilde{u}, \tilde{u} \rangle = 4 \int_0^1 \frac{\xi^2(1 - \xi)^2}{(1 + \xi)^2} d\xi = 4 \frac{25 - 36 \ln 2}{3}.$$

Тогда, согласно теореме 5.6, получим

$$\lambda_1 \leq \frac{\langle A\tilde{u}, \tilde{u} \rangle}{\langle B\tilde{u}, \tilde{u} \rangle} = \frac{1}{4(25 - 36 \ln 2)} \approx 5,3532. \quad (6.133)$$

Оценку снизу значения λ_1 можно получить, если рассмотреть стержень с круговым поперечным сечением постоянного радиуса $R' = a \leq r(x)$. Для такого стержня вместо задачи (6.132) имеем

$$\begin{cases} -\frac{d^2 u(\xi)}{d\xi^2} - 4\lambda u(\xi) = 0, & \xi \in (0, 1); \\ u(0) = u(1) = 0. \end{cases} \quad (6.134)$$

Несложно проверить, что собственными функциями задачи (6.134) будут $u_n(x) = \sin n\pi\xi$. Этим функциям соответствуют

собственные значения

$$\lambda'_n = \frac{(n\pi)^2}{4}, \quad (6.135)$$

причем наименьшее из них (при $n = 1$) равно $\lambda'_1 = \frac{\pi^2}{4} \approx 2,4674$ и соответствует значению *эйлеровой* (или первой критической) *силы* $P_{кр} = \frac{\lambda'_1 \pi E a^4}{l^2} = \frac{\pi^3 E a^4}{4 l^2}$ [VIII], вызывающей потерю устойчивости прямолинейной формы равновесия стержня с постоянным круговым поперечным сечением (задача нахождения этой силы была впервые рассмотрена Л. Эйлером в 1744 году).

Из сравнения (6.134) с (6.132) следует, что $A' = A = -\frac{d^2}{d\xi^2}$ и $B' = 4 > 0$, причем $\langle A'u, u \rangle = \langle Au, u \rangle$ и $\langle B'u, u \rangle \geq \langle Bu, u \rangle$ для любого элемента $u \in D(A)$, т.е. выполнено неравенство (6.129). Тогда с учетом (6.133) получаем двустороннюю оценку

$$\lambda'_1 = \frac{\pi^2}{4} \leq \lambda_1 \leq \frac{1}{4(25 - 36 \ln 2)} \approx 5,3532. \quad (6.136)$$

Оценка (6.136) является довольно грубой, поскольку верхняя граница превышает нижнюю более чем в два раза. Отметим, что для стержня с круговым поперечным сечением постоянного радиуса $R'' = \sqrt{2}a \geq r(x)$ получим $A'' = A = -\frac{d^2}{d\xi^2}$ и $B'' = 1 > 0$, причем $\langle A''u, u \rangle = \langle Au, u \rangle$ и $\langle B''u, u \rangle \leq \langle Bu, u \rangle$, $u \in D(A)$, т.е. будет выполнено неравенство (6.130). Но для такого стержня $\lambda''_1 = \pi^2 \approx 9,8696$, что почти в два раза выше верхней границы в (6.136). #

Для улучшения верхней границы собственных значений следует увеличивать число N базисных функций при приближенном представлении собственных функций (см. замечание 6.2). Покажем, как можно улучшить нижнюю оценку. Сначала рассмотрим случай, когда оператор B в (6.126) является тождественным, т.е. $B = I$.

Теорема 6.8. Пусть A — симметрический неотрицательный оператор с областью определения $D(A) \subset \mathcal{H}$ в гильбертовом пространстве $\mathcal{H}$ и для некоторого $\nu > 0$ отрезок $[0, \nu]$ не содержит точек спектра этого оператора. Тогда оператор $A_\nu = A - \nu I$ является положительным, т.е. $\langle A_\nu u, u \rangle > 0$ для любого ненулевого элемента $u \in D(A)$.

◀ Согласно условию теоремы, любое $\lambda \in [0, \nu]$ является регулярным значением оператора A и поэтому существует обратный оператор $R_\lambda = A_\lambda^{-1} = (A - \lambda I)^{-1}$. Из области определения $D(A)$ оператора A выберем любой ненулевой элемент $u \neq 0$ и положим $g = A_\nu u$. Тогда $g \neq 0$, поскольку в противном случае элемент u был бы собственным для оператора A , а ν — собственным значением этого оператора.

Пусть $v = v(\lambda) = R_\lambda g$ — решение операторного уравнения $A_\lambda v = g$. Введем вспомогательную функцию

$$\varphi(\lambda) = \langle g, v \rangle = \langle A_\lambda v, v \rangle.$$

В частном случае $\lambda = \nu$ имеем $v = u$ и $\varphi(\nu) = \langle A_\nu u, u \rangle$. Таким образом, необходимо доказать, что $\varphi(\nu) > 0$ для любого ненулевого элемента $u \in D(A)$.

Обозначим $v_1 = v(\lambda_1) = R_{\lambda_1} g$ и $v_2 = v(\lambda_2) = R_{\lambda_2} g$ — решения операторного уравнения $A_\lambda v = g$ для двух значений $\lambda_1, \lambda_2 \in [0, \nu]$. Тогда, учитывая, что оператор A симметрический, получим

$$\begin{aligned} \varphi(\lambda_2) - \varphi(\lambda_1) &= \langle g, v_2 \rangle - \langle g, v_1 \rangle = \langle (A - \lambda_1 I) v_1, v_2 \rangle - \\ &- \langle (A - \lambda_2 I) v_2, v_1 \rangle = \langle A v_1, v_2 \rangle - \lambda_1 \langle v_1, v_2 \rangle - \\ &- \langle A v_2, v_1 \rangle + \lambda_2 \langle v_2, v_1 \rangle = (\lambda_2 - \lambda_1) \langle v_2, v_1 \rangle. \end{aligned} \quad (6.137)$$

Так как $g = (A - \lambda_2 I) v_2 = (A - \lambda_1 I) v_1$, то

$$(A - \lambda_1 I)(v_2 - v_1) = (A - \lambda_1 I) v_2 - (A - \lambda_2 I) v_2 = (\lambda_2 - \lambda_1) v_2,$$

или $v_2 - v_1 = (\lambda_2 - \lambda_1)R_{\lambda_1}v_2$. С учетом неравенства треугольника имеем

$$\begin{aligned}\|v_2 - v_1\| &\leq |\lambda_2 - \lambda_1| \|R_{\lambda_1}\| \|v_2\| = \\ &= |\lambda_2 - \lambda_1| \|R_{\lambda_1}\| \|v_2 - v_1 + v_1\| \leq \\ &\leq |\lambda_2 - \lambda_1| \|R_{\lambda_1}\| (\|v_2 - v_1\| + \|v_1\|). \quad (6.138)\end{aligned}$$

Значение λ_1 является регулярным для оператора A , и поэтому оператор R_{λ_1} ограничен, так что при $\lambda_2 \rightarrow \lambda_1$ справедливо неравенство $|\lambda_2 - \lambda_1| \|R_{\lambda_1}\| < 1/2$. Тогда из (6.138) следует, что

$$\|v_2 - v_1\| \leq 2|\lambda_2 - \lambda_1| \|R_{\lambda_1}\| \|v_1\|.$$

Отсюда получаем, что $v_2 \rightarrow v_1$ при $\lambda_2 \rightarrow \lambda_1$, и поэтому $\langle v_2, v_1 \rangle \rightarrow \|v_1\|^2$ при $\lambda_2 \rightarrow \lambda_1$. Заменяя в (6.137) λ_1 на λ и v_1 на v , в соответствии с определением производной находим

$$\varphi'(\lambda) = \lim_{\lambda_2 \rightarrow \lambda} \frac{\varphi(\lambda_2) - \varphi(\lambda)}{\lambda_2 - \lambda} = \lim_{\lambda_2 \rightarrow \lambda} \langle v_2, v \rangle = \|v\|^2.$$

Так как $g \neq 0$, то и $v \neq 0$, а значит, $\varphi'(\lambda) > 0$ при $\lambda \in [0, \nu]$ и функция $\varphi(\lambda)$ возрастает на отрезке $[0, \nu]$. Для неотрицательного оператора A имеем $\varphi(0) = \langle A_0 v, v \rangle = \langle A v, v \rangle \geq 0$ для любого элемента $v \in D(A)$. Следовательно, $\varphi(\nu) > 0$ для любого ненулевого элемента $u \in D(A)$. ►

Теорема 6.9. Пусть A — симметрический оператор, действующий в гильбертовом пространстве $\mathcal{H}$, и отрезок $[a, b]$ не содержит точек спектра этого оператора. Тогда оператор $A_a A_b = (A - aI)(A - bI)$ является положительным.

◄ Положим $m = \frac{a+b}{2}$, $\mu = \frac{b-a}{2}$ и введем оператор $W_\lambda = A_{m+\lambda} = A - (m + \lambda)I$. Значение $\lambda \in [-\mu, \mu]$ является регулярным для оператора $W_0 = A_m$, и поэтому оператор $W_\lambda = W_0 - \lambda I$, а также оператор $W_\lambda W_{-\lambda} = W_0^2 - \lambda^2 I$ имеют ограниченные обратные

операторы W_λ^{-1} и $(W_\lambda W_{-\lambda})^{-1}$ соответственно. Это означает, что отрезок $[0, \mu^2]$ не содержит точек спектра оператора W_0^2 , который, являясь квадратом симметрического оператора $A_m = A - mI$, будет неотрицательным, так как для любого $u \in D(A)$

$$\langle W_0^2 u, u \rangle = \langle A_m(A_m u), u \rangle = \langle A_m u, A_m u \rangle = \|A_m u\|^2 \geq 0.$$

Таким образом, для оператора W_0^2 на отрезке $[0, \mu^2]$ выполнены условия теоремы 6.8. Поэтому оператор $W_0^2 - \mu^2 I = W_\mu W_{-\mu} = A_a A_b$ является положительным. ►

Теорема 6.10. Пусть A — симметрический оператор в гильбертовом пространстве $\mathcal{H}$ и элемент $u \in \mathcal{H}$ удовлетворяет условию $Au \neq 0$. Если число $\bar{\lambda}_1 = \frac{\langle Au, u \rangle}{\|u\|^2}$ принадлежит интервалу (a, b) , причем на отрезке $[a, b]$ находится единственная точка $\lambda \in (a, b)$ спектра оператора A , являющаяся собственным значением этого оператора, то справедливо неравенство

$$\psi(b) \leq \lambda \leq \psi(a), \quad (6.139)$$

где

$$\psi(t) = \frac{t \langle Au, u \rangle - \|Au\|^2}{t \|u\|^2 - \langle Au, u \rangle}.$$

◄ Оператор A на отрезках $[a, \lambda - \varepsilon]$ и $[\lambda + \varepsilon, b]$ при достаточно малом $\varepsilon > 0$ удовлетворяет условиям теоремы 6.9. Поэтому операторы $A_1 = A_a A_{\lambda - \varepsilon} = (A - aI)(A - (\lambda - \varepsilon)I)$ и $A_2 = A_{\lambda + \varepsilon} A_b = (A - (\lambda + \varepsilon)I)(A - bI)$ являются положительными. Следовательно,

$$\begin{aligned} \langle A_1 u, u \rangle &= \langle (A^2 - (a + \lambda - \varepsilon)A + a(\lambda - \varepsilon)I)u, u \rangle = \\ &= \|Au\|^2 - (a + \lambda - \varepsilon) \langle Au, u \rangle + a(\lambda - \varepsilon) \|u\|^2 = \\ &= (\lambda - \varepsilon)(a \|u\|^2 - \langle Au, u \rangle) - (a \langle Au, u \rangle - \|Au\|^2) > 0 \end{aligned}$$

и аналогично

$$\begin{aligned}\langle A_2 u, u \rangle &= \langle (A^2 - (\lambda + \varepsilon + b)A + a(\lambda + \varepsilon)I)u, u \rangle = \\ &= \|Au\|^2 - (\lambda + \varepsilon + b)\langle Au, u \rangle + b(\lambda + \varepsilon)\|u\|^2 = \\ &= (\lambda + \varepsilon)(b\|u\|^2 - \langle Au, u \rangle) - (b\langle Au, u \rangle - \|Au\|^2) > 0.\end{aligned}$$

Из этих неравенств, учитывая, что, согласно условию теоремы, $a < \frac{\langle Au, u \rangle}{\|u\|^2} < b$, т.е. $a\|u\|^2 - \langle Au, u \rangle < 0$ и $b\|u\|^2 - \langle Au, u \rangle > 0$, находим

$$\begin{aligned}\lambda - \varepsilon &< \frac{a\langle Au, u \rangle - \|Au\|^2}{a\|u\|^2 - \langle Au, u \rangle} = \psi(a), \\ \lambda + \varepsilon &> \frac{b\langle Au, u \rangle - \|Au\|^2}{b\|u\|^2 - \langle Au, u \rangle} = \psi(b).\end{aligned}$$

Переходя к пределу при $\varepsilon \rightarrow 0$, получаем (6.139). ►

Теорему 6.10 можно применить для двусторонней оценки наименьшего собственного значения λ_1 симметрического оператора A , если известна такая гарантированная оценка $\underline{\lambda}_2$ снизу следующего собственного значения λ_2 , что $\bar{\lambda}_1 = \frac{\langle Au, u \rangle}{\|u\|^2} < \underline{\lambda}_2 < \lambda_2$. Эта теорема верна для любого $a < 0$, причем в соответствии с (6.139)

$$\psi(-\infty) = \lim_{t \rightarrow -\infty} \frac{t\langle Au, u \rangle - \|Au\|^2}{t\|u\|^2 - \langle Au, u \rangle} = \frac{\langle Au, u \rangle}{\|u\|^2} = \bar{\lambda}_1$$

и поэтому

$$\psi(\underline{\lambda}_2) = \frac{\underline{\lambda}_2 \langle Au, u \rangle - \|Au\|^2}{\underline{\lambda}_2 \|u\|^2 - \langle Au, u \rangle} \leq \lambda_1 \leq \bar{\lambda}_1 = \psi(-\infty). \quad (6.140)$$

В случае положительного оператора A имеем $\langle Au, u \rangle > 0$ и $\bar{\lambda}_1 > 0$, так что вместо (6.140) получаем

$$\bar{\lambda}_1 \left(1 - \frac{\kappa - \bar{\lambda}_1}{\underline{\lambda}_2 - \bar{\lambda}_1}\right) \leq \lambda_1 \leq \bar{\lambda}_1, \quad \kappa = \frac{\|Au\|^2}{\langle Au, u \rangle}. \quad (6.141)$$

Ясно, что (6.141) имеет смысл при условии $\bar{\lambda}_1 \leq \kappa$, выполняющимся в силу *неравенства Коши — Буняковского* $|\langle Au, u \rangle| \leq \|Au\| \|u\|$. Поскольку для положительного оператора $\lambda_1 > 0$, то использование (6.141) эффективно только при условии $\kappa < \underline{\lambda}_2$, так как в противном случае левая часть в (6.141) не будет положительной.

Отметим, что оценки (6.140) и (6.141) сохраняют силу и в том случае, если собственное значение λ_1 кратное, но изолированное. В этом случае под $\underline{\lambda}_2$ следует понимать гарантированную оценку снизу наименьшего собственного значения $\lambda_2 > \lambda_1$. Аналогично можно получить двустороннюю оценку для λ_2 и т.д.

Теперь вернемся к операторному уравнению (6.126). Можно показать*, что неравенство (6.141) применимо для двусторонней оценки наименьшего собственного значения λ_1 этого уравнения, если положить

$$\bar{\lambda}_1 = \frac{\langle Au, u \rangle}{\langle Bu, u \rangle}, \quad \kappa = \frac{\langle Au, f \rangle}{\langle Au, u \rangle}, \quad (6.142)$$

где ненулевые элементы $u \in D(A)$ и $f \in D(B)$ удовлетворяют равенству $Au = Bf$. При этом по-прежнему гарантированная оценка $\underline{\lambda}_2$ снизу собственного значения λ_2 должна превышать значение κ , а оценка (6.141) применима при условии $\bar{\lambda}_1 < \kappa < \underline{\lambda}_2$.

Пример 6.12. Используем (6.141) и (6.142) для уточнения двусторонней оценки собственного значения λ_1 однородного уравнения (6.132), рассмотренного в примере 6.11. Для функции $\tilde{u}(\xi) = \xi(1 - \xi)$, удовлетворяющей однородным граничным условиям $u(0) = u(1) = 0$, найдем $\tilde{f}$ из условия $A\tilde{u} = B\tilde{f}$. Имеем

$$B\tilde{f} = \frac{4}{(1 + \xi)^2} \tilde{f} = A\tilde{u} = -\frac{d^2 \tilde{u}(\xi)}{d\xi^2} = 2,$$

*См. Ректорис К.

откуда $\tilde{f} = (1 + \xi)^2/2$. Далее вычисляем

$$\langle A\tilde{u}, \tilde{f} \rangle = \int_0^1 2 \frac{(1 + \xi)^2}{2} d\xi = \frac{7}{3}$$

и, используя найденное в примере 6.11 значение $\langle A\tilde{u}, \tilde{u} \rangle = 1/3$, в соответствии с (6.142) находим

$$\kappa = \frac{\langle A\tilde{u}, \tilde{f} \rangle}{\langle A\tilde{u}, \tilde{u} \rangle} = 7. \quad (6.143)$$

В качестве гарантированной оценки снизу собственного значения λ_2 используем (6.135) при $n = 2$, т.е.

$$\underline{\lambda}_2 = \frac{(2\pi)^2}{4} = \pi^2 \approx 9,8696, \quad (6.144)$$

а в качестве $\bar{\lambda}_1$ примем верхнюю оценку в (6.136):

$$\bar{\lambda}_1 = \frac{1}{4(25 - 36 \ln 2)} \approx 5,3532. \quad (6.145)$$

Таким образом, условие $\bar{\lambda}_1 < \kappa < \underline{\lambda}_2$ выполнено. Подставляя (6.143)–(6.145) в (6.141), получаем

$$3,4021 \leq \lambda_1 \leq 5,3532. \quad (6.146)$$

Сравнение с (6.136) показывает, что при совпадении верхних оценок нижняя оценка 3,4021 несколько улучшена по отношению к прежнему значению 2,4674.

Попытаемся улучшить верхнюю оценку $\bar{\lambda}_1$, используя функцию $u^*(\xi) = \tilde{u}(\xi) + C\tilde{u}^2(\xi)$, также удовлетворяющую граничным условиям $u(0) = u(1) = 0$. Константу C найдем, согласно теореме 5.6, из необходимого условия минимума отношения $\frac{\langle Au^*, u^* \rangle}{\langle Bu^*, u^* \rangle}$.

Предварительно найдем

$$\begin{aligned} Au^* &= -\frac{d^2 u^*(\xi)}{d\xi^2} = 2 + 4C\xi(1-\xi) - 2C(1-2\xi)^2 = \\ &= 2(1 - C + 6C\xi(1-\xi)) \end{aligned}$$

и вычислим

$$\begin{aligned} \langle Au^*, u^* \rangle &= 2 \int_0^1 (1 - C + 6C\xi(1-\xi)) \times \\ &\times (\xi(1-\xi) + C\xi^2(1-\xi)^2) d\xi = \frac{35 + 14C + 2C^2}{105}. \end{aligned}$$

Аналогично

$$\begin{aligned} \langle Bu^*, u^* \rangle &= \int_0^1 4 \frac{(\xi(1-\xi) + C\xi^2(1-\xi)^2)^2}{(1+\xi)^2} d\xi = \\ &= 4 \left(\frac{25 - 36 \ln 2}{3} - \frac{499 - 720 \ln 2}{10} C + \frac{2329 - 3360 \ln 2}{35} C^2 \right). \end{aligned}$$

В результате получаем

$$\frac{\langle Au^*, u^* \rangle}{\langle Bu^*, u^* \rangle} \approx \frac{35 + 14C + 2C^2}{6,5391 + 2,7653C + 0,3132C^2}.$$

Наименьшее значение этого выражения $\lambda_1^* \approx 5,3028$ соответствует значению $C \approx 0,9785$. Таким образом, замена функции $\tilde{u}(x)$ более сложной функцией $u^*(x)$ позволила понизить верхнюю оценку наименьшего собственного значения λ_1 менее чем на 0,5%.

Для уточнения нижней оценки значения λ_1 запишем

$$Bf^* = \frac{4f^*(x)}{(1+\xi)^2} = Au^* = 2(1 - C + 6C\xi(1-\xi)).$$

Отсюда найдем

$$f^*(\xi) = \frac{(1+\xi)^2}{2} (1 - C + 6C\xi(1-\xi)),$$

а затем с учетом полученного значения C вычислим

$$\langle Au^*, f^* \rangle = \int_0^1 (1+\xi)^2 (1 - C + 6C\xi(1-\xi))^2 d\xi = \frac{245 - 7C + 50C^2}{105}.$$

Таким образом,

$$\kappa^* = \frac{\langle Au^*, f^* \rangle}{\langle Au^*, u^* \rangle} = \frac{245 - 7C + 50C^2}{35 + 14C + 2C^2} \approx 5,6511.$$

Заменяя в (6.141) $\bar{\lambda}_1$ на λ_1^* , κ на κ^* и подставляя (6.144), получаем

$$4,8983 \leq \lambda_1 \leq 5,3028.$$

Сравнение с (6.146) показывает, что нижняя оценка возросла и теперь отличается от верхней оценки менее чем на 8%.

6.10. Особенности выбора базисных функций

Из рассмотрения различных вариантов *метода ортогональных проекций* видно, что *базисные функции* u_n в *гильбертовом пространстве* $\mathcal{H}$, которые входят в представление приближенного решения *операторного уравнения* $Au = f$, должны удовлетворять ряду требований. Сформулируем некоторые требования к этим функциям, предполагая, что ищется приближенное решение вида

$$\tilde{u}_N = \sum_{n=1}^N a_n u_n. \quad (6.147)$$

1. Конечная система $\{u_n\}_N$ используемых функций должна быть линейно независимой. Тогда такая система функций

образует базис в N -мерном подпространстве $H_N \subset H$, которое является *линейной оболочкой* этой системы, а $\tilde{u}_N \in H_N$.

2. Для сходимости приближенного решения $\tilde{u}_N \in H_N$ к *слабому решению* $u_* \in H$ уравнения $Au = f$ можно потребовать, чтобы система $\{u_n\}_N$ используемых функций входила в систему $\{u_n\}$, являющуюся ортонормированным (или ортогональным) базисом в H . Для того чтобы гарантировать, что норма *невязки операторного уравнения* при $N \rightarrow \infty$ стремится к нулю, можно потребовать выполнения некоторых дополнительных условий.

3. Базисные функции u_n должны принадлежать *области определения* $D(A)$ оператора A , так как для нахождения коэффициентов a_n в (6.147) необходимо предварительно вычислить значения $\langle Au_n, v_k \rangle$, где $v_k, k = \overline{1, N}$, — *проекционные функции*, а это возможно лишь при выполнении сформулированного требования.

Перейдем к обсуждению перечисленных требований.

Необходимость выполнения условия линейной независимости конечной системы $\{u_n\}_N$ выбранных базисных функций очевидна. При нарушении этого условия матрица СЛАУ (6.80) будет вырожденной. Однако даже если базисные функции линейно независимы, увеличение их количества N , преследующее цель уменьшить погрешность приближенного решения (6.147), может привести из-за возрастания порядка СЛАУ к росту вычислительной погрешности. Например, если с увеличением N эти функции становятся все более близкими (т.е. нормы их разностей уменьшаются), то увеличивается *число обусловленности матрицы* и растет чувствительность решения СЛАУ к погрешностям в коэффициентах и правых частях уравнений СЛАУ [IV]. Поэтому при ограниченном числе N выбор этих функций связан с проблемой минимизации возникающей погрешности. Так, если в случае линейной задачи в качестве одной из базисных функций использовать какое-либо частное решение этой задачи, то его дополнение до полного решения

может быть достаточно точно аппроксимировано меньшим числом базисных функций.

Выбор в качестве базисных функций элементов ортонормированного (или ортогонального) базиса в $\mathcal{H}$ гарантирует сходимость при $N \rightarrow \infty$ приближенного решения, полученного при помощи метода ортогональных проекций. При использовании *метода наименьших квадратов* условия сходимости устанавливает теорема 6.7. Одно из этих условий требует, чтобы функции u_n принадлежали *счетному базису* в $\mathcal{H}$. В этом случае норма невязки операторного уравнения при $N \rightarrow \infty$ стремится к нулю.

Если используются *методы Бубнова — Галеркина и Рунца*, то, согласно замечанию 6.1, для стремления к нулю невязки операторного уравнения необходимо, чтобы функции u_n были *собственными элементами положительно определенного оператора A* , хотя для сходимости приближенного решения достаточно принадлежности этих функций базису в *энергетическом пространстве $\mathcal{H}_A$* , ортонормированному относительно *энергетического скалярного произведения*. В более общем виде условия сходимости приближенного решения сформулированы в рамках *проекционного метода* (см. Д.6.1), частным случаем которого является метод ортогональных проекций.

Однако при приближенном решении прикладных задач базисные функции не всегда удается выбрать так, чтобы они были элементами ортонормированного (или ортогонального) базиса в $\mathcal{H}$ или $\mathcal{H}_A$. Тогда можно ограничиться более слабым требованием, чтобы функции u_n принадлежали некоторому *счетному базису*. Напомним, что *система $\{u_n\}$* является базисом в гильбертовом пространстве лишь в том случае, если она является *замкнутой* в этом пространстве [IX].

При приближенном решении уравнения $Au = f$ с положительно определенным оператором A методами Бубнова — Галеркина или Рунца требование выбора базисных функций из области $D(A)$ определения этого оператора можно ослабить. Дело в том, что в общем случае речь идет о поиске *слабого ре-*

шения $u_* \in \mathcal{H}_A$ этого уравнения в энергетическом пространстве $\mathcal{H}_A$, так что, вообще говоря, $u_* \notin D(A)$. Но тогда функции u_n в (6.147) можно выбирать не только из $D(A)$, но и из $\mathcal{H}_A$.

Пример 6.13. Рассмотрим краевую задачу

$$\begin{cases} -(p(x)u'(x))' + q(x)u(x) = f(x), & x \in (0, 1); \\ u(0) = u(1) = 0. \end{cases} \quad (6.148)$$

Пусть $0 < p_0 \leq p(x) \leq \bar{p}$, $|p'(x)| \leq p_1$ и $0 < q(x) \leq \bar{q}$ при любых $x \in (0, 1)$, p_0 , $\bar{p}$, p_1 , $\bar{q}$ — положительные константы. Оператор Штурма — Лиувилля A , вводимый равенством $Au = -(pu')' + qu$, определен на множестве $D(A)$ дважды непрерывно дифференцируемых на отрезке $[0, 1]$ функций, обращающихся в нуль на концах этого отрезка. Это множество является линейным многообразием гильбертова пространства $L_2[0, 1]$ функций, суммируемых с квадратом на отрезке $[0, 1]$.

На множестве $D(A)$ оператор A положительно определен (см. пример 5.10), но краевая задача (6.148) не будет иметь классического решения в классе функций из $D(A)$, если функция $f(x)$ не принадлежит области $R(A)$ значений этого оператора (например, имеет в интервале $(0, 1)$ точки разрыва). Для нахождения слабого решения операторного уравнения $Au = f$ необходимо пополнить множество $D(A)$. Пополнение $\mathcal{H}_A$ этого множества можно построить по энергетической норме $\|\cdot\|_A$, индуцированной энергетическим скалярным произведением, определяемым для функций $u, v \in D(A)$ равенством

$$\langle u, v \rangle_A = \langle Au, v \rangle = \int_0^1 (-(p(x)u'(x))' + q(x)u(x))v(x) dx.$$

Отсюда, интегрируя по частям, с учетом граничных условий находим

$$\langle u, v \rangle_A = \int_0^1 (p(x)u'(x)v'(x) + q(x)u(x)v(x)) dx.$$

В результате для энергетической нормы с учетом (5.95) и оценок функций $p(x)$ и $q(x)$ получим

$$\begin{aligned} \|u\|_A^2 &= \langle u, u \rangle_A = \int_0^1 (p(x) u'^2(x) + q(x) u(x)^2) dx \leq \\ &\leq \bar{p} \int_0^1 u'^2(x) dx + \bar{q} \int_0^1 u^2(x) dx \leq \left(\bar{p} + \frac{\bar{q}}{\pi^2} \right) \int_0^1 (u'(x))^2 dx. \end{aligned} \quad (6.149)$$

Оператор $B = -\frac{d^2}{dx^2}$ также положительно определен на множестве $D(A)$ (см. пример 5.10). Это множество, согласно теореме 4.23, всюду плотно в энергетическом пространстве $\mathcal{H}_B$, которое является пополнением множества $D(A)$ по норме $\|\cdot\|_B$, индуцированной энергетическим скалярным произведением, которое для функций $u, v \in D(A)$ определяется соотношением

$$\langle u, v \rangle_B = \langle Bu, v \rangle = \int_0^1 (-u''(x)) v(x) dx.$$

Отсюда интегрированием по частям с учетом граничных условий получим

$$\langle u, v \rangle_B = \int_0^1 u'(x) v'(x) dx, \quad \|u\|_B^2 = \int_0^1 (u'(x))^2 dx. \quad (6.150)$$

Функции $u_n(x) = \sin n\pi x$, $x \in [0, 1]$, $n \in \mathbb{N}$, являются собственными элементами оператора B (см. пример 4.20) и образуют в $D(A)$ ортогональный базис [IX]. Поэтому система этих функций является замкнутой в $D(A)$, а значит, и в $\mathcal{H}_B$, поскольку множество $D(A)$ является всюду плотным в $\mathcal{H}_B$ (см. определение 4.11 и теорему 4.23). Но из (6.149), (6.150) и оценок для функции $p(x)$ следует, что

$$p_0 \|u\|_B^2 \leq \|u\|_A^2 \leq \left(\bar{p} + \frac{\bar{q}}{\pi^2} \right) \|u\|_B^2.$$

Следовательно, $\mathcal{H}_A$ состоит из тех же элементов, что и $\mathcal{H}_B$, так что система функций $u_n(x)$ является замкнутой и в $\mathcal{H}_A$.

Систему функций $u_n(x) = \sin n\pi x$, $n \in \mathbb{N}$, можно ортонормировать на отрезке $[0, 1]$, но их использование в качестве базисных функций при приближенном решении краевой задачи (6.148) методами Бубнова — Галеркина или Ритца не обеспечит стремления к нулю при $N \rightarrow \infty$ невязки в соответствующем операторном уравнении $Au = f$. Дело в том, что эти функции, будучи собственными элементами оператора B , не являются собственными элементами оператора A (см. замечание 6.1). #

Для дифференциального оператора A условие $u_n \in D(A)$, $n = \overline{1, N}$, учитываемое при выборе базисных функций в (6.79), означает, что функции u_n должны удовлетворять не только требованиям дифференцируемости, но и граничным условиям на поверхности S , ограничивающей область V решаемой задачи. Напомним, что в основе большинства дифференциальных уравнений математической физики лежат локальные формы законов сохранения физических субстанций (см. 2), которые для процессов, не зависящих от времени, можно представить в виде

$$Au = -\nabla(A_1 u) + B_1 u = f, \quad D(A) \subset R(A) \subset \mathbb{R}, \quad (6.151)$$

где $u \in D(A)$ и $f \in R(A)$ — искомая и заданная функции пространственных координат, а A_1 и B_1 — линейные дифференциальные операторы, относительно которых предполагаем, что порядок производных по пространственным координатам в B_1 не выше, чем в A_1 (тогда порядок производных в A на единицу выше, чем в A_1). Ясно, что $A_1 u$ в данном случае будет векторной функцией (часто градиентом искомой функции u).

Выбрав в качестве проекционных функции $v_k \in L_2(V)$, $k \in \mathbb{N}$, образующие в гильбертовом пространстве $L_2(V)$ счетный базис, согласно теореме 5.1, вместо (6.151) можем записать $\langle Au, v_k \rangle = \langle f, v_k \rangle$, $k = \overline{1, N}$, или

$$\int_V (-\nabla(A_1 u) + B_1 u) v_k dV = \int_V f v_k dV, \quad k \in \mathbb{N}. \quad (6.152)$$

Равенства (6.152) вместе с *краевыми условиями* составляют *интегральную формулировку задачи*, описываемую операторным уравнением (6.151).

Пусть приближенное решение уравнения (6.151) определяется методом ортогональных проекций в виде

$$\tilde{u}_N = \sum_{n=1}^N a_n u_n. \quad (6.153)$$

Тогда, подставив (6.153) в левую часть (6.152), получим уравнение, левая часть которого представлена в виде линейной комбинации N слагаемых

$$\langle Au_n, v_k \rangle = \int_V (-\nabla(A_1 u_n) + B_1 u_n) v_k dV, \quad k, n = \overline{1, N}, \quad (6.154)$$

с неизвестными коэффициентами a_n . В результате получим СЛАУ

$$\sum_{n=1}^N a_n \langle Au_n, v_k \rangle = \langle f, v_k \rangle, \quad k = \overline{1, N}, \quad (6.155)$$

относительно неизвестных коэффициентов a_n , $n = \overline{1, N}$. Если $u_n \in D(A)$, то в соответствии с *первой формулой Грина* вместо (6.154) при $k, n = \overline{1, N}$ получим

$$\begin{aligned} \langle Au_n, v_k \rangle = & \int_V v_k B_1 u_n dV + \\ & + \int_V (A_1 u_n) \nabla v_k dV - \int_S v_k (A_1 u_n) \mathbf{n} dS, \end{aligned} \quad (6.156)$$

где $\mathbf{n}$ — единичный вектор внешней нормали к поверхности S .

Из (6.156) следует, что требования к дифференцируемости базисных функций u_n можно было бы частично переложить на проекционные функции v_k , т.е. выбрать в (6.153) $u_n \in D_1 = D(A_1) \cup D(B_1)$, $n = \overline{1, N}$, непрерывно продолжив при этом

оператор A на $D_1 \subset L_2(V)$, а от функций v_k потребовать непрерывной дифференцируемости в V . Действительно, пусть необходимые условия дифференцируемости, определяющие принадлежность $u_n \in D(A)$, $n = \overline{1, N}$, нарушены в точках $P^* \in S^*$ некоторой поверхности S^* , разделяющей область V на две подобласти V_1 и V_2 ($V = V_1 \cup V_2 \cup S^*$), но выполнены в каждой из этих подобластей. Тогда вместо (6.156) получим

$$\begin{aligned} \langle Au_n, v_k \rangle = & \int_V v_k B_1 u_n dV + \int_{V_1} (A_1 u_n) \nabla v_k dV + \\ & + \int_{V_2} (A_1 u_n) \nabla v_k dV - \int_S v_k (A_1 u_n) \mathbf{n} dS - \\ & - \int_{S^*} ((A_1 u_n)_1 \mathbf{n}^* - (A_1 u_n)_2 \mathbf{n}^*) v_k dS, \quad (6.157) \end{aligned}$$

где $\mathbf{n}^*$ — единичный вектор нормали к поверхности S^* , внешней по отношению к подобласти V_1 , а через $(A_1 u_n)_1$ и $(A_1 u_n)_2$ обозначены пределы векторной функции $A_1 u(P)$ при стремлении точки P к $P^* \in S^*$ со стороны подобластей V_1 и V_2 соответственно. Но при выборе $u_n \in D_1$ эти пределы равны, так что интеграл по поверхности S^* исчезает. Поскольку точки $P^* \in S^*$ образуют в области V множество, имеющее равную нулю меру Лебега, то

$$\int_{V_1} (A_1 u_n) \nabla v_k dV + \int_{V_2} (A_1 u_n) \nabla v_k dV = \int_V (A_1 u_n) \nabla v_k dV,$$

т.е. при $u_n \in D(A_1)$ (6.157) равносильно (6.156).

Таким образом, учитывая (6.156), СЛАУ (6.155) можно записать в виде

$$\sum_{n=1}^N G_{nk} a_n = f_k, \quad k = \overline{1, N}, \quad (6.158)$$

где при $k, n = \overline{1, N}$

$$G_{nk} = \int_V v_k B_1 u_n dV + \int_V (A_1 u_n) \nabla v_k dV - \int_S v_k (A_1 u_n) n dS, \quad (6.159)$$

а значения $f_k = \langle f, v_k \rangle$, $k = \overline{1, N}$, равны интегралам в правой части (6.152). Следует отметить, что в методе наименьших квадратов такой подход не снижает требований к дифференцируемости функций u_n , так как преобразование (6.156) скалярного произведения при выборе $v_k = Au_k$ вызовет, наоборот, повышение требований к дифференцируемости функций u_n . В самом деле, $\nabla v_k = \nabla(Au_k)$ и $u_n \in D(\nabla A)$.

При выборе базисных функций можно сделать еще один шаг по пути ослабления требований к их гладкости в области V . Дело в том, что по физическому смыслу функция $A_1 u$ обычно является вектором плотности потока некоторой физической субстанции (см. 1). Законы сохранения таких физических субстанций, как масса, энергия, заряд, количество движения, допускают разрывность этой векторной функции на некоторой *поверхности разрыва* S^* , но требуют непрерывности ее проекции $(A_1 u(P))n^*(P)$, $P \in S^*$, на направление нормали n^* к этой поверхности. Ясно, что это условие будет выполнено, если в (6.159) функции u_n выбраны так, что $u_n \in D_1 = D(A_1) \cup D(B_1)$, $n = \overline{1, N}$.

Предположим теперь, что функции u_n , $n = \overline{1, N}$, в (6.159) выбраны так, что на поверхности S^* , разделяющей область V на две подобласти V_1 и V_2 ($V = V_1 \cup V_2 \cup S^*$), проекция $(A_1 u)n^*$ терпит разрыв, т.е. в (6.157) нарушено условие

$$(A_1 u_n(P^*))_1 = (A_1 u_n(P^*))_2, \quad P^* \in S^*. \quad (6.160)$$

Допустим, что функции v_k , $k = \overline{1, N}$, также разрывны на этой поверхности, т.е. пределы $v_k^{(1)}(P^*)$ и $v_k^{(2)}(P^*)$ каждой

функции $v_k(P)$ при стремлении точки P к $P^* \in S^*$ со стороны подобластей V_1 и V_2 могут быть не равны. Тогда вместо (6.157) будем иметь

$$\begin{aligned} \langle Au_n, v_k \rangle = & \int_V v_k B_1 u_n dV + \int_{V_1} (A_1 u_n) \nabla v_k dV + \\ & + \int_{V_2} (A_1 u_n) \nabla v_k dV - \int_S v_k (A_1 u_n) \mathbf{n} dS - \\ & - \int_{S^*} (v_k^{(1)} (A_1 u_n)_1 \mathbf{n}^* - v_k^{(2)} (A_1 u_n)_2 \mathbf{n}^*) dS. \end{aligned} \quad (6.161)$$

Но для сходимости приближенного решения необходимо, чтобы условие (6.160) было выполнено. Для этого воспользуемся методом взвешенных невязок, выбрав в качестве весовых функций $\frac{1}{2}(v_k^{(1)} + v_k^{(2)})$:

$$\int_{S^*} ((A_1 u_n)_1 \mathbf{n}^* - (A_1 u_n)_2 \mathbf{n}^*) \frac{v_k^{(1)} + v_k^{(2)}}{2} dS = 0. \quad (6.162)$$

Вычитая левую часть (6.162) из правой части (6.161), получаем

$$\begin{aligned} \langle Au_n, v_k \rangle = & \int_V v_k B_1 u_n dV + \int_{V_1} (A_1 u_n) \nabla v_k dV + \int_{V_2} (A_1 u_n) \nabla v_k dV - \\ & - \int_S v_k (A_1 u_n) \mathbf{n} dS - \int_{S^*} ((A_1 u_n)_1 \mathbf{n}^* + (A_1 u_n)_2 \mathbf{n}^*) \frac{v_k^{(1)} - v_k^{(2)}}{2} dS. \end{aligned}$$

Отсюда вытекает, что для исчезновения интеграла по поверхности S^* достаточно непрерывности функций v_k , $k = \overline{1, N}$, в точках поверхности S^* .

Ясно, что поверхность S^* может быть выбрана в области V произвольно. Поэтому в случае непрерывных в области V

проекционных функций v_k , $k = \overline{1, N}$, обладающих в ней кусочно непрерывными производными, для вычисления элементов G_{nk} , $k, n = \overline{1, N}$, матрицы СЛАУ (6.158) применима формула (6.159). При этом допустимо, чтобы функции $A_1 u_n$, $n = \overline{1, N}$, были разрывны в точках, образующих в области V множество, мера Лебега которого равна нулю.

Например, если в (6.151) A является дифференциальным оператором второго порядка, т.е. A_1 имеет первый порядок, то для вычисления интегралов в (6.159) от базисных функций u_n , $n = \overline{1, N}$, достаточно потребовать, чтобы они обладали лишь кусочно непрерывными производными. Полученный результат справедлив и для частного случая метода ортогональных проекций — метода Бубнова — Галеркина при выборе в (6.153) $v_k = u_k$, $k = \overline{1, N}$. В случае, когда A является положительно определенным оператором, для приближенного решения уравнения (6.151) можно применить метод Ритца. Тогда функции u_n будут принадлежать соответствующему энергетическому пространству $\mathcal{H}_A$, в котором определена энергетическая норма $\|\cdot\|_A$, а функционал энергии (5.30) в рассматриваемом случае примет вид $J[u] = \|u\|_A^2 - 2(f, u)$. Отметим, что указанные возможности ослабления требований к гладкости базисных функций важны при численном решении задач математической физики, в частности *методом конечных элементов*.

Теперь обратимся к вопросу об удовлетворении граничных условий на поверхности S , ограничивающей область V . Пусть на участках $S_1 \subset S$ этой поверхности задано граничное условие

$$u(P) = f_1(P), \quad P \in S_1, \quad (6.163)$$

а на участках $S_2 = S \setminus S_1$ — граничное условие

$$(A_1 u(P))n(P) + \beta(P)u(P) = f_2(P), \quad P \in S_2, \quad (6.164)$$

где f_1 , β и f_2 — заданные функции. Тогда приближенное решение в виде (6.153) должно удовлетворять условию (6.163) в

каждой точке $P \in S_1$, а для выполнения (6.164) можно потребовать равенства нулю взвешенной невязки, возникающей при подстановке (6.153) в (6.164). При выборе функций v_k в качестве весовых будем иметь

$$\int_{S_2} ((A_1 \tilde{u}_N) n + \beta \tilde{u}_N - f_2) v_k dS = 0, \quad k = \overline{1, N}. \quad (6.165)$$

В этом случае вместо (6.153) целесообразно искать приближенное решение в виде

$$\tilde{u}_N = u_0 + \sum_{n=1}^N a_n u_n, \quad (6.166)$$

где $u_0(P) = f_1(P)$ при $P \in S_1$, а функции u_n и v_k выбрать так, чтобы $u_n(P) = v_k(P) = 0$, $P \in S_1$, $k, n = \overline{1, N}$. Подставим (6.166) в (6.152) и выполним преобразования, аналогичные тем, которые привели к СЛАУ (6.158). Тогда с учетом (6.165) получим систему уравнений (6.158), в которой при $k, n = \overline{1, N}$

$$G_{nk} = \int_V (v_k B_1 u_n + (A_1 u_n) \nabla v_k) dV + \int_{S_2} \beta u_n v_k dS,$$

$$f_k = \int_V (f v_k - (A_1 u_0) \nabla v_k - v_k B_1 u_0) dV + \int_{S_2} (f_2 - \beta u_0) v_k dS.$$

Описанный подход позволяет не накладывать на выбор функций $u_n \in D_1 = D(A_1) \cup D(B_1)$, $n = \overline{0, N}$, ограничений, вытекающих из условия (6.164). В связи с этим такое **граничное условие** обычно называют **естественным**. В противоположность этому **граничное условие** (6.163), которому должно удовлетворять (6.166), называют **главным**, а иногда **предварительным** в том смысле, что оно должно быть учтено еще до нахождения коэффициентов a_n при построении приближенного

решения (6.166). Хотя приближенное решение можно строить, не соблюдая граничных условий, функции u_n все же лучше выбирать с учетом этих условий, так как при их выполнении последовательность приближенных решений будет иметь более высокую сходимость*, особенно в окрестности точек $P \in S_2$.

Более строгое разделение граничных условий на главные и естественные следует из вариационной формулировки задачи для *положительно определенного оператора* A . Естественные граничные условия можно получить из условий стационарности функционала, входящего в эту формулировку, если приравнять нулю его *первую вариацию*. Граничные условия, не представленные в условиях стационарности функционала, являются главными. Преобразуя вариационную формулировку задачи, можно получить *функционал Лагранжа*, называемый иногда полным. Этот функционал отличается тем, что в условия его стационарности входят все граничные условия (отметим, что в теории упругости термин „функционал Лагранжа“ имеет иной смысл). При использовании такого функционала все граничные условия будут естественными и базисные функции можно выбирать лишь с учетом ограничений по их дифференцируемости.

Если в вариационную формулировку задачи входит минимизируемый функционал, который допустимо рассматривать на множестве функций, удовлетворяющих главному граничному условию вида (6.163), то это ограничение можно снять **методом штрафа**. Суть этого метода состоит в добавлении к функционалу положительного слагаемого, которое быстро возрастает, когда приближенное решение $\tilde{u}_N$ не удовлетворяет условию (6.163). При минимизации такого функционала $\tilde{u}_N(P) \rightarrow f_1(P)$ при $N \rightarrow \infty$, причем скорость сходимости к $f_1(P)$ в случае $P \in S_1$ будет зависеть от коэффициента, устанавливающего „тариф штрафа“ за нарушение условия (6.163).

*См., например: Мизлин С.Г., 1966, а также Ректорис К.

Пример 6.14. Рассмотрим краевую задачу для уравнения Пуассона

$$-\nabla^2 u(M) = f(M), \quad M \in V, \quad (6.167)$$

с заданными на кусочно гладкой поверхности S , ограничивающей область V , граничными условиями

$$u(P) = g(P), \quad P \in S_1 \subset S, \quad (6.168)$$

$$\nabla u(P) \mathbf{n}(P) + \alpha(P)u(P) = \alpha(P)h(P), \quad P \in S_2 = S \setminus S_1, \quad (6.169)$$

где $\mathbf{n}(P)$ — единичный вектор внешней нормали к поверхности S в точке $P \in S$. Одной из интерпретаций этой задачи может быть нахождение установившегося распределения температуры $u(M)$ в области V при заданном распределении $g(P)$ температуры на участках S_1 поверхности S и заданных на участках S_2 условиях теплообмена с окружающей средой, имеющей температуру $h(P)$ (функция $\alpha(P) > 0$ характеризует интенсивность теплообмена).

Краевой задаче (6.167)–(6.169) соответствует вариационная формулировка, включающая минимизируемый функционал [XV]

$$J[u] = \int_V \left(\frac{1}{2} (\nabla u)^2 - fu \right) dV + \frac{1}{2} \int_{S_2} \alpha (u - h)^2 dS, \quad (6.170)$$

который рассматривают на множестве функций $u(M)$, непрерывных на замыкании $\bar{V} = V \cup S$ области V , имеющих в V кусочно непрерывные производные по координатам точки $M \in V$ и принимающих значения $u(P) = g(P)$ в точках $P \in S_1$. Эти функции могут не удовлетворять граничному условию (6.169), являющемуся для этого функционала естественным. Если главное граничное условие (6.168) заменить естественным граничным условием

$$\nabla u(P) \mathbf{n}(P) + \beta u(P) = \beta g(P), \quad P \in S_1, \quad \beta > 0, \quad (6.171)$$

то вместо (6.170) получим функционал

$$J_{\beta}[u] = \int_V \left(\frac{1}{2} (\nabla u)^2 - fu \right) dV + \\ + \frac{1}{2} \int_{S_2} \alpha (u - h)^2 dS + \frac{\beta}{2} \int_{S_1} (u - g)^2 dS, \quad (6.172)$$

который можно рассматривать на множестве функций $u(M)$, не удовлетворяющих граничным условиям на всей поверхности S .

Последний интеграл в (6.172) равен среднеквадратичной невязке, вызванной нарушением главного граничного условия (6.168). При этом значение параметра β в (6.172) определяет „тариф штрафа“ за нарушение этого условия, а по физическому смыслу характеризует в (6.171) интенсивность теплообмена на участках S_1 поверхности S с условной окружающей средой, имеющей температуру $g(P)$, $P \in S_1$. Так как функция $\nabla u(P)n(P)$, $P \in S_1$, ограничена, то из (6.171) следует, что $u(P) \rightarrow g(P)$ при $\beta \rightarrow \infty$, т.е. при предельном переходе будет удовлетворено главное граничное условие (6.168) исходной краевой задачи. Ясно, что при минимизации (6.172) с фиксированным конечным значением β полученное приближенное решение $u_{\beta}(M)$, $M \in \bar{V}$, не будет удовлетворять (6.168), но по мере возрастания β значение последнего интеграла в (6.172) будет стремиться к нулю и $u_{\beta}(P) \rightarrow g(P)$, $P \in S_1$. #

Отметим, что построить базисные функции u_n и проекционные функции v_k так, чтобы $u_n(P) = v_k(P) = 0$, $P \in S_1$, $k, n = \overline{1, N}$, можно следующим образом. Если подобрана такая функция $\omega \in D_1 = D(A_1) \cup D(B_1)$, что $\omega(P) = 0$ при $P \in S_1$, а $\bar{u}_n \in D_1$ и функции $\bar{v}_k$ имеют в области V кусочно непрерывные производные, то функции $u_n = \omega \bar{u}_n$ и $v_k = \omega \bar{v}_k$ будут удовлетворять всем необходимым требованиям.

При решении многомерных задач функции $u_n(M)$ и $v_k(M)$ можно представить в виде произведения сомножителей, завися-

щих только от одной координаты точки $M \in V$. Если область V в направлении одной из координатных осей (например, оси Ox_1) ограничена плоскостями, перпендикулярными этой оси, то для упрощения подбора таких сомножителей можно принять коэффициенты a_n в (6.166) зависящими от этой координаты. Тогда вместо СЛАУ (6.158) получим систему N обыкновенных дифференциальных уравнений (ОДУ) относительно N неизвестных $a_n = a_n(x_1)$, $n = \overline{1, N}$, причем порядок этих ОДУ будет совпадать с порядком старшей производной по x_1 в дифференциальном операторе A . Для решения такой системы ОДУ необходимо использовать граничные условия, заданные на указанных плоскостях.

Описанный способ построения приближенного решения многомерной задачи, известный как **метод Канторовича**^{*}, аналогичен методу разделения переменных (*методу Фурье*). Этот способ можно использовать и в случае, если в выбранной системе криволинейных координат область V в направлении одной из них ограничена координатными поверхностями.

Рассмотрим случай, когда физические процессы, описываемые операторным уравнением вида (6.151), зависят от времени. Тогда оператор B_1 в (6.151) будет включать дифференцирование по времени, т.е. искомая функция $u(t, M)$ будет зависеть не только от пространственных координат точки $M \in V$ в области V , но и от времени t . При этом в (6.151) и в граничные условия (6.163), (6.164) могут входить функции $f(t, M)$, $f_1(t, P)$ и $f_2(t, P)$. Тогда, используя метод Канторовича, приближенное решение $\tilde{u}_N$ можно искать в виде

$$\tilde{u}_N(t, M) = u_0(t, M) + \sum_{n=1}^N a_n(t) u_n(M), \quad t > 0, \quad M \in V, \quad (6.173)$$

^{*}Л.В. Канторович (1912–1986) — отечественный математик и экономист, лауреат Нобелевской премии за работы по математической экономике.

где функции u_0 и u_n удовлетворяют условиям $u_0(t, P) = f_1(t, P)$ и $u_n(P) = 0$ при $P \in S_1$.

В этом случае, подставляя (6.173) в (6.152) и проводя при условии $v_k(P) = 0$, $P \in S_1$, преобразования, аналогичные предыдущим, получаем систему N ОДУ

$$\begin{aligned} \sum_{n=1}^N \left(\int_V v_k B_1(a_n u_n) dV + a_n \left(\int_V (A_1 u_n) \nabla v_k dV + \int_{S_2} \beta u_n v_k dS \right) \right) = \\ = \int_V (f v_k - (A_1 u_0) \nabla v_k - v_k B_1 u_0) dV + \int_{S_2} (f_2 - \beta u_0) v_k dS, \quad (6.174) \end{aligned}$$

где $k = \overline{1, N}$, относительно функций $a_n(t)$, $n = \overline{1, N}$. Важно подчеркнуть, что теперь оператор B_1 действует не только на функции u_n , зависящие от пространственных координат, но и на функции $a_n(t)$.

Для решения системы ОДУ (6.174) необходимо задать начальные распределения $\bar{u}^{(r)}(M) = u^{(r)}(0, M)$, $M \in V$, искомой функции (при $r = 0$) и ее производных по времени до порядка R включительно, если в операторе B_1 старшая производная по времени имеет порядок $R + 1$. Например, при $R = 0$ начальные значения $a_n(0)$ можно найти из условий равенства нулю взвешенной невязки, возникшей при подстановке (6.173) при $t = 0$ в начальное условие $u(0, M) = \bar{u}^{(0)}(M)$. Это приведет к СЛАУ

$$\begin{aligned} \sum_{n=1}^N a_n(0) \int_V u_n(M) v_k(M) dV = \\ = \int_V (\bar{u}^{(0)}(M) - u_0(0, M)) v_k(M) dv, \quad k = \overline{1, N}. \quad (6.175) \end{aligned}$$

Пример 6.15. Рассмотрим задачу нестационарной теплопроводности в твердом теле, описываемую дифференциальным

уравнением (при $t > 0$, $M \in V$) вида (2.53)

$$c(M) \frac{\partial T(t, M)}{\partial t} - \nabla(\lambda(M) \nabla T(t, M)) = I_V^{(q)}(t, M), \quad (6.176)$$

где $c(M)$ и $\lambda(M)$ — объемная теплоемкость и коэффициент теплопроводности материала тела, а $I_V^{(q)}(t, M)$ — объемная мощность источников энерговыделения. Пусть искомая функция $T(t, M)$, описывающая распределение температуры на участках $S_1 \subset S$ и $S_2 = S \setminus S_1$ кусочно гладкой поверхности S , ограничивающей область V , удовлетворяет граничным условиями вида (2.55) и (2.56)

$$T(t, P) = f_1(t, P), \quad P \in S_1, \quad (6.177)$$

$$\lambda(P) \nabla T(t, P) \mathbf{n}(P) + \beta(t, P) T(t, P) = f_2^*(t, P), \quad P \in S_2, \quad (6.178)$$

и начальному условию

$$T(0, M) = f_0(M), \quad M \in V \cup S_2, \quad (6.179)$$

в момент времени $t = 0$, принимаемый за начало отсчета. В (6.176)–(6.178) все функции за исключением искомой $T(t, M)$ считаем заданными.

Сопоставляя (6.176) с (6.151), устанавливаем, что в данном случае правая часть (6.151) соответствует функции $I_V^{(q)}$, оператор B_1 определен первым слагаемым в левой части (6.176), а оператор A_1 определен равенством $A_1 u = \lambda \nabla u$. Граничные условия (6.163), (6.164) и (6.177), (6.178) совпадают по форме. Поэтому если приближенное решение задачи (6.176)–(6.179) представить аналогично (6.173) в форме

$$\tilde{T}_N(t, M) = T_0(t, M) + \sum_{n=1}^N a_n(t) u_n(M), \quad t > 0, \quad M \in V, \quad (6.180)$$

где функции T_0 и u_n удовлетворяют условиям $T_0(t, P) = f_1(t, P)$ и $u_n(P) = 0$ при $P \in S_1$, то система ОДУ (6.174) для нахождения

функций $a_n(t)$ примет вид

$$\begin{aligned} \sum_{n=1}^N \left(\frac{da_n}{dt} \int_V c u_n v_k dV + a_n \int_V \lambda (\nabla u_n) \nabla v_k dV + a_n \int_{S_2} \beta u_n v_k dS \right) = \\ = \int_V \left(I_V^{(q)} v_k - \lambda (\nabla T_0) \nabla v_k - c \frac{\partial T_0}{\partial t} v_k \right) dV + \int_{S_2} (f_2 - \beta T_0) v_k dS, \end{aligned}$$

где $k = \overline{1, N}$. Необходимые для решения этой системы начальные значения $a_n(0)$ получим из СЛАУ вида (6.175) при $k = \overline{1, N}$

$$\sum_{n=1}^N a_n(0) \int_V u_n(M) v_k(M) dV = \int_V (f_0(M) - T_0(0, M)) v_k(M) dV.$$

Изложенный подход применен для приближенного решения одномерной задачи нестационарной теплопроводности в плоской стенке (см. примеры 6.7 и 6.10).

Дополнение 6.1. Проекционный метод

Рассмотрим обобщение метода ортогональных проекций, известное как **проекционный метод**. Этот метод есть метод приближенного решения операторного уравнения $Au = f$ в гильбертовом пространстве $\mathcal{H}$, где A — линейный оператор, имеющий область определения $D(A)$ и область значений $R(A)$, всюду плотные в $\mathcal{H}$. Будем также предполагать, что существует ограниченный обратный оператор A^{-1} .

Последовательность подпространств $H_N \subset \mathcal{H}$ назовем предельно плотной в $\mathcal{H}$, если для любого элемента $u \in \mathcal{H}$ последовательность расстояний $\{\rho(u, H_N)\}$ от u до подпространств H_N удовлетворяет условию

$$\rho(u, H_N) = \inf_{w \in H_N} \|u - w\| \rightarrow 0 \text{ при } N \rightarrow \infty. \quad (6.181)$$

Суть проекционного метода состоит в следующем. Для операторного уравнения $Au = f$ выберем две последовательности $\{U_N\}$ и $\{V_N\}$ подпространств $U_N \subset D(A)$ и $V_N \subset R(A)$, $N \in \mathbb{N}$, предельно плотные в $\mathcal{H}$. Для каждого подпространства V_N выберем ограниченный линейный оператор $P_N: \mathcal{H} \rightarrow V_N$ с областью значений $R(P_N) = V_N$, удовлетворяющий условию $P_N^2 = P_N$ (такие линейные операторы называют **проекционными**, или **проекторами**). Последовательность решений $u_N \in U_N$ операторных уравнений $P_N Au = P_N f$, $N \in \mathbb{N}$, при некоторых условиях можно рассматривать как последовательность приближений для решения операторного уравнения $Au = f$.

Описанная схема приближенного решения операторного уравнения $Au = f$ является корректной при выполнении двух требований:

- 1) каждое операторное уравнение $P_N Au = P_N f$ имеет решение $u_N \in U_N$, и притом единственное;
- 2) последовательность $\{u_N\}$ сходится к искомому решению u операторного уравнения $Au = f$ в каком-либо смысле (например, по норме гильбертова пространства).

Теорема 6.11. Пусть $A: D(A) \rightarrow R(A)$ — взаимно однозначный линейный оператор в гильбертовом пространстве $\mathcal{H}$, P — проектор в $\mathcal{H}$, удовлетворяющий условию $V = R(P) \subset R(A)$, и $U \subset D(A)$ — подпространство в $\mathcal{H}$. Если выполнены условия $PAU = V$ и

$$\|Pv\| \geq \tau \|v\|, \quad v \in AU, \quad (6.182)$$

где $\tau > 0$ — некоторое число, то для любого элемента $f \in \mathcal{H}$ операторное уравнение $PAu = Pf$ имеет решение в U , и притом единственное.

◀ Пусть $\tilde{P}$ — сужение проектора P на подпространство AU , т.е. $\tilde{P}v = Pv$ при $v \in AU$. В силу условия (6.182) линейный оператор $\tilde{P}$ имеет нулевое ядро, т.е. $\ker \tilde{P} = \{0\}$. Поэтому он

является взаимно однозначным и отображает подпространство AU взаимно однозначно на V . Следовательно, существует обратный оператор $\tilde{P}^{-1}: V \rightarrow AU$. Так как линейный оператор A обратим, то определен линейный оператор $A^{-1}\tilde{P}^{-1}: V \rightarrow U$, обратный к линейному оператору $\tilde{P}A$. Существование такого оператора означает, что для любого элемента $f \in \mathcal{H}$ операторное уравнение $\tilde{P}Au = Pf$ имеет решение u , принадлежащее U , и притом единственное. При этом $u = A^{-1}\tilde{P}^{-1}(Pf)$. ►

Теорема 6.12. Пусть $A: D(A) \rightarrow R(A)$ — взаимно однозначный линейный оператор в гильбертовом пространстве $\mathcal{H}$, для которого $D(A)$ и $R(A)$ всюду плотны в $\mathcal{H}$; $\{U_N\}$ и $\{V_N\}$ — некоторые последовательности подпространств в $\mathcal{H}$ и $P_N: \mathcal{H} \rightarrow V_N$, $N \in \mathbb{N}$, — проекторы, образующие равномерно ограниченную последовательность, т.е. $\|P_N\| \leq C$, $N \in \mathbb{N}$, для некоторого числа $C > 0$. Тогда, для того чтобы при любом $f \in \mathcal{H}$ каждое операторное уравнение $P_N Au = P_N f$, $N \in \mathbb{N}$, имело единственное решение $u_N \in U_N$, причем $\|Au_N - f\| \rightarrow 0$ при $N \rightarrow \infty$, необходимо и достаточно, чтобы выполнялись следующие условия:

- 1) последовательность подпространств AU_N предельно плотна в $\mathcal{H}$;
- 2) $P_N AU_N = V_N$, $N \in \mathbb{N}$;
- 3) существует такое $\tau > 0$, что для каждого номера N выполняется неравенство $\|P_N v\| \geq \tau \|v\|$, $v \in AU_N$.

При выполнении указанных трех условий скорость сходимости последовательности $\{Au_N - f\}$ к нулю определяется соотношением

$$\|Au_N - f\| \leq \left(1 + \frac{C}{\tau}\right) \rho(f, AU_N).$$

◄ **Необходимость.** Предположим, что для любого элемента $f \in \mathcal{H}$ каждое уравнение $P_N Au = P_N f$, $N \in \mathbb{N}$, имеет единственное решение $u_N \in U_N$, причем $\|Au_N - f\| \rightarrow 0$ при

$N \rightarrow \infty$. Тогда, согласно неравенству

$$\rho(f, AU_N) = \inf_{v \in AU_N} \|f - v\| \leq \|f - Au_N\|,$$

заключаем, что $\rho(f, AU_N) \rightarrow 0$ и последовательность подпространств $\{AU_N\}$ является предельно плотной в $\mathcal{H}$. Таким образом, первое условие теоремы выполняется.

Зафиксируем произвольный номер N . Поскольку $P_N \mathcal{H} = D(P_N) = V_N$, то для любого элемента $f^* \in V_N$ существует такой элемент $f \in \mathcal{H}$, что $P_N f = f^*$. Уравнение $P_N A u = f^*$ эквивалентно уравнению $P_N A u = P_N f$, а потому, согласно предположению, имеет решение $u_N \in U_N$, и притом единственное. Это означает, что $P_N A U_N = V_N$, т.е. выполняется второе условие теоремы.

Докажем, что выполняется и последнее, третье условие теоремы. Обозначим через $\tilde{P}_N$ сужение оператора P_N на подпространство AU_N . Поскольку линейный оператор $\tilde{P}_N$ ограничен и взаимно однозначно отображает банахово пространство AU_N на банахово пространство V_N , то, согласно *теореме Банаха об обратном операторе*, существует ограниченный обратный оператор $\tilde{P}_N^{-1}: V_N \rightarrow AU_N$. Покажем, что последовательность норм $\|\tilde{P}_N^{-1}\|$ этих операторов ограничена. Пусть $u_N \in U_N$ -- решение операторного уравнения $P_N A u = P_N f$. Тогда $Au_N = \tilde{P}_N^{-1} P_N f$, и условие $\|Au_N - f\| \rightarrow 0$ при $N \rightarrow \infty$ означает, что $\tilde{P}_N^{-1} P_N f - f \rightarrow 0$ при $N \rightarrow \infty$. Следовательно, $\tilde{P}_N^{-1} P_N f \rightarrow f$ при $N \rightarrow \infty$. Последнее условие верно для любого элемента $f \in \mathcal{H}$. Согласно *теореме Банаха -- Штейнгауза*, последовательность $\{\|\tilde{P}_N^{-1} P_N\|\}$ ограничена, т.е. для некоторого числа $C^* > 0$ имеем $\|\tilde{P}_N^{-1} P_N\| \leq C^*$, $n \in \mathbb{N}$. Поэтому для любого элемента $v \in V_N$, учитывая, что $P_N v = v$, имеем

$$\|P_N^{-1} v\| = \|P_N^{-1} P_N v\| \leq \|P_N^{-1} P_N\| \|v\| \leq C^* \|v\|.$$

Из этих неравенств вытекает, что $\|P_N^{-1}\| \leq C^*$, $N \in \mathbb{N}$.

Выбрав произвольный элемент $f \in AU_N$ и положив $v = P_N f$, получим

$$\|f\| = \|\tilde{P}_N^{-1} v\| \leq C^* \|v\| = C^* \|P_N f\|,$$

т.е. третье условие теоремы выполнено с $\tau = 1/C^*$.

Достаточность. Согласно теореме 6.11, из второго и третьего условий теоремы 6.12 вытекает, что каждое уравнение $P_N A u = P_N f$, $N \in \mathbb{N}$, имеет решение u_N , и притом единственное. Это решение можно представить в виде $u_N = A^{-1} \tilde{P}_N^{-1} P_N f$, где $\tilde{P}_N$ — сужение проектора P_N на подпространство AU_N . В соответствии с третьим условием теоремы имеем $\|\tilde{P}_N^{-1}\| \leq \frac{1}{\tau}$. Кроме того, по условиям теоремы $\|P_N\| \leq C$, $N \in \mathbb{N}$. Следовательно, $\|\tilde{P}_N^{-1} P_N\| \leq \frac{C}{\tau}$ и для любого элемента $v \in AU_N$, учитывая равенство $v = \tilde{P}_N^{-1} P_N v$, получаем

$$\begin{aligned} \|Au_N - f\| &= \|\tilde{P}_N^{-1} P_N f - f\| \leq \\ &\leq \|\tilde{P}_N^{-1} P_N f - \tilde{P}_N^{-1} P_N v\| + \|v - f\| \leq \\ &\leq \|\tilde{P}_N^{-1} P_N\| \|f - v\| + \|v - f\| = \\ &= (\|\tilde{P}_N^{-1} P_N\| + 1) \|f - v\| \leq \left(\frac{C}{\tau} + 1\right) \|f - v\|. \end{aligned}$$

Так как элемент $v \in AU_N$ выбирался произвольно, то

$$\|Au_N - f\| \leq \left(\frac{C}{\tau} + 1\right) \inf_{v \in AU_N} \|f - v\| = \left(\frac{C}{\tau} + 1\right) \rho(f, AU_N).$$

Согласно первому условию теоремы $\rho(f, AU_N) \rightarrow 0$, при $N \rightarrow \infty$. Поэтому и $\|Au_N - f\| \rightarrow 0$ при $N \rightarrow \infty$. ►

Замечание 6.3. Если U_N и V_N — конечномерные линейные пространства одинаковой размерности, то второе условие теоремы 6.12 является следствием ее третьего условия. Действительно, из третьего условия вытекает, что все операторы P_N имеют нулевое ядро. В этом случае для каждого номера N

оператор $P_N A$, рассматриваемый на конечномерном линейном пространстве U_N , также имеет нулевое ядро. Значит, его образ имеет ту же размерность, что и линейное пространство U_N . Если $\dim U_N = \dim V_N$, то размерность подпространства $P_N A U_N$ в V_N совпадает с размерностью V_N . Но это возможно лишь при $P_N A U_N = V_N$.

Замечание 6.4. Если в условиях теоремы 6.12 оператор A имеет ограниченный обратный оператор, то для любого элемента $f \in R(A)$ последовательность $\{u_N\}$ решений операторных уравнений $P_N A u = P_N f$ сходится по норме к решению u уравнения $Au = f$. Действительно, в этом случае

$$\|u_N - u\| = \|u_N - A^{-1}f\| = \|A^{-1}(Au_N - f)\| \leq \|A^{-1}\| \|Au_N - f\|.$$

Согласно теореме 6.12, $\|Au_N - f\| \rightarrow 0$ при $N \rightarrow \infty$. Поэтому и $\|u_N - u\| \rightarrow 0$ при $N \rightarrow \infty$. #

Рассмотрим проекционный метод в частном случае, когда выбор последовательностей подпространств $\{U_N\}$ и $\{V_N\}$, а также последовательности проекторов $\{P_N\}$ определяется парой ортогональных базисов $\{u_n\}$ и $\{v_n\}$ в гильбертовом пространстве $\mathcal{H}$ следующим образом. Подпространства U_N и V_N являются линейными оболочками конечных ортогональных систем элементов $u_1, u_2, \dots, u_N$ и $v_1, v_2, \dots, v_N$. Оператор P_N является оператором ортогонального проектирования на V_N (или **ортпроектором**), т.е. каждому элементу $f \in \mathcal{H}$ с разложением $f = \sum_{k=1}^{\infty} f_k v_k$ в соответствие ставится его проекция f_N

на V_N , равная $f_N = \sum_{k=1}^N f_k v_k$, т.е. $Pf = f_N$.

Отметим, что в рассматриваемом случае $\|P_N\| = 1$, $N \in \mathbb{N}$, и, следовательно, последовательность проекторов P_N равномерно ограничена. Приближенное операторное уравнение $P_N A u = P_N f$, решение которого ищется в конечномерном линейном пространстве U_N , равносильно уравнению $P_N(Au - f) = 0$. Но

равенство $P_N v = 0$ означает, что все коэффициенты разложения элемента v в ортогональном базисе $\{v_k\}$ с номерами $k \leq N$ равны нулю, т.е. $\langle v, v_k \rangle = 0$, $k = \overline{1, N}$. Таким образом, решение $u_N \in U_N$ операторного уравнения $P_N A u = P_N f$ удовлетворяет системе уравнений $\langle A u - f, v_k \rangle = 0$, $k = \overline{1, N}$. Заменяя элемент $u \in U_N$ его разложением $u = \sum_{n=1}^N a_n u_n$ в базисе $v_1, v_2, \dots, v_N$ этого линейного пространства, приходим к системе линейных уравнений относительно неизвестных коэффициентов разложения $a_1, a_2, \dots, a_N$:

$$\sum_{n=1}^N \langle A u_n, v_k \rangle a_n = \langle f, v_k \rangle, \quad k = \overline{1, N}. \quad (6.183)$$

Нетрудно увидеть, что в рассматриваемом случае проекционный метод совпал с *методом ортогональных проекций*, а система (6.183) не отличается от системы линейных уравнений (6.80), к которой приводит метод ортогональных проекций. Это позволяет рассматривать проекционный метод как обобщение метода ортогональных проекций, а для исследования метода ортогональных проекций использовать теоремы 6.11 и 6.12, дополнительно учитывая замечания 6.3 и 6.4.

Вопросы и задачи

6.1. Одномерное установившееся распределение температуры $T(x)$ в стенке толщиной h удовлетворяет обыкновенному дифференциальному уравнению

$$\frac{d}{dx} \left(F(x) \frac{dT(x)}{dx} \right) = 0, \quad x \in [0, h],$$

где

$$F(x) = F_0 \left(1 + \frac{x}{R_1} \right) \left(1 + \frac{x}{R_2} \right), \quad x \in [0, h],$$

а R_1 и R_2 — главные радиусы кривизны одной из поверхностей стенки. На этой поверхности выбраны участок площадью F_0 и начало отсчета координатной оси Ox , направленной по нормали к поверхностям стенки, на которых заданы значения температур $T(0) = T_0$ и $T(h) = T_h$. При условиях $\frac{h}{|R_1|} \ll 1$ и $\frac{h}{|R_2|} \ll 1$ методом малого параметра найти распределение $T(x)$ и сравнить его с точным решением задачи.

При тех же условиях методом малого параметра решить одномерную нестационарную краевую задачу

$$F(x) \frac{\partial T(x, t)}{\partial t} = \tilde{a} \frac{\partial}{\partial x} \left(F(x) \frac{\partial T(x, t)}{\partial x} \right), \quad x \in [0, h], \quad t > 0,$$

$$T(x, 0) = T(0, t) = T_0, \quad T(h, t) = T_h,$$

где t — время, $\tilde{a}$ — температуропроводность (см. 2.3) материала стенки.

6.2. Упругая балка длиной l с постоянной жесткостью на изгиб EI_* (см. рис. 6.3) закреплена таким образом, что ее прогиб $w(x)$ удовлетворяет условиям $w(0) = w'(0) = w(l) = w''(l) = 0$. Найти форму прогиба оси балки под действием распределенной нагрузки интенсивностью $q(x) = q_0 \left(1 + \frac{x}{l}\right)$ методами коллокации в подобластях и в точках и сравнить результаты при одинаковом числе слагаемых в представлении приближенного решения, а также провести сравнение с точным решением задачи.

6.3. Методами коллокации в подобластях и в точках найти одномерное стационарное распределение температуры $T(r)$, $r \in [R_1, R_2]$, по толщине стенки трубы с внутренним R_1 и наружным R_2 радиусами, описываемое обыкновенным дифференциальным уравнением

$$\frac{d^2 T(r)}{dr^2} + \frac{1}{r} \frac{dT(r)}{dr} + \frac{I_V^{(q)} r^2}{\lambda R_2^2} = 0,$$

где $I_V^{(q)}$ — объемная мощность энергоснабжения (см. 2.3) в стенке при $r = R_2$; λ — теплопроводность (см. 1.3) материала трубы. На поверхностях трубы принять $T(R_1) = T(R_2) = T_0$. Результаты сравнить с точным решением задачи.

6.4. Методами коллокации в подобластях и в точках найти одномерное установившееся распределение температуры $T(x)$ в стенке с поверхностью двоякой кривизны (см. задачу 6.1).

6.5. Найти решение уравнения (6.85) с граничными условиями $u(0) = u(1) = 0$ методом наименьших квадратов и сравнить с решением, полученным в примере 6.8.

6.6. Вязкая несжимаемая жидкость движется в трубе квадратного поперечного сечения, причем вектор $\mathbf{v}$ скорости жидкости направлен вдоль оси Ox_3 трубы, т.е. его проекции на координатные оси Ox_1 и Ox_2 , перпендикулярные стенкам трубы, равны нулю ($v_1 = v_2 = 0$). Методом наименьших квадратов и методом Бубнова — Галеркина найти распределение в поперечном сечении трубы проекции $v_3(x_1, x_2)$ вектора скорости жидкости на ось Ox_3 . Функция $v_3(x_1, x_2)$ удовлетворяет уравнению

$$\eta \left(\frac{\partial^2 v_3}{\partial x_1^2} + \frac{\partial^2 v_3}{\partial x_2^2} \right) - \frac{\partial p}{\partial x_3} = 0,$$

где η — коэффициент сдвиговой вязкости жидкости (см. 3.2); $p = p(x_3)$ — заданная функция изменения давления жидкости вдоль оси трубы. На стенках трубы принять $v_3 = 0$. Сравнить результаты с точным решением задачи.

6.7. Решить задачу 6.1 методом наименьших квадратов и методом Бубнова — Галеркина и сравнить результаты с точным решением.

6.8. Решить задачу 6.2 при условиях $w(0) = w'(0) = w(l) = w'(l) = 0$ методом наименьших квадратов и методом Бубнова — Галеркина и сравнить результаты с точным решением.

6.9. Убедиться, что применение метода Бубнова — Галеркина для решения уравнения (6.85) с граничными условиями $u(0) = u(1) = 0$ и для решения краевой задачи (6.84) дает одинаковый результат.

6.10. Найти двустороннюю оценку критической силы, вызывающей потерю устойчивости прямолинейной формы равновесия стержня с круговым поперечным сечением, рассмотренного в примерах 6.11 и 6.12, если

$$r(x) = a \sqrt[4]{1 + \frac{mx(l-x)}{l^2}}, \quad m = 4.$$

ЧАСТЬ III

Сеточные методы

7. ОСНОВЫ МЕТОДА КОНЕЧНЫХ РАЗНОСТЕЙ

Используемые в задачах математической физики модели физических процессов (см. 1–3) предполагают в большинстве случаев непрерывность распределения искомых величин в пространстве и их непрерывное изменение во времени. Вместе с тем можно получить некоторое приближенное представление о пространственном распределении и эволюции во времени этих величин, если оперировать совокупностью их значений в фиксированные моменты времени на конечном множестве точек пространства. Ясно, что уменьшение интервалов между выбранными фиксированными моментами времени и сокращение расстояний между выбранными точками пространства должны приближать такое дискретное представление к непрерывному распределению искомых величин.

7.1. Понятие о сеточных методах

Множество точек пространства, используемых для приближенного представления непрерывного пространственного распределения какой-либо величины, называют **пространственной сеткой**, а точки — **узлами** (или узловыми точками) этой **сетки**. Аналогично множество фиксированных моментов времени называют **временной сеткой**, а такие моменты времени — узлами этой сетки. Объединение пространственных сеток, рассматриваемых в выбранные фиксированные моменты времени, образует множество узлов **пространственно-временной сетки**. Множество узлов пространственной сетки в фиксированный момент времени называют **слоем простран-**

ственно-временной сетки. Значение величины в узле сетки называют **узловым**.

При необходимости значения величин в промежутках между узлами сетки можно найти *интерполированием*. Это позволяет получить по дискретной информации об искомым величинах их приближенные непрерывные зависимости от пространственных координат и времени.

При рассмотрении пространственного распределения искомым величин их зависимость может быть существенной не от всех трех пространственных координат. Тогда наряду с общим случаем **трехмерной сетки** в частных случаях она может быть **двумерной** или даже **одномерной**. В стационарных задачах искомые величины не зависят от времени. Поэтому необходимость в использовании временной сетки при решении таких задач отпадает.

Понятия сетки и сеточного узла являются основными при построении большой группы приближенных методов решения задач математической физики, называемых **сеточными методами** (иногда используют собирательный термин — **метод сеток**). В таких методах непрерывное пространственное распределение искомым величин и описание их непрерывного изменения во времени представляют совокупностью их значений в узлах пространственно-временной сетки. При этом производные искомым функций, входящие в дифференциальные уравнения математической физики и краевые условия, приближенно заменяют (аппроксимируют) в каждом узле **конечными разностями**. В итоге исходную математическую формулировку задачи сводят к системе уравнений (в общем случае нелинейных) относительно неизвестных узловых значений. Такие **уравнения** называют **разностными**, а их систему вместе с правилами их построения называют **разностной схемой**. Одной и той же краевой задаче могут соответствовать различные разностные схемы. В случае линейной задачи разностная схема включает систему линейных алгебраических уравнений (СЛАУ).

Описанный подход приводит к одному из наиболее широко применяемых вариантов метода сеток — *методу конечных разностей* (МКР) приближенного решения задач математической физики. Но математическая формулировка таких задач может и не содержать дифференциальных уравнений, а включать *интегральные уравнения* (в общем случае — *интегро-дифференциальные*) или *функционалы*, в которых искомые функции входят в подынтегральное выражение [XV]. В таких случаях узлы пространственно-временной сетки используют для построения *квадратурных формул*, что позволяет приближенно заменить интегралы соответствующими *квадратурными суммами*, содержащими узловые значения искомых функций. В итоге метод сеток также приводит к системе уравнений относительно неизвестных узловых значений.

Отметим, что группу соседних узлов пространственно-временной сетки можно использовать для построения непрерывной функции, имеющей так называемый конечный носитель (например, являющейся *интерполяционным многочленом*, принимающим некоторое значение в фиксированном узле и нулевое значение во всех соседних). Из таких функций, построенных для каждого узла сетки с конечным числом узлов, можно составить базис конечномерного *функционального пространства*, в котором применимы *проекционные методы* приближенного решения задач математической физики. При таком сочетании эти *методы* иногда называют *проекционно-сеточными**.

В частности, подобный подход приводит к *методам конечных* или *граничных элементов*, которые также обычно относят к группе сеточных методов. При этом под элементом в общем случае понимают подобласть пространственно-временной области, содержащую группу соседних узлов соответствующей сетки, используемую для построения упомянутой непрерывной функции, т.е. конечный или граничный элемент является конечным носителем этой функции.

*См., например: Марчук Г.И., Агошков В.И.

7.2. Аппроксимация производных конечными разностями

Можно выделить два основных способа построения приближенных формул для производных функции по ее значениям в узлах пространственно-временной сетки. Такие формулы принято называть аппроксимирующими, а замену ими производных в уравнениях — *аппроксимацией производных*.

Рассмотрим эти способы применительно к действительной функции $u(x)$ одного переменного x . Как и в случае численного дифференцирования, они основаны на использовании *интерполяционных многочленов* или на применении формулы Тейлора [II]. Отличие состоит в том, что при численном дифференцировании известны *узловые значения* функции, а при аппроксимации производных эти значения еще предстоит найти в процессе приближенного решения задачи.

Пусть u_n — значения функции $u(x)$ в узлах x_n , $n = \overline{1, N}$, *одномерной сетки*. Известно [II], что для функции $u(x)$ можно построить *интерполяционный многочлен Лагранжа*

$$\tilde{u}_N(x) = \sum_{n=1}^N u_n L_n(x), \quad L_n(x) = \prod_{m=1, m \neq n}^N \frac{x - x_m}{x_n - x_m}, \quad (7.1)$$

степени $N - 1$, принимающий в узлах значения u_n . При этом для значений x , не совпадающих с узлами, погрешность на равномерной сетке с расстоянием h между соседними узлами (*шагом сетки*) пропорциональна h^n . Дифференцированием (7.1) можно получить $u^{(k)}(x) \approx \tilde{u}_N^{(k)}(x)$, $x \in (x_1, x_N)$, $k = \overline{1, N-1}$, но погрешность этой приближенной формулы оценить сложно.

Погрешность выражений, аппроксимирующих производные функции $u(x)$, называемую *погрешностью аппроксимации*, удобно оценивать, используя второй способ построения приближенных выражений для производных путем представления этой функции формулой Тейлора в окрестности фиксированно-

го узла x_n :

$$u(x) = u_n + u'_n(x - x_n) + u''_n \frac{(x - x_n)^2}{2} + \dots + u_n^{(k)} \frac{(x - x_n)^k}{k!} + \dots \\ \dots + u_n^{(m)} \frac{(x - x_n)^m}{m!} + O((x - x_n)^{m+1}), \quad (7.2)$$

где $u_n^{(k)} = u^{(k)}(x)|_{x=x_n}$, $k = \overline{0, m}$, причем $u_n^{(0)} = u_n = u(x_n)$. Слагаемое $O((x - x_n)^{m+1})$ указывает, что погрешность представления функции $u(x) \in C^m[x_1, x_N]$ при помощи (7.2) имеет $(m+1)$ -й порядок малости при $x \rightarrow x_n$.

Этим способом установлено [II], что для первой производной

$$u'_n = \frac{u_{n+1} - u_n}{h} + O(h) = \frac{u_n - u_{n-1}}{h} + O(h) = \frac{u_{n+1} - u_{n-1}}{2h} + O(h^2),$$

а для второй производной

$$u''_n = \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + O(h^2), \quad (7.3)$$

где $h = |x_{n+1} - x_n| = \text{const}$. Таким образом, погрешность аппроксимации первой производной *правой и левой конечными разностями* имеет первый порядок малости при $h \rightarrow 0$, а *центральной конечной разностью* — второй порядок малости. Такой же порядок малости имеет погрешность аппроксимации в (7.3). Для краткости обычно говорят просто о **порядке погрешности аппроксимации**, опуская слово „малости“. Отметим, что правую и левую конечные разности (или *конечные разности вперед и назад* соответственно) можно рассматривать как центральные, но при аппроксимации первой производной в **промежуточных узлах** $x_{n\pm 1/2} = x_n \pm h/2$, т.е.

$$u'_{n+1/2} = \frac{u_{n+1} - u_n}{h} + O(h^2), \quad u'_{n-1/2} = \frac{u_n - u_{n-1}}{h} + O(h^2). \quad (7.4)$$

Порядок погрешности аппроксимации производных можно повысить, построив конечно-разностные соотношения, содержащие большее число узловых значений функции $u(x)$ [11].

В случае сетки с переменным шагом между соседними узлами, используя (7.4), приближенно получаем

$$u_n'' \approx \frac{u'_{n+1/2} - u'_{n-1/2}}{h_n} \approx \frac{1}{h_n} \left(\frac{u_{n+1} - u_n}{h_{n+1/2}} + \frac{u_{n-1} - u_n}{h_{n-1/2}} \right), \quad (7.5)$$

где $h_n = \frac{1}{2}(x_{n+1} - x_{n-1})$, $h_{n\pm 1/2} = |x_{n\pm 1} - x_n|$. Для оценки порядка возникающей в (7.5) погрешности сначала при помощи представления (7.2) функции $u(x)$ в окрестности узла x_n запишем единым выражением значения этой функции в узлах $x_{n\pm 1}$:

$$\begin{aligned} u_{n\pm 1} = u_n \pm u_n' h_{n\pm 1/2} + u_n'' \frac{h_{n\pm 1/2}^2}{2} \pm u_n''' \frac{h_{n\pm 1/2}^3}{6} + \\ + u_n^{iv} \frac{h_{n\pm 1/2}^4}{24} \pm u_n^v \frac{h_{n\pm 1/2}^5}{120} + u_n^{vi} \frac{h_{n\pm 1/2}^6}{720} \pm \dots \end{aligned} \quad (7.6)$$

Затем, разделив равенство (7.6) с верхними знаками „+“ в индексах и перед слагаемыми на $h_{n+1/2}$, а равенство с нижними знаками „-“ на $h_{n-1/2}$ и сложив почленно результаты, получим

$$\begin{aligned} u_n'' = \frac{1}{h_n} \left(\frac{u_{n+1} - u_n}{h_{n+1/2}} + \frac{u_{n-1} - u_n}{h_{n-1/2}} \right) - u_n''' \frac{h_{n+1/2} - h_{n-1/2}}{3} - \\ - u_n^{iv} \frac{h_{n+1/2}^2 - h_{n+1/2} h_{n-1/2} + h_{n-1/2}^2}{12} - \dots \end{aligned} \quad (7.7)$$

Таким образом, аппроксимация (7.5) имеет первый порядок погрешности. Нетрудно убедиться, что при постоянном шаге сетки $h_n = h_{n\pm 1/2} = h$ (7.7) переходит в (7.3).

Чтобы аппроксимировать в узле x_n неравномерной сетки четвертую производную функции $u(x)$, в соответствии с (7.5)

можно записать

$$u_n^{iv} \approx \frac{1}{h_n} \left(\frac{u''_{n+1} - u''_n}{h_{n+1/2}} + \frac{u''_{n-1} - u''_n}{h_{n-1/2}} \right),$$

а затем при помощи (7.5) выразить вторые производные в узлах x_n и $x_{n\pm 1}$:

$$\begin{aligned} u_n^{iv} \approx & \frac{1}{h_n h_{n+1/2} h_{n+1}} \left(\frac{u_{n+2} - u_{n+1}}{h_{n+3/2}} + \frac{u_n - u_{n+1}}{h_{n+1/2}} \right) - \\ & - \frac{1}{h_n^2 h_{n+1/2}} \left(\frac{u_{n+1} - u_n}{h_{n+1/2}} - \frac{u_{n-1} - u_n}{h_{n-1/2}} \right) + \\ & + \frac{1}{h_n h_{n-1/2} h_{n-1}} \left(\frac{u_{n-2} - u_{n-1}}{h_{n-3/2}} + \frac{u_n - u_{n-1}}{h_{n-1/2}} \right) - \\ & - \frac{1}{h_n^2 h_{n-1/2}} \left(\frac{u_{n+1} - u_n}{h_{n+1/2}} - \frac{u_{n-1} - u_n}{h_{n-1/2}} \right), \quad (7.8) \end{aligned}$$

где $h_{n\pm 1} = \frac{1}{2}|x_{n\pm 2} - x_n|$ и $h_{n\pm 3/2} = |x_{n\pm 2} - x_{n\pm 1}|$. В случае равномерной сетки ($h_{n\pm 1/2} = h_{n\pm 3/2} = h_n = h = \text{const}$) получим

$$u_n^{iv} \approx \frac{u_{n+2} - 4u_{n+1} + 6u_n - 4u_{n-1} + u_{n-2}}{h^4}. \quad (7.9)$$

Можно показать, что погрешность аппроксимации в (7.8) имеет первый порядок, а в (7.9) — второй порядок.

7.3. Метод баланса

В дифференциальные уравнения задач математической физики, вытекающие из законов сохранения физических субстанций и описывающие процессы в неоднородной среде, часто входят выражения вида

$$(p(x)u'(x))', \quad (r(x)u''(x))'' \quad (7.10)$$

и аналогичные им, причем функции $p(x)$ и $r(x)$ в своих областях определения могут иметь точки разрыва первого рода. Но даже в случае дифференцируемости этих функций при аппроксимации таких выражений нецелесообразно предварительным дифференцированием выделять в них явно старшую производную функции $u(x)$. Дело в том, что при этом в разностном виде будут нарушены условия баланса, соответствующие закону сохранения рассматриваемой физической субстанции, или условия равновесия действующих нагрузок. Кроме того, функции $v(x) = -p(x)u'(x)$ и $w(x) = r(x)u''(x)$ имеют определенный физический смысл (например, плотности теплового потока, если u — температура, а p — теплопроводность среды, или изгибающего момента, если u — прогиб балки, а r — ее жесткость на изгиб), так что эти функции также желательно аппроксимировать на *одномерной сетке* (обычно в ее *промежуточных узлах*).

Рассмотрим аппроксимацию первого выражения в (7.10) в узле $x_n \in [x_{n-1}, x_{n+1}]$ сетки с переменным (в общем случае) шагом, предполагая, что функция $p(x)$ на отрезке $[x_{n-1}, x_{n+1}]$ положительна и имеет конечное число точек разрыва первого рода, но функция $v(x) = -p(x)u'(x)$ непрерывна на этом отрезке и непрерывно дифференцируема в интервалах между точками разрыва. Сначала проинтегрируем равенство $u'(x) = -\frac{v(x)}{p(x)}$ на отрезке $[x_n, x_{n+1}]$:

$$\int_{x_n}^{x_{n+1}} \frac{du(x)}{dx} dx = u_{n+1} - u_n = - \int_{x_n}^{x_{n+1}} \frac{v(x)}{p(x)} dx$$

и приближенно примем, что $v(x) \approx v_{n+1/2} = v(x_{n+1/2}) = \text{const}$ при $x \in [x_n, x_{n+1}]$. Тогда получим

$$u_{n+1} - u_n \approx -v_{n+1/2} \int_{x_n}^{x_{n+1}} \frac{dx}{p(x)},$$

и отсюда

$$v_{n+1/2} \approx -\tilde{p}_{n+1/2} \frac{u_{n+1} - u_n}{h_{n+1/2}}, \quad (7.11)$$

где

$$\frac{1}{\tilde{p}_{n+1/2}} = \frac{1}{h_{n+1/2}} \int_{x_n}^{x_{n+1}} \frac{dx}{p(x)}.$$

Аналогично находим

$$v_{n-1/2} \approx -\tilde{p}_{n-1/2} \frac{u_n - u_{n-1}}{h_{n-1/2}}, \quad (7.12)$$

где

$$\frac{1}{\tilde{p}_{n-1/2}} = \frac{1}{h_{n-1/2}} \int_{x_{n-1}}^{x_n} \frac{dx}{p(x)}.$$

Используя *центральную конечную разность* в узле x_n , с учетом (7.11) и (7.12) получаем

$$\begin{aligned} (p(x) u'(x))' \Big|_{x=x_n} &= -v'(x) \Big|_{x=x_n} \approx -\frac{v_{n+1/2} - v_{n-1/2}}{h_n} \approx \\ &\approx \frac{1}{h_n} \left(\tilde{p}_{n+1/2} \frac{u_{n+1} - u_n}{h_{n+1/2}} + \tilde{p}_{n-1/2} \frac{u_{n-1} - u_n}{h_{n-1/2}} \right). \end{aligned} \quad (7.13)$$

Описанную процедуру построения аппроксимации левой части (7.13) называют **методом баланса** (иногда интегро-интерполяционным методом), поскольку правая часть (7.13) позволяет в разностной форме выполнить условия баланса, соответствующие тому или иному закону сохранения. Если функция $u(x)$ имеет смысл температуры, а $p(x)$ — теплопроводности среды, то отношения $\frac{h_{n\pm 1/2}}{\tilde{p}_{n\pm 1/2}}$ в (7.11) и (7.12) равны суммарным термическим сопротивлениям слоев неоднородной среды, расположенных между соседними узлами, а выражение в скобках в правой части (7.13) является суммой притекающих в узел x_n

тепловых потоков. Если же $u(x)$ трактовать как распределение электрического потенциала, а $p(x)$ — как электрическую проводимость, то в скобках в правой части (7.13) получим сумму электрических токов, притекающих в этот узел (отношения $\frac{h_{n\pm 1/2}}{\tilde{p}_{n\pm 1/2}}$ в (7.11) и (7.12) будут суммарными электрическими сопротивлениями слоев среды между соседними узлами). Аналогичную трактовку имеет (7.11)–(7.13) применительно и к другим физическим процессам.

Оценка *порядка погрешности*, возникающей при аппроксимации левой части (7.13), в общем случае довольно сложна*. В частном случае $p(x) = p_0 = \text{const}$ имеем $\tilde{p}_{n\pm 1/2} = p_0$ и аппроксимация (7.13) равносильна (7.5), т.е. имеет первый порядок погрешности, который возрастает на единицу при $h_{n\pm 1/2} = h = \text{const}$, поскольку тогда аппроксимация (7.13) будет равносильна (7.3).

Если функция $p(x)$ непрерывна, то можно получить аналог формулы (7.13), заменив в ней $\tilde{p}_{n\pm 1/2}$ на $p_{n\pm 1/2} = p(x_{n\pm 1/2})$. Оценим возникающий при такой аппроксимации порядок погрешности, полагая, что функции $u(x)$ и $p(x)$ непрерывно дифференцируемы необходимое число раз на отрезке $x \in [x_{n-1}, x_{n+1}]$. Для этого сначала при помощи представления вида (7.2) функции $v(x) = -p(x)u'(x)$ в окрестности узла x_n запишем единым выражением значения этой функции, вычисленные в точках $x_{n\pm 1/2} = x_n \pm h_{n\pm 1/2}/2$:

$$\begin{aligned} p(x)u'(x) \Big|_{x=x_{n\pm 1/2}} &= p(x)u'(x) \Big|_{x=x_n} \pm \\ &\pm (p(x)u'(x))' \Big|_{x=x_n} \frac{h_{n\pm 1/2}}{2} + \frac{1}{2} (p(x)u'(x))'' \Big|_{x=x_n} \frac{h_{n\pm 1/2}^2}{4} \pm \\ &\pm \frac{1}{6} (p(x)u'(x))''' \Big|_{x=x_n} \frac{h_{n\pm 1/2}^3}{8} + \dots \end{aligned}$$

*См., например: Самарский А.А.

Затем после почленного вычитания из равенства с верхними знаками „+“ в индексах и перед слагаемыми равенства с нижними знаками „-“ получим

$$\begin{aligned} (p(x) u'(x))' \Big|_{x=x_n} &= \frac{p(x) u'(x) \Big|_{x=x_{n+1/2}} - p(x) u'(x) \Big|_{x=x_{n-1/2}}}{h_n} - \\ &- (p(x) u'(x))'' \Big|_{x=x_n} \frac{h_{n+1/2}^2 - h_{n-1/2}^2}{4(h_{n+1/2} + h_{n-1/2})} - \\ &- (p(x) u'(x))''' \Big|_{x=x_n} \frac{h_{n+1/2}^3 + h_{n-1/2}^3}{24(h_{n+1/2} + h_{n-1/2})} + O(\tilde{h}_n^3), \quad (7.14) \end{aligned}$$

где $\tilde{h}_n = \max\{h_{n-1/2}, h_{n+1/2}\}$.

Теперь при помощи разложения функции $u(x)$ в ряд Тейлора в окрестности точек $x_{n\pm 1/2}$ запишем значения этой функции, вычисленные в узлах x_n и $x_{n\pm 1}$:

$$\begin{aligned} u_n = u_{n\pm 1/2} \mp u'(x) \Big|_{x=x_{n\pm 1/2}} \frac{h_{n\pm 1/2}}{2} + \frac{1}{2} u''(x) \Big|_{x=x_{n\pm 1/2}} \frac{h_{n\pm 1/2}^2}{4} \mp \\ \mp \frac{1}{6} u'''(x) \Big|_{x=x_{n\pm 1/2}} \frac{h_{n\pm 1/2}^3}{8} + \frac{1}{24} u^{IV}(x) \Big|_{x=x_{n\pm 1/2}} \frac{h_{n\pm 1/2}^4}{16} \mp \dots \\ u_{n\pm 1} = u_{n\pm 1/2} \pm u'(x) \Big|_{x=x_{n\pm 1/2}} \frac{h_{n\pm 1/2}}{2} + \frac{1}{2} u''(x) \Big|_{x=x_{n\pm 1/2}} \frac{h_{n\pm 1/2}^2}{4} \pm \\ \pm \frac{1}{6} u'''(x) \Big|_{x=x_{n\pm 1/2}} \frac{h_{n\pm 1/2}^3}{8} + \frac{1}{24} u^{IV}(x) \Big|_{x=x_{n\pm 1/2}} \frac{h_{n\pm 1/2}^4}{16} \pm \dots \end{aligned}$$

Вычитая почленно из последнего равенства предыдущее, получим

$$\pm u'(x) \Big|_{x=x_{n\pm 1/2}} = \frac{u_{n\pm 1} - u_n}{h_{n\pm 1/2}} \mp u'''(x) \Big|_{x=x_{n\pm 1/2}} \frac{h_{n\pm 1/2}^2}{24} + O(\tilde{h}_n^4).$$

Если умножить это равенство на значение $p_{n\pm 1/2} = p(x_{n\pm 1/2})$ и учесть, что

$$p(x) u'''(x) \Big|_{x=x_{n\pm 1/2}} = p(x) u'''(x) \Big|_{x=x_n} \pm \\ \pm (p(x) u'''(x))' \Big|_{x=x_n} \frac{h_{n\pm 1/2}}{2} + O(\tilde{h}_n^2),$$

то получим

$$\pm p(x) u'(x) \Big|_{x=x_{n\pm 1/2}} = p_{n\pm 1/2} \frac{u_{n\pm 1} - u_n}{h_{n\pm 1/2}} \mp p(x) u'''(x) \Big|_{x=x_n} \frac{h_{n\pm 1/2}^2}{24} - \\ - (p(x) u'''(x))' \Big|_{x=x_n} \frac{h_{n\pm 1/2}^3}{48} + O(\tilde{h}_n^4).$$

После почленного сложения строк с верхним и нижним знаками соответственно и подстановки результата в (7.14) находим

$$(p(x) u'(x))' \Big|_{x=x_n} = \frac{1}{h_n} \left(p_{n+1/2} \frac{u_{n+1} - u_n}{h_{n+1/2}} + p_{n-1/2} \frac{u_{n-1} - u_n}{h_{n-1/2}} \right) - \\ - \left((p(x) u'(x))'' + \frac{p(x)}{3} u'''(x) \right) \Big|_{x=x_n} \frac{h_{n+1/2} - h_{n-1/2}}{4} - \\ - \left((p(x) u'(x))''' + (p(x) u'''(x))' \right) \Big|_{x=x_n} \times \\ \times \frac{h_{n+1/2}^2 - h_{n+1/2} h_{n-1/2} + h_{n-1/2}^2}{24} + O(\tilde{h}_n^3). \quad (7.15)$$

Итак, при равномерно расположенных узлах ($h_{n\pm 1/2} = h_n = h = \text{const}$) погрешность аппроксимации левой части (7.15) первым слагаемым правой части (7.15) имеет второй порядок, а если узлы расположены неравномерно, то лишь первый порядок.

Теперь перейдем к аппроксимации второго выражения в (7.10) в узле $x_n \in [x_{n-2}, x_{n+2}]$ неравномерной сетки, предполагая, что функция $r(x)$ на отрезке $[x_{n-2}, x_{n+2}]$ положительна и имеет конечное число точек разрыва первого рода, но при этом

функция $w(x) = r(x) u''(x)$ непрерывна на этом отрезке и дважды непрерывно дифференцируема в интервалах между точками разрыва. Сначала проинтегрируем равенство $u''(x) = \frac{w(x)}{r(x)}$ на отрезке $[x_{n-1/2}, x_{n+1/2}]$, приближенно приняв на этом отрезке $w(x) \approx w_n = w(x_n)$:

$$\int_{x_{n-1/2}}^{x_{n+1/2}} \frac{d^2 u(x)}{dx^2} dx = u'_{n+1/2} - u'_{n-1/2} = \int_{x_{n-1/2}}^{x_{n+1/2}} \frac{w(x)}{r(x)} dx \approx w_n \int_{x_{n-1/2}}^{x_{n+1/2}} \frac{dx}{r(x)}.$$

Отсюда, используя (7.4), получаем

$$w_n \approx \frac{\tilde{r}_n}{h_n} \left(\frac{u_{n+1} - u_n}{h_{n+1/2}} + \frac{u_n - u_{n-1}}{h_{n-1/2}} \right), \quad (7.16)$$

где

$$\frac{1}{\tilde{r}_n} = \frac{1}{h_n} \int_{x_{n-1/2}}^{x_{n+1/2}} \frac{dx}{r(x)}.$$

Аналогично находим

$$w_{n+1} \approx \frac{\tilde{r}_{n+1}}{h_{n+1}} \left(\frac{u_{n+2} - u_{n+1}}{h_{n+3/2}} + \frac{u_n - u_{n+1}}{h_{n+1/2}} \right), \quad (7.17)$$

$$w_{n-1} \approx \frac{\tilde{r}_{n-1}}{h_{n-1}} \left(\frac{u_{n-2} - u_{n-1}}{h_{n-3/2}} + \frac{u_n - u_{n-1}}{h_{n-1/2}} \right). \quad (7.18)$$

Здесь

$$\frac{1}{\tilde{r}_{n+1}} = \frac{1}{h_{n+1}} \int_{x_{n+1/2}}^{x_{n+3/2}} \frac{dx}{r(x)}, \quad \frac{1}{\tilde{r}_{n-1}} = \frac{1}{h_{n-1}} \int_{x_{n-3/2}}^{x_{n-1/2}} \frac{dx}{r(x)}$$

и $h_{n\pm 3/2} = |x_{n\pm 2} - x_{n\pm 1}|$, $h_{n\pm 1} = \frac{1}{2}|x_{n\pm 2} - x_n|$, $x_{n\pm 3/2} = x_{n\pm 1} \pm \frac{1}{2}h_{n\pm 3/2}$. В итоге, используя (7.5), получаем

$$(r(x) u''(x))'' \Big|_{x=x_n} \approx \frac{1}{h_n} \left(\frac{w_{n+1} - w_n}{h_{n+1/2}} + \frac{w_{n-1} - w_n}{h_{n-1/2}} \right). \quad (7.19)$$

Аппроксимация производных при помощи метода баланса приводит к единообразным выражениям вне зависимости от того, как расположены возможные точки разрыва функций $p(x)$ и $r(x)$ в (7.10). В случае неоднородной среды использование метода баланса обычно позволяет ограничиться применением равномерной сетки с постоянным шагом.

7.4. Пример простейшей разностной схемы

В одномерных стационарных краевых задачах математической физики искомые функции зависят лишь от одной пространственной координаты и не зависят от времени. В математическую формулировку таких задач входят обыкновенные дифференциальные уравнения (ОДУ) с граничными условиями. Рассмотрим построение *разностной схемы* для сравнительно простой краевой задачи, описываемой линейным ОДУ второго порядка

$$-u''(x) + q(x)u(x) = f(x), \quad x \in [0, l], \quad (7.20)$$

где $q(x), f(x) \in C[0, l]$, с граничными условиями

$$u(0) = \bar{u}_0, \quad u(l) = \bar{u}_l. \quad (7.21)$$

К задаче (7.20), (7.21) можно прийти при рассмотрении установившегося распределения температуры $u(x)$ в тонком цилиндрическом стержне длиной l , торцы которого имеют заданные значения температуры $\bar{u}_0$ и $\bar{u}_l$, а на его боковой поверхности происходит теплообмен с окружающей средой (рис. 7.1).

При этом интенсивность теплообмена задает функция $q(x) > 0$, $x \in [0, l]$, а изменение температуры $u_c(x)$ среды вдоль стержня — функция $f(x)/q(x)$.

Разобьем отрезок $[0, l]$ внутренними точками $x_n = nh$, $n = \overline{1, N-1}$, на N частичных от-

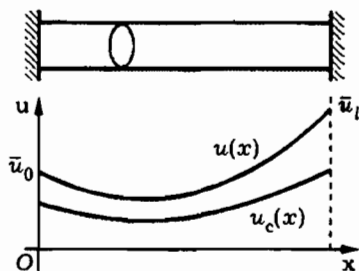


Рис. 7.1

резков равной длины $h = l/N$, т.е. введем равномерную *одномерную сетку* с номерами узлов $n = \overline{0, N}$. Для каждого внутреннего узла x_n , $n = \overline{1, N-1}$, используем *аппроксимацию* (7.3) второй *производной* $u''(x)$, имеющую второй *порядок погрешности*. Тогда из (7.20) получим систему $N-1$ *разностных уравнений*

$$-\frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} + q_n u_n = f_n, \quad n = \overline{1, N-1}, \quad (7.22)$$

относительно неизвестных *узловых значений* u_n искомой функции $u(x)$, причем $q_n = q(x_n)$, $f_n = f(x_n)$ и в соответствии с (7.21) $u_0 = \bar{u}_0$, $u_N = \bar{u}_l$.

Итак, разностная схема в данном случае состоит из равномерной одномерной сетки с $N+1$ узлами и системы (7.22) разностных уравнений при заданных значениях u_0 и u_N . Ясно, что (7.22) образует систему $N-1$ линейных алгебраических уравнений (СЛАУ) относительно узловых значений u_n , $n = \overline{1, N-1}$, *матричная запись* которой

$$Au = y \quad (7.23)$$

включает квадратную *трехдиагональную матрицу*

$$A = \begin{pmatrix} 2+q_1h^2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2+q_2h^2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2+q_3h^2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2+q_{N-2}h^2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 2+q_{N-1}h^2 \end{pmatrix}$$

порядка $N-1$ и векторы $u = (u_1 \ u_2 \ \dots \ u_n \ \dots \ u_{N-1})^T$, $y = (y_1 \ y_2 \ \dots \ y_n \ \dots \ y_{N-1})^T$, где $y_1 = f_1h^2 + u_0$, $y_{N-1} = f_{N-1}h^2 + u_N$ и $y_n = f_nh^2$, $n = \overline{2, N-2}$. Из $(N-1)^2$ элементов матрицы

А ненулевыми являются лишь $3N - 5$ и соответствуют коэффициентам СЛАУ (7.22), которую можно представить в виде

$$\begin{cases} b_1 u_1 - c_1 u_2 = y_1, \\ -a_n u_{n-1} + b_n u_n - c_n u_{n+1} = y_n, & n = \overline{2, N-2}, \\ -a_{N-1} u_{N-2} + b_{N-1} u_{N-1} = y_{N-1}, \end{cases} \quad (7.24)$$

где $a_n = c_{n-1} = 1$, $n = \overline{2, N-1}$, и $b_n = a_n + c_n + q_n h^2 \neq 0$, $n = \overline{1, N-1}$, если учесть, что $a_1 = c_{N-1} = 0$. При $q(x) \geq 0$, $x \in (0, l)$, коэффициенты в (7.24) удовлетворяют неравенствам

$$\begin{cases} b_n - a_n - c_n \geq 0, \\ a_n \geq 0, & n = \overline{1, N-1}, \\ c_n > 0, & n = \overline{1, N-2}. \end{cases} \quad (7.25)$$

В рассматриваемом случае выполнены неравенства

$$|b_n| \geq |a_n| + |c_n| > |a_n|, \quad n = \overline{1, N-1}, \quad (7.26)$$

причем $a_1 = c_{N-1} = 0$, т.е. при $n = 1$ и $n = N - 1$ первое неравенство в (7.25) является заведомо строгим. Если это неравенство является строгим хотя бы для одного значения n , будем говорить о трехдиагональной **матрице с частичным диагональным преобладанием**. В случае $q(x) > 0$, $x \in (0, l)$, все неравенства в (7.26) строгие, т.е. A — матрица с **диагональным преобладанием**. СЛАУ, у которой матрица с диагональным преобладанием, имеет решение, и притом единственное [III]. Это решение можно найти *методом Гаусса*.

Для решения СЛАУ, матрица которой имеет частичное диагональное преобладание, наиболее эффективным является *метод прогонки*, основанный на возможности выразить любое узловое значение u_n , $n = \overline{1, N-2}$, через значение u_{n+1} в соседнем

узле, т.е.

$$u_n = \mu_n u_{n+1} + \nu_n, \quad n = \overline{1, N-2}, \quad (7.27)$$

причем из первого равенства (7.24) имеем $\mu_1 = \frac{c_1}{b_1}$ и $\nu_1 = \frac{y_1}{b_1}$, а далее из (7.24) при помощи (7.27) находим

$$\mu_n = \frac{c_n}{b_n - a_n \mu_{n-1}}, \quad \nu_n = \frac{y_n + a_n \nu_{n-1}}{b_n - a_n \mu_{n-1}}, \quad n = \overline{2, N-2}. \quad (7.28)$$

Подставляя $u_{N-2} = \mu_{N-2} u_{N-1} + \nu_{N-2}$ в последнее равенство (7.24), получаем

$$u_{N-1} = \frac{y_{N-1} + a_{N-1} \nu_{N-2}}{b_{N-1} - a_{N-1} \mu_{N-2}}. \quad (7.29)$$

Это позволяет затем при помощи (7.27) и предварительно вычисленных по формуле (7.28) коэффициентов μ_n и ν_n найти остальные значения u_n , где номер n последовательно принимает значения $N-2, \dots, 1$.

Решение СЛАУ (7.24) существует и единственно, если по формулам (7.27)–(7.29), полученным из СЛАУ эквивалентными преобразованиями, можно однозначно найти неизвестные значения u_n , $n = \overline{1, N-1}$. Это будет в том случае, когда ни один из знаменателей $\Delta_n = b_n - a_n \mu_{n-1}$ в (7.28) и (7.29) в процессе вычислений не обращается в нуль. С помощью метода математической индукции покажем, что для существования и единственности решения СЛАУ достаточно выполнения условия (7.26). Действительно, $\Delta_1 = b_1 \neq 0$ и $|\mu_1| = \frac{|c_1|}{|b_1|} \leq 1$. Предположим, что $\Delta_{n-1} \neq 0$ и $|\mu_{n-1}| \leq 1$ при $2 < n \leq N-1$. Тогда получаем

$$|\Delta_n| = |b_n - a_n \mu_{n-1}| \geq |b_n| - |a_n \mu_{n-1}| \geq |b_n| - |a_n| > 0,$$

т.е. $\Delta_n \neq 0$. Кроме того, имеем $|\Delta_n| \geq |c_n|$, $n = \overline{1, N-2}$, и в данном случае $|\Delta_n| \geq |a_n|$, $n = \overline{2, N-1}$. Поэтому $|\mu_n| = \frac{|c_n|}{|\Delta_n|} \leq 1$ и $|a_n| \leq |\Delta_n|$. Следовательно, алгоритм метода прогонки, исполь-

зующий рекуррентные формулы (7.27) и (7.28), не приводит к накоплению вычислительной погрешности, связанной, например, с ошибками округления. В таком случае говорят, что алгоритм обладает **вычислительной устойчивостью**.

В рассматриваемом случае алгоритм метода прогонки обладает **устойчивостью** и **по входным данным**, поскольку возможные погрешности в задании исходной информации при формулировке краевой задачи (7.20), (7.21) не возрастают в процессе вычислений благодаря выполнению неравенств $|\mu_n| \leq 1$ и $|a_n| \leq |\Delta_n| \leq 1$. **Алгоритм** называют **устойчивым**, если он обладает одновременно и вычислительной устойчивостью и устойчивостью по входным данным, и **неустойчивым** в противном случае.

Для СЛАУ (7.24) справедлив так называемый **принцип максимума**, состоящий в том, что $u_n \leq 0$ при выполнении неравенств (7.25) и $u_0 \leq 0$, $u_N \leq 0$, $y_n \leq 0$, $n = \overline{1, N-1}$. Докажем это от противного. Предположим, что $u_n > 0$ в одном или нескольких внутренних узлах x_n , $n = \overline{1, N-2}$. Обозначим через x_m узел, в котором значение $u_m > 0$ является наибольшим. Тогда в соответствии с (7.25) имеем

$$a_m u_m \geq a_m u_{m-1}, \quad c_m u_m > c_m u_{m+1}$$

и с учетом (7.24) запишем

$$0 \leq (b_m - a_m - c_m) u_m < < b_m u_m - a_m u_{m-1} - c_m u_{m+1} = y_m \leq 0. \quad (7.30)$$

Полученное противоречие ($0 < 0$) доказывает, что $u_n \leq 0$, $n = \overline{1, N-2}$. Если же наибольшим является значение $u_{N-1} > 0$, то при $m = N-1$ также приходим к противоречию в (7.30), поскольку $a_{N-1} > 0$ и $c_{N-1} = 0$.

СЛАУ (7.24) получена эквивалентными преобразованиями из СЛАУ (7.22). Поэтому принцип максимума по отношению к (7.22) можно сформулировать так: $u_n \leq 0$ при выполнении неравенств (7.25) и $u_0 \leq 0$, $u_N \leq 0$, $f_n \leq 0$, $n = \overline{1, N-1}$.

Если в (7.22) $f_n = 0$, $n = \overline{1, N-1}$, то для узловых значений v_n , $n = \overline{1, N-1}$, являющихся решением такой СЛАУ при $v_0 = u_0$ и $v_N = u_N$, справедлива оценка

$$\max_{n=\overline{1, N-1}} |v_n| \leq \max\{|u_0|, |u_N|\} = M \geq 0. \quad (7.31)$$

Действительно, рассмотрим совокупность узловых значений $\xi_n = v_n - M$, $n = \overline{1, N-1}$, которые будут удовлетворять (7.22) при $f_n = -Mq_n h^2 \leq 0$, $n = \overline{1, N-1}$, причем $\xi_0 = v_0 - M \leq 0$ и $\xi_N = v_N - M \leq 0$. Поэтому, согласно принципу максимума, имеем $\xi_n \leq 0$, или $v_n \leq M$, $n = \overline{1, N-1}$, откуда следует (7.31).

Пусть теперь множество узловых значений w_n , $n = \overline{1, N-1}$, удовлетворяет СЛАУ (7.24) при условии $w_0 = w_N = 0$. Обозначим $Y = \max_{n=\overline{1, N-1}} |f_n|$ и рассмотрим совокупность узловых значений

$$z_n = \frac{1}{2} Y x_n (l - x_n) = \frac{1}{2} Y h^2 n (N - n) \geq 0, \quad n = \overline{1, N-1}.$$

Непосредственной проверкой убедимся, что тогда значения $\zeta_n = w_n - z_n$, $n = \overline{1, N-1}$, будут удовлетворять СЛАУ (7.24) при условии $\zeta_0 = \zeta_N = 0$, если в ее правой части f_n заменить на $f_n - Y - q_n z_n \leq 0$, $n = \overline{1, N-1}$. Следовательно, согласно принципу максимума, получим $\zeta_n \leq 0$, или

$$\begin{aligned} w_n &\leq z_n = Y \frac{x_n(l - x_n)}{2} = \\ &= \frac{x_n(l - x_n)}{2} \max_{n=\overline{1, N-1}} |f_n|, \quad n = \overline{1, N-1}. \end{aligned} \quad (7.32)$$

Отсюда с учетом $Nh = l$ находим

$$\max_{n=\overline{1, N-1}} |w_n| \leq \max_{n=\overline{1, N-1}} |z_n| \leq \frac{Y l^2}{8} = \frac{l^2}{8} \max_{n=\overline{1, N-1}} |f_n|. \quad (7.33)$$

Ясно, что при заданных значениях u_0 и u_N решение СЛАУ (7.24) можно представить в виде $u_n = v_n + w_n$. Поэтому с учетом (7.31) и (7.33) справедлива оценка

$$\begin{aligned} \max_{n=1, N-1} |u_n| &= \max_{n=1, N-1} |v_n + w_n| \leq \\ &\leq \max_{n=1, N-1} |v_n| + \max_{n=1, N-1} |w_n| \leq \\ &\leq \max\{|u_0|, |u_N|\} + \frac{l^2}{8} \max_{n=1, N-1} |f_n|. \quad (7.34) \end{aligned}$$

Теперь вывод об устойчивости по входным данным алгоритма метода прогонки при решении СЛАУ (7.22) можно перенести на соответствующую этой СЛАУ разностную схему. Пусть $\tilde{u}_n$, $n = \overline{1, N-1}$, — решение СЛАУ (7.22) при правых частях $\tilde{f}_n$ и заданных $\tilde{u}_0$, $\tilde{u}_N$. Тогда значения $\eta_n = u_n - \tilde{u}_n$, $n = \overline{1, N-1}$, будут удовлетворять СЛАУ (7.22) при правых частях $\Delta f_n = f_n - \tilde{f}_n$ и заданных $\Delta u_0 = u_0 - \tilde{u}_0$, $\Delta u_N = u_N - \tilde{u}_N$, а вместо (7.34) получим

$$\begin{aligned} \max_{n=1, N-1} |\eta_n| &= \max_{n=1, N-1} |u_n - \tilde{u}_n| \leq \\ &\leq \max\{|\Delta u_0|, |\Delta u_N|\} + \frac{l^2}{8} \max_{n=1, N-1} |\Delta f_n|. \quad (7.35) \end{aligned}$$

Значения Δf_n , $n = \overline{1, N-1}$, и Δu_0 , Δu_N можно рассматривать как погрешности при задании исходной информации для краевой задачи (7.20), (7.21). Выполнение неравенства (7.35) означает, что рассматриваемая разностная схема обладает устойчивостью по входным данным.

Перейдем к оценке погрешностей, возникающих при приближенном решении краевой задачи (7.20), (7.21). Пусть функция $\bar{u}(x)$ является точным решением этой задачи и предположим, что $\bar{u}(x)$ имеет на отрезке $[0, l]$ непрерывную производную четвертого порядка. Тогда, подставляя в разностные уравне-

ния (7.22) вместо u_n значения $\bar{u}_n = \bar{u}(x_n)$ и используя приближенный вариант (7.7) при $h_{n\pm 1/2} = h$, получим

$$-\frac{\bar{u}_{n-1} - 2\bar{u}_n + \bar{u}_{n+1}}{h^2} + q_n \bar{u}_n - f_n \approx -\bar{u}_n'' + \bar{u}_n^{\text{IV}} \frac{h^2}{12} + \\ + q_n \bar{u}_n - f_n = \bar{u}_n^{\text{IV}} \frac{h^2}{12} = \psi_n, \quad n = \overline{1, N-1}, \quad (7.36)$$

поскольку, согласно (7.20), $-\bar{u}_n'' + q_n \bar{u}_n - f_n = 0$. Значение

$$\psi = \max_{n=\overline{1, N-1}} |\psi_n| = \frac{h^2}{12} \max_{n=\overline{1, N-1}} |\bar{u}_n^{\text{IV}}| \leq \frac{M_4}{12} h^2, \quad M_4 = \max_{x \in [0, l]} |\bar{u}^{\text{IV}}(x)|,$$

назовем **погрешностью аппроксимации ОДУ** (7.20) рассматриваемой разностной схемой. Говорят, что разностное уравнение аппроксимирует ОДУ, если $\psi \rightarrow 0$ при $h \rightarrow 0$. При этом показатель степени m' в неравенстве $\psi \leq C' h^{m'}$, где $C' > 0$ — некоторая константа, называют **порядком аппроксимации**. В данном случае $m' = 2$.

Из (7.36) следует, что значения $\bar{u}_n$, $n = \overline{1, N-1}$, удовлетворяют (7.22) при замене f_n на $\bar{f}_n = f_n + \psi_n$, т.е. значения $\Delta u_n = \bar{u}_n - u_n$ удовлетворяют условиям $\Delta u_0 = \Delta u_N = 0$ и СЛАУ (7.22) при замене в ней f_n на ψ_n . Тогда в соответствии с (7.32) и (7.36) имеем

$$|\bar{u}_n - u_n| \leq x_n \frac{l - x_n}{2} \max_{n=\overline{1, N-1}} |\psi_n| \leq \\ \leq M_4 h^2 x_n \frac{l - x_n}{24}, \quad n = \overline{1, N-1}, \quad (7.37)$$

откуда следует

$$\Delta u = \max_{n=\overline{1, N-1}} |\bar{u}_n - u_n| \leq \frac{M_4 l^2}{96} h^2. \quad (7.38)$$

Значение Δu называют **погрешностью разностной схемы**. Если $\Delta u \rightarrow 0$ при $h \rightarrow 0$, то говорят, что решение разностных уравнений сходится к решению соответствующей краевой

задачи, или короче: разностная схема сходится. При этом показатель степени m в неравенстве $\Delta u \leq C'h^m$, $C > 0$, называют **порядком точности разностной схемы**. Для рассматриваемой разностной схемы $m = 2$. Совпадение порядка точности разностной схемы с порядком аппроксимации является в данном случае следствием непрерывной зависимости решения u_n , $n = \overline{1, N-1}$, разностных уравнений от их правых частей*.

Оценки (7.37) и (7.38) погрешности, называемые *априорными*, обычно трудно использовать на практике, поскольку значение M_4 сложно оценить до решения задачи. *Апостериорную оценку погрешности* можно получить методом Рунге, используя результаты решений СЛАУ (7.22) на так называемых сгущающихся сетках, т.е. на нескольких сетках с увеличивающимся числом узлов при условии сохранения ранее введенных узлов. Пусть u_n^* — значения в узлах x_n^* , $n^* = \overline{1, N^*-1}$, удовлетворяющие (7.22) при $u_0^* = \bar{u}_0$, $u_{N^*}^* = \bar{u}_1$, числе $N^* = 2N$ узлов сетки и ее шаге $h^* = h/2$. Оценивая левую часть в (7.37) сверху, т.е. заменяя неравенство (7.37) равенством, для узла $x_n = x_n^* = x_{2n}^*$ находим

$$|\bar{u}_n - u_n| = M_4 h^2 x_n \frac{l - x_n}{24}, \quad n = \overline{1, N-1},$$

$$|\bar{u}_{n^*}^* - u_{n^*}^*| = M_4 (h^*)^2 x_{n^*}^* \frac{l - x_{n^*}^*}{24} = M_4 \frac{h^2}{4} x_n \frac{l - x_n}{24},$$

где $\bar{u}_{n^*}^* = \bar{u}_n$ и $n^* = \overline{1, N^*-1}$. Отсюда получаем

$$4|\bar{u}_{n^*}^* - u_{n^*}^*| = |\bar{u}_n - u_n|,$$

или, обозначая $\Delta u_n^* = |\bar{u}_{n^*}^* - u_{n^*}^*| = |\bar{u}_n - u_n^*|$,

$$4\Delta u^*(x_n) = |\bar{u}_n - u_n| = |(\bar{u}_n - u_n^*) + u_n^* - u_n| \leq \Delta u_n^* + |u_n^* - u_n|.$$

В итоге для узла x_n получаем оценку погрешности

$$\Delta u_n^* = |\bar{u}_n - u_n^*| \leq \frac{|u_n^* - u_n|}{3} = \Delta u_n. \quad (7.39)$$

*См., например: Самарский А.А.

Пример 7.1. Рассмотрим краевую задачу

$$-u''(x) + x^2 u(x) = \left(\frac{\pi^2}{4} + x^2\right) \cos \frac{\pi x}{2}, \quad x \in (0, 1),$$

$$u(0) = 1, \quad u(1) = 0,$$

имеющую точное решение $\bar{u}(x) = \cos \frac{\pi x}{2}$. Используем разностную схему (7.22) при $q_n = x_n^2$ и $f_n = \left(\frac{\pi^2}{4} + x_n^2\right) \cos \frac{\pi x_n}{2}$, где $x_n = \frac{n}{4}$, $n = \overline{0, 4}$, т.е. разобьем отрезок $[0, 1]$ тремя внутренними точками на $N = 4$ частичных отрезков равной длины $h = \frac{1}{N} = \frac{1}{4}$. Тогда при заданных значениях $u_0 = 1$ и $u_N = 0$ получим СЛАУ из трех уравнений

$$-u_{n-1} + \left(2 + \frac{n^2}{256}\right) u_n - u_{n+1} = \frac{4\pi^2 + n^2}{256} \cos \frac{\pi n}{8}, \quad n = \overline{1, N-1}.$$

Ее решение методом прогонки представлено в табл. 7.1.

Таблица 7.1

n	u_n	n^*	$u_{n^*}^*$	$\bar{u}(x_{n^*}^*)$	Δ_n^*	Δ_n	C_n^*
0	1,000000	0	1,000000	1,000000	0	0	0
...	—	1	0,981114	0,980785	3,29	—	4,34
1	0,926080	2	0,924413	0,923880	5,33	5,56	7,43
—	—	3	0,832097	0,831470	6,27	—	9,29
2	0,709703	4	0,707733	0,707107	6,26	6,57	9,91
—	—	5	0,556116	0,555570	5,46	—	9,29
3	0,384324	6	0,383082	0,382683	3,99	4,14	7,43
—	—	7	0,195300	0,195090	2,10	—	4,34
4	0	8	0	0	0	0	0

В случае $N^* = 8$ та же разностная схема приведет к СЛАУ из семи уравнений

$$-u_{n^*-1}^* + \left(2 + \frac{(n^*)^2}{4096}\right) u_{n^*}^* - u_{n^*+1}^* = \frac{16\pi^2 + (n^*)^2}{4096} \cos \frac{\pi n^*}{16},$$

$$n^* = \overline{1, N^*-1},$$

причем $u_0^* = 1$ и $u_{N^*}^* = 0$. Решение этой СЛАУ и значения $\bar{u}(x_{n^*})$ точного решения в узлах $x_{n^*} = \frac{n^*}{N^*}$ приведены в табл. 7.1. Там же представлены значения $\Delta_n^* = 10^4 |\bar{u}(x_{n^*}) - u_{n^*}^*|$. Их сравнение со значениями $\Delta_n = 10^4 \Delta u_n$, вычисленными в соответствии с (7.39), показывает, что метод Рунге дает в данном случае хорошие результаты. Вместе с тем погрешности значения $C_n^* = 10^4 \cdot \frac{M_4}{24} (h^*)^2 x_{n^*}^* (1 - x_{n^*}^*)$, соответствующие априорной оценке (7.37), где в данном случае

$$M_4 = \max_{x \in [0,1]} |\bar{u}^{IV}(x)| = \frac{\pi^4}{16} \max_{x \in [0,1]} \left| \cos\left(\frac{\pi x}{2}\right) \right| = \frac{\pi^4}{16},$$

заметно выше (см. табл. 7.1).

Вопросы и задачи

7.1. Какой порядок погрешности имеет формула $u''_{n \pm 1} \approx \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2}$, где h — постоянный шаг одномерной сетки?

7.2. Показать, что погрешность аппроксимации в (7.8) имеет первый порядок, а в (7.9) — второй порядок.

8. ОДНОМЕРНЫЕ КРАЕВЫЕ ЗАДАЧИ

8.1. Разностные схемы для стационарных задач

Рассмотрим одномерную стационарную краевую задачу, которая включает линейное обыкновенное дифференциальное уравнение (ОДУ) второго порядка

$$-(p(x)u'(x))' + q(x)u(x) = f(x), \quad x \in [0, l], \quad (8.1)$$

с переменными коэффициентами и граничные условия

$$-\alpha_1 u'(x) \Big|_{x=0} + \alpha u(0) = \tilde{\alpha}, \quad \beta_1 u'(x) \Big|_{x=l} + \beta u(l) = \tilde{\beta}, \quad (8.2)$$

где $\alpha_1, \alpha, \tilde{\alpha}, \beta_1, \beta, \tilde{\beta}$ — действительные числа, удовлетворяющие условиям $\alpha_1^2 + \alpha^2 \neq 0$ и $\beta_1^2 + \beta^2 \neq 0$, а $p(x), q(x)$ и $f(x)$ являются заданными функциями независимого переменного x . Далее примем, что $p(x) \geq p^0 > 0$ и $q(x) \geq 0$ при $x \in [0, l]$, $\alpha_1 \alpha \geq 0$ и $\beta_1 \beta \geq 0$.

К краевой задаче (8.1), (8.2) можно прийти, изучая, например, распределение температуры $u(x)$ при стационарной теплопроводности в стержне длиной l , аналогичном изображенному на рис. 7.1, но с переменной площадью поперечного сечения, пропорциональной значениям функции $p(x)$ (функция $q(x)$ описывает изменение периметра этого сечения). На боковой поверхности стержня происходит теплообмен с окружающей средой, температура которой пропорциональна значениям функции $f(x)$, а (8.2) задают условия теплообмена на торцах стержня. При $\alpha_1 = 0$ и (или) $\beta_1 = 0$ получаем частный случай, когда на торце стержня $x = 0$ задана температура $u(0) = \tilde{\alpha}/\alpha$ и (или) на торце $x = l$ задана температура $u(l) = \tilde{\beta}/\beta$.

Если функции $p(x)$, $q(x)$ и $f(x)$ имеют на отрезке $[0, l]$ конечное число точек разрыва, то при аппроксимации ОДУ (8.1) целесообразно использовать *метод баланса*. Разобьем отрезок $[0, l]$ внутренними точками x_n , $n = \overline{1, N-1}$, на N частичных отрезков длиной $h_{n\pm 1/2} = |x_{n\pm 1} - x_n|$, т.е. введем неравномерную (в общем случае) *одномерную сетку* с номерами *узлов* $n = \overline{0, N}$. В соответствии с методом баланса вычислим интегралы от левой и правой частей (8.1) по отрезку $[x_{n-1/2}, x_{n+1/2}]$, где $x_{n\pm 1/2} = x_n \pm \frac{1}{2}h_{n\pm 1/2}$. Тогда, учитывая (7.11)–(7.13), получим систему $N - 1$ *разностных уравнений*

$$-\frac{1}{h_n} \left(\tilde{p}_{n+1/2} \frac{u_{n+1} - u_n}{h_{n+1/2}} + \tilde{p}_{n-1/2} \frac{u_{n-1} - u_n}{h_{n-1/2}} \right) + \tilde{q}_n u_n = \tilde{f}_n, \quad (8.3)$$

$n = \overline{1, N-1}$, где $h_n = \frac{1}{2}(h_{n+1/2} + h_{n-1/2})$,

$$\tilde{q}_n = \frac{1}{h_n} \int_{x_{n-1/2}}^{x_{n+1/2}} q(x) dx, \quad \tilde{f}_n = \frac{1}{h_n} \int_{x_{n-1/2}}^{x_{n+1/2}} f(x) dx, \quad (8.4)$$

а $\tilde{p}_{n\pm 1/2}$ соответствуют (7.11) и (7.12).

В общем случае *порядок погрешности аппроксимации* первого слагаемого в ОДУ (8.1), полученной в соответствии с (7.13), не превышает единицы. Поэтому производные в граничных условиях (8.2) достаточно аппроксимировать с тем же порядком погрешности, т.е. использовать *правую и левую конечные разности* соответственно:

$$-\alpha_1 \frac{u_1 - u_0}{h_{1/2}} + \alpha u_0 = \tilde{\alpha}, \quad \beta_1 \frac{u_N - u_{N-1}}{h_{N-1/2}} + \beta u_N = \tilde{\beta}. \quad (8.5)$$

Ясно, что в частных случаях $\alpha_1 = 0$ и (или) $\beta_1 = 0$ граничные условия (8.2) при $x = 0$ и (или) $x = l$ будут удовлетворены в разностной схеме точно.

Если же (8.3) аппроксимируют ОДУ со вторым порядком погрешности и $\alpha_1 \neq 0$ и (или) $\beta_1 \neq 0$, то целесообразно и (8.2) аппроксимировать также со вторым порядком погрешности. В этом случае соотношения

$$-\tilde{p}_{1/2} \frac{u_1 - u_0}{h_{1/2}} - v_0 + \tilde{q}_0 \frac{h_{1/2}}{2} u_0 = \tilde{f}_0 \frac{h_{1/2}}{2}, \quad (8.6)$$

$$\tilde{p}_{N-1/2} \frac{u_N - u_{N-1}}{h_{N-1/2}} + v_N + \tilde{q}_N \frac{h_{N-1/2}}{2} u_N = \tilde{f}_N \frac{h_{N-1/2}}{2}, \quad (8.7)$$

полученные при помощи метода баланса интегрированием (8.1) на отрезках $[0, x_{1/2}]$ и $[x_{N-1/2}, l]$, будут иметь порядок погрешности аппроксимации не менее двух. В соотношениях (8.6), (8.7) обозначено $h_{1/2} = x_1$, $h_{N-1/2} = l - x_{N-1}$,

$$\begin{aligned} \frac{1}{\tilde{p}_{1/2}} &= \frac{1}{h_{1/2}} \int_0^{x_1} \frac{dx}{p(x)}, & \tilde{q}_0 &= \frac{2}{h_{1/2}} \int_0^{x_{1/2}} q(x) dx, \\ \tilde{f}_0 &= \frac{2}{h_{1/2}} \int_0^{x_{1/2}} f(x) dx, & \frac{1}{\tilde{p}_{N-1/2}} &= \frac{1}{h_{1/2}} \int_{x_{N-1}}^l \frac{dx}{p(x)}, \\ \tilde{q}_N &= \frac{2}{h_{N-1/2}} \int_{x_{N-1/2}}^l q(x) dx, & \tilde{f}_N &= \frac{2}{h_{N-1/2}} \int_{x_{N-1/2}}^l f(x) dx. \end{aligned}$$

Из (8.2) имеем

$$\begin{aligned} v_0 &= -p(0) u'(x) \Big|_{x=0} = p(0) \frac{\tilde{\alpha} - \alpha u_0}{\alpha_1}, \\ v_N &= -p(l) u'(x) \Big|_{x=l} = -p(l) \frac{\tilde{\beta} - \beta u_N}{\beta_1} \end{aligned}$$

и после подстановки в (8.6), (8.7) получаем

$$b_0 u_0 - c_0 u_1 = y_0, \quad -a_N u_{N-1} + b_N u_N = y_N, \quad (8.8)$$

где

$$b_0 = c_0 + \tilde{q}_0 \frac{h_{1/2}}{2} + p(0) \frac{\alpha}{\alpha_1}, \quad c_0 = \frac{\tilde{p}_{1/2}}{h_{1/2}},$$

$$y_0 = \tilde{f}_0 \frac{h_{1/2}}{2} + p(0) \frac{\tilde{\alpha}}{\alpha_1}, \quad a_N = \frac{\tilde{p}_{N-1/2}}{h_{N-1/2}},$$

$$b_N = a_N + \tilde{q}_N \frac{h_{N-1/2}}{2} + p(l) \frac{\beta}{\beta_1}, \quad y_N = \tilde{f}_N \frac{h_{N-1/2}}{2} + p(l) \frac{\tilde{\beta}}{\beta_1}.$$

Из физического смысла краевой задачи (8.1), (8.2) вытекает, что если $\alpha_1 \neq 0$ и (или) $\beta_1 \neq 0$, то обычно $\alpha_1 = p(0)$ и (или) $\beta_1 = p(l)$.

Аппроксимацию (8.5) граничных условий (8.2) с первым порядком погрешности также можно привести к виду (8.8), если принять

$$c_0 = \frac{\tilde{p}_{1/2}}{h_{1/2}}, \quad b_0 = c_0 + \tilde{p}_{1/2} \frac{\alpha}{\alpha_1}, \quad y_0 = \tilde{p}_{1/2} \frac{\tilde{\alpha}}{\alpha_1},$$

$$a_N = \frac{\tilde{p}_{N-1/2}}{h_{N-1/2}}, \quad b_N = a_N + \tilde{p}_{N-1/2} \frac{\beta}{\beta_1}, \quad y_N = \tilde{p}_{N-1/2} \frac{\tilde{\beta}}{\beta_1}.$$

После умножения каждого разностного уравнения (8.3) на h_n запишем

$$-a_n u_{n-1} + b_n u_n - c_n u_{n+1} = y_n, \quad n = \overline{1, N-1}, \quad (8.9)$$

где

$$a_n = \frac{\tilde{p}_{n-1/2}}{h_{n-1/2}}, \quad b_n = a_n + c_n + \tilde{q}_n h_n, \quad c_n = \frac{\tilde{p}_{n+1/2}}{h_{n+1/2}}, \quad y_n = \tilde{f}_n h_n.$$

Итак, (8.8) и (8.9) образуют систему $N+1$ линейных алгебраических уравнений (СЛАУ) относительно неизвестных узловых значений u_n , $n = \overline{0, N}$, искомой функции $u(x)$. В матричную запись

$$Au = y \quad (8.10)$$

этой СЛАУ входят квадратная *трехдиагональная матрица*

$$A = \begin{pmatrix} b_0 & -c_0 & 0 & \dots & 0 & 0 \\ -a_1 & b_1 & -c_1 & \dots & 0 & 0 \\ 0 & -a_2 & b_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & b_{N-1} & -c_{N-1} \\ 0 & 0 & 0 & \dots & -a_N & b_N \end{pmatrix} \quad (8.11)$$

порядка $N+1$ и $(N+1)$ -мерные векторы $\mathbf{u} = (u_0 \ u_1 \ \dots \ u_N)^T$, $\mathbf{y} = (y_0 \ y_1 \ \dots \ y_N)^T$. Так как $a_n = c_{n-1}$, $n = \overline{1, N}$, то эта матрица симметрическая. Она содержит не более $2N+1$ не равных между собой ненулевых элементов и является *матрицей с частичным диагональным преобладанием*, поскольку $b_0 > c_0$ и $b_N > a_N$. СЛАУ с такой матрицей может быть решена *методом прогонки*.

Замечание 8.1. Если функция $p(x)$ непрерывно дифференцируема на отрезке $[0, l]$, то, казалось бы, целесообразно продифференцировать первое слагаемое в (8.1), используя затем (7.3) для аппроксимации $u''(x)$ и центральную конечную разность для аппроксимации $u'(x)$ на равномерной сетке с шагом h :

$$\begin{aligned} -(p(x) u'(x))' &= -p(x) u''(x) - p(x)' u'(x) \approx \\ &\approx -p_n \frac{u_{n+1} - 2u_n + u_{n-1}}{h^2} - p'_n \frac{u_{n+1} - u_{n-1}}{2h}, \end{aligned}$$

где p_n и p'_n — значения функции $p(x)$ и ее производной в узле x_n . Но при этом коэффициенты в (8.9) примут вид

$$a_n = \frac{p_n}{h} - \frac{p'_n}{2}, \quad b_n = 2\frac{p_n}{h} + q_n h, \quad c_n = \frac{p_n}{h} + \frac{p'_n}{2},$$

так что матрица (8.11) утратит симметричность, а для выполнения неравенств (7.25), связанных с существованием для

данной разностной схемы *принципа максимума*, шаг сетки необходимо будет выбирать из условия

$$h \max_{n=1, N-1} \frac{|p'_n|}{p_n} < 2. \quad (8.12)$$

Это условие может оказаться весьма обременительным при резком изменении функции $p(x)$ на отрезке $[0, l]$, когда $|p'(x)|$ достаточно велико при $x \in [0, l]$. Нарушение условия (8.12) может привести к утрате матрицей (8.11) свойства частичного диагонального преобладания, что не гарантирует эффективности метода прогонки и ставит под сомнение возможность использования данной разностной схемы на практике.

Замечание 8.2. Другим примером утраты свойства частичного диагонального преобладания матрицей (8.11) при аппроксимации дифференциального уравнения являются задачи, описываемые ОДУ (7.20) при $q(x) < 0$ или (8.1) при $p(x) > 0$ и $q(x) < 0$, $x \in [0, 1]$.

Ограничение на выбор шага сетки возникает и для задач, в которых наряду с диффузионным переносом физической субстанции необходимо учитывать и конвективный перенос движущейся средой (см. 1.2). В качестве примера рассмотрим установившийся процесс переноса вещества с концентрацией u в одномерной системе, описываемый ОДУ

$$-Du''(x) + v(x)u'(x) = f(x), \quad x \in [0, l], \quad (8.13)$$

где v — скорость среды, D — коэффициент диффузии вещества в среде, а функция $f(x)$ характеризует в данном случае интенсивность подвода вещества в систему. Это же ОДУ описывает перенос тепловой энергии, если u считать температурой среды, D — коэффициентом температуропроводности, а f — величиной, пропорциональной мощности распределенных источников теплоты.

Аппроксимация производных в (8.13) *центральными разностями* на равномерной одномерной сетке с узлами $x_n = nh$, $n = \overline{0, N}$, и шагом h приводит к разностным уравнениям

$$-\left(\frac{D}{h^2} + \frac{v_n}{2h}\right)u_{n-1} + \frac{2D}{h^2}u_n - \left(\frac{D}{h^2} - \frac{v_n}{2h}\right)u_{n+1} = f_n$$

со вторым *порядком аппроксимации*, которые соответствуют (8.9) при

$$a_n = \frac{D}{h^2} + \frac{v_n}{2h}, \quad b_n = \frac{2D}{h^2}, \quad c_n = \frac{D}{h^2} - \frac{v_n}{2h}, \quad n = \overline{1, N-1}. \quad (8.14)$$

Теперь условие частичного диагонального преобладания матрицы A в СЛАУ (8.10) выполнимо лишь при $\frac{|v_n|h}{2D} \leq 1$, причем хотя бы для одного значения n это неравенство должно быть строгим.

Это условие будет выполняться в любом случае, если в (8.13) аппроксимировать первую производную *левой* или *правой разностью*, ориентированной против движения среды, т.е. при $v_n > 0$ использовать соотношение

$$v(x)u'(x)\Big|_{x=x_n} = v_n \frac{u_n - u_{n-1}}{h} + O(h),$$

а при $v_n < 0$ —

$$v(x)u'(x)\Big|_{x=x_n} = v_n \frac{u_{n+1} - u_n}{h} + O(h).$$

Объединяя эти соотношения в одно

$$v(x)u'(x)\Big|_{x=x_n} = (|v_n| + v_n) \frac{u_n - u_{n-1}}{2h} + (|v_n| - v_n) \frac{u_n - u_{n+1}}{2h} + O(h),$$

получаем разностные уравнения вида (8.9) с первым порядком аппроксимации и коэффициентами

$$a_n = \frac{D}{h^2} + \frac{|v_n| + v_n}{2h}, \quad b_n = \frac{2D}{h^2} + \frac{|v_n|}{h}, \quad c_n = \frac{D}{h^2} + \frac{|v_n| - v_n}{2h},$$

$n = \overline{1, N-1}$. Нетрудно проверить, что теперь матрица (8.11) сохраняет свойство частичного диагонального преобладания при любых значениях v_n , но порядок аппроксимации разностной схемой уравнения (8.13) снижен на единицу. #

Вернемся к рассмотрению разностной аппроксимации ОДУ (8.1). Если входящие в него функции $p(x)$, $q(x)$ и $f(x)$ являются периодическими с периодом l , т.е.

$$p(x) = p(x+l), \quad q(x) = q(x+l), \quad f(x) = f(x+l), \quad x \in \mathbb{R},$$

то это ОДУ может иметь периодическое решение $u(x) = u(x+l)$, $x \in \mathbb{R}$, с тем же периодом. Для поиска такого решения вместо граничных условий (8.2) на концах отрезка $[0, l]$ следует задать условия „сшивания“

$$u(0) = u(l), \quad p(x)u'(x)\Big|_{x=+0} = p(x)u'(x)\Big|_{x=l-0}. \quad (8.15)$$

Если $p(+0) = p(l-0)$, то второе равенство в (8.15) переходит в равенство производных $u'(0) = u'(l)$. Это означает, что решение в точках $x = 0$ и $x = l$ будет гладким.

Такая задача возникает, например, при нахождении распределения температуры при установившейся теплопроводности в упомянутом выше тонком стержне длиной l , если он изогнут так, что его торцы находятся между собой в идеальном тепловом контакте. При этом температура $u(x)$ и тепловой поток $v(x) = -p(x)u'(x)$ изменяются вдоль стержня непрерывно.

Разностную схему для поиска периодического решения можно построить исходя из следующих соображений. Если на каждом из отрезков $[0, l]$, $[l, 2l]$, ... ввести неравномерную (в общем случае) сетку с $N+1$ узлами, причем узлы $x_N, x_{2N}, \dots$ будут общими для соседних отрезков, то коэффициенты и правые части в (8.9), а значит, и искомые узловые значения u_n станут повторяться с периодом N , т.е.

$$a_n = a_{n+N}, \quad b_n = b_{n+N}, \quad c_n = c_{n+N}, \quad y_n = y_{n+N}, \quad u_n = u_{n+N}.$$

Тогда вместо (8.8) в дополнение к (8.9) получим

$$-a_0 u_{N-1} + b_0 u_0 - c_0 u_1 = y_0 \quad \text{и} \quad u_N = u_0. \quad (8.16)$$

СЛАУ (8.8), (8.16) можно привести к матричной форме (8.10), но теперь N -мерные векторы $\mathbf{u}$ и $\mathbf{y}$ будут иметь координаты u_n и y_n , $n = \overline{0, N-1}$, а квадратная матрица A порядка N примет вид

$$A = \begin{pmatrix} b_0 & -c_0 & 0 & \dots & 0 & -a_0 \\ -a_1 & b_1 & -c_1 & \dots & 0 & 0 \\ 0 & -a_2 & b_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & b_{N-2} & -c_{N-2} \\ -c_{N-1} & 0 & 0 & \dots & -a_{N-1} & b_{N-1} \end{pmatrix}. \quad (8.17)$$

Таким образом, матрица A теперь не является трехдиагональной, что не позволяет применить для решения этой СЛАУ обычный вариант метода прогонки. Вместе с тем матрица (8.17), как и матрица (8.11), является симметрической, т.е. $a_0 = c_{N-1}$, а число не равных между собой ненулевых элементов будет не более $2N$.

К СЛАУ с матрицей, отличной от трехдиагональной, приводит и разностная схема краевой задачи для ОДУ (8.1) с граничными условиями, которые зависят не только от граничных значений искомой функции и ее производной. Такие **граничные условия** называют **связанными**. В частности, такие граничные условия могут содержать интегралы от искомой функции. Рассмотрим на примере, каким образом может войти такой интеграл в граничные условия.

Пример 8.1. Если для ОДУ (8.1) вместо (8.2) заданы оба условия на одном конце отрезка $[0, l]$ при $x = 0$ в виде $u(0) = u_0$ и $u'(x)|_{x=0} = u'_0$, то вместо краевой задачи получим задачу Коши для ОДУ второго порядка [VIII] с указанными начальными условиями при $x = 0$. Такие условия могут быть заданы, например, как результаты измерения температуры одного из торцов

стержня (см. рис. 7.1) и проходящего через этот торец теплового потока. Тогда целью решения задачи будет нахождение распределения температуры по длине стержня, соответствующее этим результатам измерения.

В общем случае неравномерной одномерной сетки с номерами узлов $n = \overline{0, N}$ из второго начального условия, используя правую конечную разность, находим $u_1 = u_0 + u'_0 h_{1/2}$ и далее, разрешая (8.3) относительно u_{n+1} , $n = \overline{1, N-1}$, приходим к рекуррентной формуле

$$u_{n+1} = \left(1 + \frac{\tilde{p}_{n-1/2} h_{n+1/2}}{\tilde{p}_{n+1/2} h_{n-1/2}} + \frac{q_n h_n h_{n+1/2}}{\tilde{p}_{n+1/2}} \right) u_n - \frac{\tilde{p}_{n-1/2} h_{n+1/2}}{\tilde{p}_{n+1/2} h_{n-1/2}} u_{n-1} - \frac{f_n h_n h_{n+1/2}}{\tilde{p}_{n+1/2}}. \quad (8.18)$$

Однако алгоритм, использующий (8.18), не обладает ни вычислительной устойчивостью, ни устойчивостью по входным данным, т.е. является неустойчивым, поскольку коэффициент при u_n заведомо превышает единицу.

Рассматриваемую задачу можно свести к задаче Коши для нормальной системы двух ОДУ [VIII]

$$u'(x) = \frac{v(x)}{p(x)}, \quad v'(x) = q(x)u(x) - f(x), \quad x \in [0, l],$$

с начальными условиями $u(0) = u_0$ и $v(0) = p(0)u'_0 = v_0$. Тогда, используя правые конечные разности, получаем разностную схему

$$\frac{u_{n+1} - u_n}{h_{n+1/2}} = \frac{v_n}{p_n}, \quad \frac{v_{n+1} - v_n}{h_{n+1/2}} = q_n u_n - f_n, \quad n = \overline{0, N-1}, \quad (8.19)$$

с заданными значениями u_0 и v_0 . Выражая v_n из второго равенства (8.19) и подставляя в первое равенство, приходим к рекуррентной формуле

$$u_{n+1} = \left(1 - \frac{q_n h_{n+1/2}^2}{p_n} \right) u_n + \frac{h_{n+1/2}}{p_n} v_{n+1} + \frac{f_n h_{n+1/2}^2}{p_n}.$$

Аналогично, выражая u_n из первого равенства (8.19) и подставляя во второе равенство, находим

$$v_{n+1} = \left(1 - \frac{q_n h_{n+1/2}^2}{p_n}\right) v_n + q_n h_{n+1/2} u_{n+1} - f_n h_{n+1/2}.$$

Из последних двух формул следует, что для обеспечения вычислительной устойчивости алгоритма, построенного на основе разностной схемы (8.19), значения $h_{n+1/2}$ должны удовлетворять условию

$$\max_{n=0, N-1} \frac{q_n h_{n+1/2}^2}{p_n} < 2.$$

Это условие при больших значениях отношения q_n/p_n будет весьма обременительным, поскольку его выполнение может привести к необходимости выбора чрезвычайно малого шага сетки. Но рассматриваемую задачу можно сформулировать как краевую, если для ОДУ (8.1) наряду с граничным условием $u(0) = u_0$ при $x = 0$ второе граничное условие задать при $x = l$ в виде

$$p(x) u'(x) \Big|_{x=l} - \int_0^l q(x) u(x) dx = p(0) u'_0 - \int_0^l f(x) dx. \quad (8.20)$$

Условие (8.20) получено интегрированием ОДУ (8.1) на отрезке $[0, l]$ с использованием заданного второго начального условия $u'(x) \Big|_{x=0} = u'_0$. При построении разностной схемы входящие во второе граничное условие интегралы можно приближенно представить соответствующими *квадратурными суммами*.

Предположим, что для интегрируемой на отрезке $[0, l]$ функции $F(x)$

$$\int_0^l F(x) dx = \sum_{n=0}^N A_n F(x_n) + R, \quad (8.21)$$

где A_n — весовые коэффициенты квадратурной формулы, а R — ее погрешность, причем узлы x_n квадратурной формулы совпадают с узлами сетки. Тогда, пренебрегая в (8.21) значением R и используя для аппроксимации производной $u'(x)$ в (8.20) левую конечную разность, можно записать

$$p_N \frac{u_N - u_{N-1}}{h_{N-1/2}} - \sum_{n=0}^N A_n q_n u_n = p_0 u'_0 - \sum_{n=0}^N A_n f_n = y_N,$$

где $p_N = p(l)$ и $p_0 = p(0)$, или

$$-\varphi_N u_0 - \sum_{n=1}^{N-1} g_n u_n + b_N u_N = y_N. \quad (8.22)$$

Здесь

$$\varphi_N = A_0 q_0, \quad b_N = \frac{p_N}{h_{N-1/2}} - A_N q_N, \quad g_{N-1} = \frac{p_N}{h_{N-1/2}} + A_{N-1} q_{N-1}$$

и $g_n = A_n q_n$, $n = \overline{1, N-2}$. Ясно, что (8.22) в сочетании с разностными уравнениями вида (8.3) для внутренних узлов образуют СЛАУ, для которой матрица не будет трехдиагональной. #

В некоторых случаях в ОДУ вида (8.1) могут входить слагаемые, зависящие от неизвестных граничных значений искомой функции:

$$-(p(x)u'(x))' + q(x)u(x) = f(x) + r^*(x)u(0) + s^*(x)u(l), \quad x \in [0, l],$$

где $r^*(x)$ и $s^*(x)$ — функции, интегрируемые на отрезке $[0, l]$. Тогда вместо (8.3) получим систему $N-1$ разностных уравнений

$$\begin{aligned} -\frac{1}{h_n} \left(\tilde{p}_{n+1/2} \frac{u_{n+1} - u_n}{h_{n+1/2}} + \tilde{p}_{n-1/2} \frac{u_{n-1} - u_n}{h_{n-1/2}} \right) + \tilde{q}_n u_n = \\ = \tilde{f}_n + \tilde{r}_n u_0 + \tilde{s}_n u_N, \end{aligned}$$

$n = \overline{1, N-1}$, где дополнительно обозначено

$$\tilde{r}_n = \frac{1}{h_n} \int_{x_{n-1/2}}^{x_{n+1/2}} r^*(x) dx, \quad \tilde{s}_n = \frac{1}{h_n} \int_{x_{n-1/2}}^{x_{n+1/2}} s^*(x) dx.$$

Сравнивая эту систему с (8.9), видим, что в левой части уравнений появились еще два слагаемых, и систему можно записать в виде

$$-\varphi_n u_0 - a_n u_{n-1} + b_n u_n - c_n u_{n+1} - \psi_n u_N = y_n, \quad n = \overline{1, N-1}, \quad (8.23)$$

где $\varphi_n = \tilde{r}_n h_n$, $\psi_n = \tilde{s}_n h_n$.

В случае связанных граничных условий аналогично (8.22) при $x = 0$ имеем (см. пример 8.1)

$$b_0 u_0 - \sum_{n=1}^{N-1} g_n^\circ u_n - \psi_0 u_N = y_0, \quad (8.24)$$

где g_n° , $n = \overline{1, N-1}$, ψ_0 и y_0 — известные коэффициенты. Тогда (8.22)–(8.24) образуют СЛАУ вида (8.10) с квадратной матрицей

$$\begin{pmatrix} b_0 & -g_1^\circ & -g_2^\circ & -g_3^\circ & \dots & -g_{N-1}^\circ & -\psi_0 \\ -a_1 - \varphi_1 & b_1 & -c_1 & 0 & \dots & 0 & -\psi_1 \\ -\varphi_2 & -a_2 & b_2 & -c_2 & \dots & 0 & -\psi_2 \\ -\varphi_3 & 0 & -a_3 & b_3 & \dots & 0 & -\psi_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -\varphi_{N-1} & 0 & 0 & 0 & \dots & b_{N-1} & -c_{N-1} - \psi_{N-1} \\ -\varphi_N & -g_1 & -g_2 & -g_3 & \dots & -g_{N-1} & b_N \end{pmatrix} \quad (8.25)$$

порядка $N+1$, которую можно получить из трехдиагональной матрицы (8.11) того же порядка заменой окаймляющих ее строк и столбцов строками и столбцами с ненулевыми (в общем случае) элементами.

Для одномерных стационарных задач, описываемых линейными ОДУ выше второго порядка, разностная схема представляет собой СЛАУ с матрицей, в большей мере заполненной ненулевыми элементами, чем трехдиагональная матрица, характерная для задач с ОДУ второго порядка. Это связано с необходимостью использовать для аппроксимации производных порядка выше второго значения функций более чем в трех соседних узлах.

Пример 8.2. Рассмотрим построение разностной схемы краевой задачи для линейного ОДУ четвертого порядка

$$(r(x)u''(x))'' + q(x)u(x) = f(x), \quad x \in [0, l], \quad (8.26)$$

где $r(x) \geq r_0 > 0$ и $q(x) \geq 0$ на отрезке $[0, l]$. Таким уравнением описывается, например, поперечный прогиб $u(x)$ балки под действием распределенной поперечной нагрузки $f(x)$. Балка имеет переменную жесткость $r(x)$ на изгиб и лежит на упругом основании, реакцию которого определяет слагаемое $q(x)u(x)$.

Для задачи с дифференциальным уравнением (8.26) должны быть заданы четыре граничных условия. Если речь идет об изгибе балки, граничные условия отражают то, как закреплены ее концы. Так, для консольной балки с жестко заштырленным левым и свободным правым концами (рис. 8.1) граничные условия имеют вид $u(0) = u'(0) = u''(l) = u'''(l) = 0$. Если балка имеет на концах опоры, допускающие (в отличие от жесткого заштырления) поворот ее поперечного сечения пропорционально действующему в этом сечении изгибающему моменту (рис. 8.2), то в этом случае граничные условия принимают следующий вид: $u(0) = u(l) = 0$, $u'(0) = \alpha r(0)u''(0)$, $u'(l) = \beta r(l)u''(l)$. Из

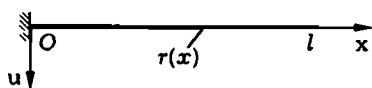


Рис. 8.1

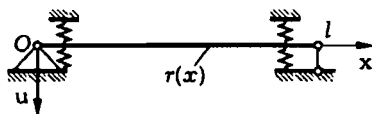


Рис. 8.2

этих уравнений при $\alpha = \beta = 0$ вытекают условия жесткого заземления, а при $\alpha \rightarrow \infty$ или $\beta \rightarrow \infty$ — шарнирного опирания: $u''(0) = 0$ или $u''(l) = 0$.

Предположим, что четыре граничных условия для ОДУ (8.26) заданы попарно на каждом из концов отрезка $[0, l]$ и содержат (в общем случае) линейные комбинации первых трех производных функции $u(x)$ в точках $x = 0$ и $x = l$, причем третья производная входит лишь в одно условие из каждой пары. Тогда на одномерной и, вообще говоря, неравномерной сетке с узлами x_n , $n = \overline{0, N}$, граничные условия можно аппроксимировать при помощи правых и левых разностей (см. 7.2). В результате получим уравнения

$$\begin{cases} c_0 u_0 - d_0 u_1 + e_0 u_2 = y_0, \\ -b_1 u_0 + c_1 u_1 - d_1 u_2 + e_1 u_3 = y_1, \\ a_{N-1} u_{N-3} - b_{N-1} u_{N-2} + c_{N-1} u_{N-1} - d_{N-1} u_N = y_{N-1}, \\ a_N u_{N-1} - b_N u_{N-1} + c_N u_N = y_N. \end{cases} \quad (8.27)$$

Аппроксимация четвертой производной в ОДУ (8.26) возможна при помощи центральных разностей лишь в узлах с номерами $n = \overline{2, N-2}$. Для этих узлов из (8.26) в соответствии с (7.16)–(7.19) следует

$$a_n u_{n-2} - b_n u_{n-1} + c_n u_n - d_n u_{n+1} + e_n u_{n+2} = y_n, \quad (8.28)$$

где с учетом того, что $h_{n-1/2} + h_{n+1/2} = 2h_n$,

$$a_n = \frac{\tilde{r}_{n-1}}{h_{n-3/2} h_{n-1} h_{n-1/2}}, \quad b_n = a_n + \frac{\tilde{r}_{n-1}}{h_{n-1} h_{n-1/2}^2} + \frac{2\tilde{r}_n}{h_{n-1/2}^2 h_{n+1/2}},$$

$$e_n = \frac{\tilde{r}_{n+1}}{h_{n+3/2} h_{n+1} h_{n+1/2}}, \quad d_n = e_n + \frac{\tilde{r}_{n+1}}{h_{n+1} h_{n+1/2}^2} + \frac{2\tilde{r}_n}{h_{n+1/2}^2 h_{n-1/2}},$$

$$c_n = b_n - a_n + d_n - e_n + \tilde{q}_n h_n, \quad y_n = \tilde{f}_n h_n, \quad n = \overline{2, N-1}.$$

Разностные уравнения (8.27) и (8.28) образуют СЛАУ вида (8.10) с **пятидиагональной матрицей**

$$\begin{pmatrix} c_0 & -d_0 & e_0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ -b_1 & c_1 & -d_1 & e_1 & 0 & \dots & 0 & 0 & 0 & 0 \\ a_2 & -b_2 & c_2 & -d_2 & e_2 & \dots & 0 & 0 & 0 & 0 \\ 0 & a_3 & -b_3 & c_3 & -d_3 & \dots & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & \dots & a_{N-1} & -b_{N-1} & c_{N-1} & -d_{N-1} \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 & a_N & -b_N & c_N \end{pmatrix}$$

порядка $N+1$. Ясно, что

$$e_n = a_{n+2}, \quad n = \overline{2, N-4}, \quad \text{и} \quad d_n = b_{n+1}, \quad n = \overline{2, N-3},$$

но в целом эта матрица будет симметрической лишь в некоторых частных случаях задания граничных условий. #

СЛАУ вида (8.10) с полученными выше матрицами можно решить методами линейной алгебры (в частности *методом Гаусса*), на основе которых разработаны алгоритмы и составлены программы, входящие в математическое обеспечение ЭВМ. Однако эти программы предусматривают обычно запись в памяти ЭВМ всех элементов матрицы СЛАУ, содержащей для системы из N уравнений N^2 элементов. Вместе с N значениями правых частей алгебраических уравнений исходная информация составляет $N(N+1)$ чисел. При решении СЛАУ методом Гаусса количество арифметических операций пропорционально N^3 .

Для решения СЛАУ с трехдиагональной матрицей наиболее рациональным является *метод прогонки*. Ниже (см. Д.8.1) рассмотрены различные модификации этого метода, позволяющие в некоторых случаях получить решение СЛАУ с матрицей, отличающейся от трехдиагональной, например с пятидиагональной матрицей.

8.2. Задача Штурма — Лиувилля

Вернемся к случаю $p(x) \geq p_0 > 0$ и $q_0 \leq q(x) \leq 0$ ($x \in [0, 1]$) в (8.1), характерному для систем, в которых возможны явления бифуркации, резонанса, теплового взрыва и т.п. Для таких систем может быть поставлена задача нахождения критических условий, при которых возникают подобные явления, т.е. вычисления критических нагрузок, резонансных частот колебаний, предельных значений теплового воздействия и т.д. Такие задачи включают однородное обыкновенное дифференциальное уравнение (ОДУ) вида (8.1) при $f(x) = 0$ с однородными граничными условиями. Входящий в ОДУ искомый критический параметр находят из условия существования нетривиального решения однородной задачи, которую называют *задачей Штурма — Лиувилля*.

Применим метод конечных разностей к приближенному решению такой задачи с ОДУ

$$-(p(x)u'(x))' - (\lambda - q(x))u(x) = 0, \quad x \in [0, l], \quad (8.29)$$

и граничными условиями $u(0) = 0$ и $u'(x)|_{x=l} = 0$ на концах отрезка $[0, l]$. При этих условиях и при выполнении неравенства $\pi^2 p_0 > -q_0$ оператор A , заданный при помощи (8.29), будет *положительно определенным* (см. пример 5.10) с *положительными собственными значениями*. При невыполнении этого неравенства свойство положительной определенности может быть утрачено, но оператор A останется *симметрическим*, а его собственные значения будут действительными числами [XV].

В случае равномерной одномерной *сетки с шагом* $h = l/N$ и узлами $x_n = nh$, $n = \overline{0, N}$, рассматриваемой задаче отвечают разностные уравнения (при $u_0 = 0$)

$$-a_n u_{n-1} + (b_n - \lambda)u_n - c_n u_{n+1} = 0, \quad n = \overline{1, N}, \quad (8.30)$$

где в соответствии с методом баланса $c_N = 0$, а также

$$\frac{1}{c_n} = \frac{1}{a_{n+1}} = h \int_{x_n}^{x_{n+1}} \frac{dx}{p(x)}, \quad b_n = a_n + c_n + \frac{1}{h} \int_{x_n - h/2}^{x_n + h/2} q(x) dx, \quad (8.31)$$

$n = \overline{1, N-1}$, и, кроме того,

$$\frac{1}{a_1} = h \int_0^h \frac{dx}{p(x)}, \quad b_N = a_N + \frac{1}{h} \int_{l-h/2}^l q(x) dx.$$

Отметим, что переход от (8.29) к (8.30) позволяет найти приближенные значения λ_m^* лишь N собственных значений λ_m , $m = \overline{1, N}$, оператора рассматриваемой задачи.

Условием существования нетривиального решения следующей из (8.30) однородной системы линейных алгебраических уравнений (СЛАУ) $Au - \lambda Iu = 0$, где $u = (u_1 \ u_2 \ \dots \ u_N)^T$ — вектор узловых значений, 0 — нулевой вектор размерности N , I — единичная матрица порядка N , A — *трехдиагональная матрица*

$$A = \begin{pmatrix} b_1 & -c_1 & 0 & \dots & 0 & 0 \\ -a_2 & b_2 & -c_2 & \dots & 0 & 0 \\ 0 & -a_3 & b_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & b_{N-1} & -c_{N-1} \\ 0 & 0 & 0 & \dots & -a_N & b_N \end{pmatrix} \quad (8.32)$$

порядка N , будет равенство нулю определителя $\det(A - \lambda I)$. Так как в (8.31) $a_{n+1} = c_n$, $n = \overline{1, N-1}$, то матрица A является симметрической. Поэтому все корни λ_m^* , $m = \overline{1, N}$, характеристического уравнения $\det(A - \lambda I) = 0$ этой матрицы действительные и являются ее *собственными значениями*.

Перед нахождением собственных значений и соответствующих им собственных векторов матрицы A целесообразно оценить промежуток их возможного изменения. Сначала рассмотрим произвольную симметрическую матрицу (a_{ij}) порядка N с элементами $a_{ij} = a_{ji}$, $i, j = \overline{1, N}$. Пусть λ — собственное значение этой матрицы, а x_j — координаты соответствующего ему ненулевого собственного вектора. Тогда для $i = \overline{1, N}$ будем иметь (по индексу i нет суммирования!)

$$\sum_{j=1}^N a_{ij} x_j = \lambda x_i, \quad \text{или} \quad (\lambda - a_{ii}) x_i = \sum_{\substack{j=\overline{1, N} \\ j \neq i}} a_{ij} x_j.$$

Отсюда с учетом неравенства треугольника находим

$$|\lambda - a_{ii}| |x_i| \leq \sum_{\substack{j=\overline{1, N} \\ j \neq i}} |a_{ij}| |x_j|, \quad i = \overline{1, N}. \quad (8.33)$$

Выбирая i так, чтобы $|x_i| = \max_{j=\overline{1, N}} |x_j|$, из (8.33) получаем

$$|\lambda - a_{ii}| \leq \sum_{\substack{j=\overline{1, N} \\ j \neq i}} |a_{ij}| \frac{|x_j|}{|x_i|} \leq \sum_{\substack{j=\overline{1, N} \\ j \neq i}} |a_{ij}| = \rho_i, \quad (8.34)$$

поскольку $\frac{|x_j|}{|x_i|} \leq 1$, $j = \overline{1, N}$. Таким образом, все собственные значения рассматриваемой матрицы лежат в промежутке, являющемся объединением отрезков $[a_{ii} - \rho_i, a_{ii} + \rho_i]$, $i = \overline{1, N}$, а для наибольшего собственного значения имеем оценку

$$\lambda_{\max} \leq \max_{i=\overline{1, N}} (a_{ii} + \rho_i). \quad (8.35)$$

Установленные свойства собственных значений составляют содержание **теоремы Гершгорина*** применительно к симметрической матрице.

*С.А. Гершгорин (1901–1933) — российский математик.

Для трехдиагональной матрицы (8.32) из (8.34) и (8.35) получаем

$$|\lambda - b_n| \leq |a_n| + |c_n|, \quad n = \overline{1, N}, \quad (8.36)$$

$$\lambda_{\max} \leq \max_{n=\overline{1, N}} (|a_n| + b_n + |c_n|). \quad (8.37)$$

Если в (8.85) выполнено третье неравенство и $b_n > 0$, $n = \overline{1, N}$, то из неравенств (8.36), (8.37) следует, что $\lambda \geq b_n - |a_n| - |c_n| \geq 0$, т.е. все собственные значения неотрицательны. Если к тому же A является *матрицей с частичным диагональным преобладанием*, то $\det A \neq 0$ (см. Д.8.1), т.е. она не имеет нулевого собственного значения и все ее собственные значения положительны, причем с учетом соотношений $|b_n| = b_n \geq |a_n| + |c_n|$, $n = \overline{1, N}$, из неравенств (8.36), (8.37) следует более грубая, но и более простая оценка

$$\lambda_{\max} \leq 2 \max_{n=\overline{1, N}} b_n \quad (8.38)$$

наибольшего собственного значения матрицы A по ее наибольшему диагональному элементу.

Замечание 8.3. Покажем, что любая трехдиагональная симметрическая матрица $A = (a_{nm})$ порядка N , у которой отличны от нуля все элементы $a_{n,n+1}$, стоящие над диагональю (такую *матрицу* относят к *неразложимым*), имеет простые собственные значения. Для этого рассмотрим матрицу $S = A - \xi I$, где ξ — некоторое число.

Сначала убедимся, что *минор* этой матрицы, соответствующий элементу s_{N1} , отличен от нуля. Действительно, определитель матрицы S_1 порядка $N - 1$, полученной вычеркиванием N -й строки и первого столбца в матрице S , равен взятому с обратным знаком произведению элемента $s_{12} = a_{12} \neq 0$ на определитель матрицы S_2 порядка $N - 2$, образованной вычеркиванием первой строки и второго столбца в матрице S_1 . В

свою очередь, этот определитель равен взятому с обратным знаком произведению элемента $s_{23} = a_{23} \neq 0$ на определитель матрицы S_3 порядка $N - 3$, полученной вычеркиванием в S_2 второй строки и третьего столбца, и так до тех пор, пока не дойдем до матрицы S_{N-1} , имеющей единственный элемент $s_{N-1,N} = a_{N-1,N} \neq 0$. Следовательно, матрица S имеет ненулевой минор порядка $N - 1$, равный произведению стоящих над ее диагональю элементов и числа $(-1)^{N-1}$. Для матрицы (8.32) этот минор равен произведению всех элементов c_n , $n = \overline{1, N-1}$.

Таким образом, согласно определению ранга матрицы [III], $\text{Rg } S \geq N - 1$ при любом ξ . Но если ξ — собственное значение матрицы A , т.е. $\det S = \det(A - \xi I) = 0$, то $\text{Rg } S < N$. Отсюда следует, что $\text{Rg } S = N - 1$, а размерность ядра оператора, которому соответствует матрица S , равна единице [IV]. Это означает, что каждому собственному значению матрицы A соответствует лишь один (с точностью до константы) собственный вектор. Но для любого самосопряженного оператора с симметрической матрицей существует ортонормированный базис [IV], состоящий из N линейно независимых собственных векторов этой матрицы, каждый из которых не может соответствовать двум различным ее собственным значениям. Так как матрица A имеет всего N собственных значений, то кратность каждого из них должна быть равна 1, т.е. все собственные значения простые. Для неразложимой трехдиагональной матрицы с частичным диагональным преобладанием при $a_{nn} > 0$, $n = \overline{1, N}$, все собственные значения положительны и попарно различны. #

Нахождение приближенных значений λ_m^* при $N \gg 1$ итерационными методами связано с многократным вычислением значения характеристического многочлена $\det(A - \lambda I)$ матрицы A при пробных значениях λ . Применение в данном случае формулы вида (8.38) затруднено в силу нарушения условия частичного диагонального преобладания применительно к матрице $A - \lambda I$, что может вызвать неустойчивость процесса вычислений. Су-

существует устойчивый и экономичный способ, требующий для вычисления $\det(A - \lambda I)$ всего $5N$ арифметических операций (сложений и умножений). Он основан на рекуррентной формуле

$$M_n = b_n M_{n-1} - a_n c_{n-1} M_{n-2}, \quad n = \overline{1, N}, \quad (8.39)$$

где M_n — *угловой минор* порядка n трехдиагональной матрицы $A - \lambda I$. Такие *миноры* иногда называют *главными* или *главными диагональными*. Для начала вычислений при $n = 1$ следует положить $M_0 = 1$ и $M_{-1} = 0$. Тогда, последовательно увеличивая n , при $n = N$ получаем $\det(A - \lambda I) = M_N$.

Отметим, что наряду с решением уравнения $\det(A - \lambda I) = 0$ для нахождения собственных значений матрицы A и соответствующих им собственных векторов применяют итерационные методы, не использующие вычисление значений характеристического многочлена*.

Иной путь нахождения собственных значений связан с применением алгоритма, характерного для численного решения задачи Коши для ОДУ первого порядка [VIII]. Однако использование для этой цели рекуррентной формулы

$$u_{n+1} = \frac{b_n - \lambda}{c_n} u_n - \frac{a_n}{c_n} u_{n-1}, \quad n = \overline{1, N-1}, \quad (8.40)$$

которая следует из (8.30), приведет в общем случае к неустойчивому алгоритму. Действительно, если задаться пробным значением λ и при $n = 1$ выбрать произвольное ненулевое значение u_1 (например, $u_1 = 1$), что возможно, так как любому собственному значению λ_m соответствует собственная функция $u^{(m)}(x)$, определенная в однородной задаче с точностью до постоянного множителя, то из (8.40) при $u_0 = 0$ можно найти значение u_2 и затем, увеличивая номер n , последовательно все значения u_{n+1} вплоть до u_N . Но так как в (8.40) коэффициенты b_n/c_n при u_n и $-a_n/c_n$ при u_{n-1} по абсолютной величине

*См.: Амосов А.А., Дубинский Ю.А., Копченкова Н.В.

могут превышать единицу, то из-за накопления вычислительной погрешности попытка уточнить значение λ , удовлетворяя последнее уравнение

$$-a_N u_{N-1} + (b_N - \lambda) u_N = 0 \quad (8.41)$$

однородной СЛАУ, может не дать положительного результата.

Преодолеть возникающие трудности можно следующим образом. Обозначим

$$p(x) u'(x) = -v(x) \quad (8.42)$$

и после подстановки в (8.29) получим

$$v'(x) - (\lambda - q(x)) u(x) = 0, \quad (8.43)$$

причем (8.42) и (8.43) образуют систему двух ОДУ первого порядка. На неравномерной одномерной сетке с узлами x_n , $n = \overline{0, N}$, ей соответствуют разностные уравнения

$$\tilde{p}_{n-1/2} \frac{u_n - u_{n-1}}{h_{n-1/2}} = -v_{n-1/2}, \quad n = \overline{1, N}; \quad (8.44)$$

$$v_{n+1/2} = v_{n-1/2} + \frac{h_{n-1/2} + h_{n+1/2}}{2} (\lambda - \tilde{q}_n) u_n, \quad n = \overline{1, N-1}, \quad (8.45)$$

где $\tilde{p}_{n-1/2}$ и $\tilde{q}_n$ вычислены, согласно методу баланса, в соответствии с (7.12) и (8.4).

Приняв $u_1 = 1$, из (8.44) при $n = 1$ и $u_0 = 0$ получим $v_{1/2} = -\frac{\tilde{p}_{1/2}}{h_{1/2}}$. Тогда, задаваясь пробным значением λ , последовательно при каждом значении $n = \overline{1, N-1}$ находим сначала из (8.45) $v_{n+1/2}$, а затем из (8.44)

$$u_{n+1} = u_n - \frac{h_{n+1/2}}{\tilde{p}_{n+1/2}} v_{n+1/2}, \quad n = \overline{1, N-1}. \quad (8.46)$$

Если при $n = N - 1$ вычисление по (8.46) даст $u_{n+1} = u_N$, не удовлетворяющее (8.41), то следует задать новое значение λ и

повторять описанную процедуру до тех пор, пока это уравнение не будет удовлетворено с заданной точностью. В итоге получим одно из собственных значений λ_m . Такой алгоритм решения задачи получил название **метода стрельбы**. В данном случае ведется „пристрелка“ граничного условия при $x = l$ путем подбора параметра λ_m .

Отметим, что вычисления этим методом можно вести в обратном направлении, задавшись пробным значением λ и приняв $u_N = 1$. Тогда из (8.41) получим $u_{N-1} = \frac{b_N - \lambda}{a_N}$, затем из (8.44) найдем $v_{N-1/2}$, а из (8.45) —

$$v_{N-3/2} = v_{N-1/2} - \frac{h_{N-3/2} + h_{N-1/2}}{2} (\lambda - \tilde{q}_{N-1}) u_{N-1}.$$

Это позволит из (8.44) вычислить u_{N-2} , из (8.45) найти $v_{N-5/2}$ и т.д. Если при последующем уменьшении n вычисленное из (8.44) значение u_0 будет отлично от нуля, т.е. не будет удовлетворять граничному условию при $x = 0$, то следует корректировать значение λ до тех пор, пока это граничное условие не будет удовлетворено с заданной точностью. Таким образом, в этом варианте метода стрельбы подбором параметра λ_m ведется „пристрелка“ граничного условия при $x = 0$.

Процесс вычислений по формулам (8.44), (8.46) устойчив и погрешности не накапливаются при переходе от n -го узла к $(n+1)$ -му, если коэффициенты, связывающие $v_{n+1/2}$ с $v_{n-1/2}$ и u_{n+1} с u_n , не превышают по модулю единицы. Подставив второе равенство (8.44) в (8.46), получим

$$u_{n+1} = \left(1 - \frac{h_{n-1/2} + h_{n+1/2}}{2\tilde{p}_{n+1/2}} (\lambda - \tilde{q}_n) h_{n+1/2} \right) u_n - \frac{h_{n+1/2}}{\tilde{p}_{n+1/2}} v_{n-1/2}.$$

Отсюда следует условие устойчивости

$$\left| 1 - \frac{h_{n-1/2} + h_{n+1/2}}{2\tilde{p}_{n+1/2}} (\lambda - \tilde{q}_n) h_{n+1/2} \right| \leq 1, \quad n = \overline{1, N-1},$$

или в случае равномерной сетки ($h_{n\pm 1/2} = h = \text{const}$)

$$h^2 \leq \min_{n=1, N-1} \frac{2\tilde{p}_{n+1/2}}{\lambda - \tilde{q}_n}. \quad (8.47)$$

Аналогичное условие устойчивости получим, если во второе равенство (8.44) подставим (8.46), заменив n на $n - 1$.

Условие (8.47) задает ограничение на выбор шага сетки, причем это ограничение ужесточается при поиске больших собственных значений λ_m с большими номерами m . Это связано с тем, что каждому собственному значению λ_m рассматриваемой задачи соответствует собственный вектор $u^{(m)} = (u_1^{(m)} \ u_2^{(m)} \ \dots \ u_N^{(m)})^T$ узловых значений $u_n^{(m)}$, $n = \overline{1, N}$, которые удовлетворяют (8.44). В непрерывной системе, описываемой ОДУ (8.29), этому вектору отвечает собственная функция $u^{(m)}(x)$, которая в интервале $(0, l)$ имеет $m - 1$ нулей. При больших m функция $u^{(m)}(x)$ становится быстро переменной и для ее удовлетворительной аппроксимации на одномерной сетке требуется большое число узлов. Поэтому описанный алгоритм на сетке с $N + 1$ узлами обычно позволяет достаточно точно вычислить лишь несколько первых собственных значений λ_m при $m \ll N$. Для нахождения больших значений λ_m системе (8.42), (8.43) можно поставить в соответствие другие разностные схемы или проинтегрировать эту систему численно одним из вариантов метода Рунге — Кутты (см. 8.3).

Пример 8.3. Рассмотрим задачу Штурма — Лиувилля для однородного ОДУ

$$-u''(x) - \lambda u(x) = 0, \quad x \in (0, 1), \quad (8.48)$$

с однородными граничными условиями $u(0) = u(1) = 0$. Несложно проверить, что собственными функциями этой задачи будут $\tilde{u}_m(x) = \sin m\pi x$, $m \in \mathbb{N}$, соответствующие собственным значениям $\lambda = (m\pi)^2$.

Разобьем отрезок $[0, 1]$ внутренними точками $x_n = nh$, $n = \overline{1, N-1}$, на N частичных отрезков одинаковой длины $h = 1/N$ и на одномерной равномерной сетке с номерами узлов $n = \overline{0, N}$ используем аппроксимацию (7.3) второй производной $u''(x)$, имеющую второй *порядок погрешности*. Тогда из (8.48) получим однородную систему $N - 1$ *разностных уравнений*

$$-\frac{u_{n-1} - 2u_n + u_{n+1}}{h^2} - \lambda u_n = 0, \quad n = \overline{1, N-1}.$$

Отсюда с учетом граничных условий $u_0 = u_N = 0$ приходим к однородной СЛАУ $(A - \lambda I)u = 0$ вида (8.30), в которой элементы симметрической трехдиагональной матрицы A порядка $N - 1$ имеют следующие значения: $a_n = c_{n-1} = \frac{1}{h^2} = N^2$, $n = \overline{2, N-1}$, и $b_n = \frac{2}{h^2} = 2N^2$, $n = \overline{1, N-1}$. Таким образом, для этой матрицы получаем характеристическое уравнение

$$\begin{vmatrix} 2N^2 - \lambda & -N^2 & 0 & \dots & 0 & 0 \\ -N^2 & 2N^2 - \lambda & -N^2 & \dots & 0 & 0 \\ 0 & -N^2 & 2N^2 - \lambda & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2N^2 - \lambda & -N^2 \\ 0 & 0 & 0 & \dots & -N^2 & 2N^2 - \lambda \end{vmatrix} = 0,$$

левая часть которого содержит определитель порядка $N - 1$.

При $N = 2$ это уравнение принимает вид $2N^2 - \lambda = 0$, откуда $\lambda = 2N^2 = 8$, что является достаточно грубым приближением к первому собственному значению $\lambda_1 = \pi^2 \approx 9,8696$ рассматриваемой задачи Штурма — Лиувилля. Для $N = 3$ имеем $(2N^2 - \lambda)^2 - N^4 = 0$. Отсюда получаем два собственных значения: $\lambda_1^* = N^2 = 9$ и $\lambda_2^* = 3N^2 = 27$. Значение λ_1^* приблизилось к λ_1 , но λ_2^* достаточно сильно отличается от второго собственного значения $\lambda_2 = 4\pi^2 \approx 39,4784$. При $N = 4$ уравнение $(2N^2 - \lambda)^3 - 2N^4(2N^2 - \lambda) = 0$ имеет три действительных корня:

$\lambda_1^* = (2 - \sqrt{2})N^2 \approx 9,3726$, $\lambda_2^* = 2N^2 = 32$ и $\lambda_3^* = (2 + \sqrt{2})N^2 \approx 54,6274$. Для сравнения приведем третье собственное значение $\lambda_3 = 9\pi^2 \approx 88,8264$. В случае $N = 5$ характеристическое уравнение

$$(2N^2 - \lambda)^4 - 3N^4(2N^2 - \lambda)^2 + N^8 = 0$$

является биквадратным относительно $2N^2 - \lambda$, что позволяет найти четыре корня

$$\lambda_{1,2,3,4}^* = \left(2 \pm \sqrt{\frac{3 \pm \sqrt{5}}{2}} \right) N^2,$$

или $\lambda_1^* \approx 9,5492$, $\lambda_2^* \approx 34,5492$, $\lambda_3^* \approx 65,4508$ и $\lambda_4^* \approx 90,4508$. Ясно, что для получения удовлетворительного приближения даже к первому собственному значению λ_1 необходимо выбрать достаточно большое значение N .

Отметим, что корни рассматриваемого характеристического уравнения можно представить в виде $\lambda_k^* = 4N^2 \sin \frac{\pi k}{2N}$, $k = \overline{1, N-1}$. Из этого представления вытекает, что $\lambda_1^* \rightarrow \lambda_1$ при $N \rightarrow \infty$.

Используем метод стрельбы для получения приближенного значения λ_1 . В данном случае, выбрав $u_1 = 1$, из (8.44) при $n = 1$ и $u_0 = 0$ найдем $v_{1/2} = -1/h = -N$ и, задавшись пробным значением λ , последовательно при каждом значении $n = \overline{1, N-1}$ вычислим из (8.45) $v_{n+1/2} = v_{n-1/2} + \lambda h u_n = v_{n-1/2} + \frac{\lambda u_n}{N}$, а затем из (8.46) $u_{n+1} = u_n - h v_{n+1/2} = u_n - \frac{v_{n+1/2}}{N}$. При $N = 10$, задаваясь пробным значением $\lambda = 10$, получаем $u_N \approx -0,1089$, а для пробного значения $\lambda = 9,7$ находим $u_N \approx 0,0470$. Это означает, что условию $u_N = 0$ будет удовлетворять промежуточное значение λ , т.е. $9,7 < \lambda_1^* < 10$. Последовательными приближениями устанавливаем, что $|u_N| \leq 10^{-6}$ при $\lambda = \lambda_1^* \approx 9,7887$. В случае $N = 100$ такая же точность достигнута при $\lambda_1^* \approx 9,8688$, что по сравнению с λ_1 меньше лишь на 0,0008.

Использование метода стрельбы для вычисления второго собственного значения при условии, что $|u_N| \leq 10^{-6}$, дает $\lambda_2^* \approx \approx 38,1966$ при $N = 10$ и $\lambda_2^* \approx 39,4654$ при $N = 100$. Последний результат меньше значения λ_2 на 0,0130. Отметим, что условие (8.47) устойчивости выполнено для всех вариантов расчетов, проведенных в этом примере.

8.3. Нестационарная задача теплопроводности

В качестве характерного примера рассмотрим одномерную нестационарную краевую задачу теплопроводности в твердом теле, описываемую линейным дифференциальным уравнением с частными производными (см. 2.3)

$$c(x) \frac{\partial T(t, x)}{\partial t} = \frac{\partial}{\partial x} \left(\lambda(x) \frac{\partial T(t, x)}{\partial x} \right) + I_V^{(q)}(t, x), \quad t > 0, \quad (8.49)$$

где $T(t, x)$ — зависящая от времени t и пространственной координаты $x \in (0, l)$ температура тела; $c(x) > 0$ и $\lambda(x) > 0$ — объемная теплоемкость и теплопроводность; $I_V^{(q)}(t, x)$ — объемная мощность внутренних источников тепловой энергии. Помимо (8.49) в формулировку нестационарной краевой задачи теплопроводности должны входить начальные условия в виде заданного распределения температуры $T(0, x) = T^0(x)$ в момент времени $t = 0$, принимаемый за начальный, и граничные условия, заданные на концах отрезка $[0, l]$. В данном случае примем граничные условия вида (8.2):

$$-\lambda(x) \frac{\partial T(t, x)}{\partial x} \Big|_{x=0} + \alpha T(t, 0) = \tilde{\alpha}(t), \quad (8.50)$$

$$\lambda(x) \frac{\partial T(t, x)}{\partial x} \Big|_{x=l} + \beta T(t, l) = \tilde{\beta}(t). \quad (8.51)$$

Можно указать три основных подхода к решению такой краевой задачи с использованием метода конечных разностей (МКР).

Первый из них, называемый **методом прямых**, связан с аппроксимацией производной в левой части (8.49):

$$\frac{\partial T(t, x)}{\partial t} = \frac{T_k(x) - T_{k-1}(x)}{\Delta t_k} + O(\Delta t_k), \quad (8.52)$$

где $T_{k-1}(x)$ и $T_k(x)$ — распределения температуры в теле в моменты времени t_{k-1} и $t_k = t_{k-1} + \Delta t_k$ в начале и конце k -го интервала времени Δt_k . При рассмотрении процесса теплопроводности в теле на k -м интервале времени предположим, что распределение $T_{k-1}(x)$ известно из решения задачи на предшествующем, $(k-1)$ -м интервале, а при $k=1$ оно определено начальными условиями, т.е. $T_0(x) = T^\circ(x)$. Тогда, подставив (8.52) в (8.49) при условии, что правая часть (8.49) соответствует моменту времени t_k , получим линейное обыкновенное дифференциальное уравнение (ОДУ) второго порядка

$$-\frac{d}{dx} \left(\lambda(x) \frac{dT_k(x)}{dx} \right) + \frac{c(x)}{\Delta t_k} T_k(x) = I_V^{(q)}(t_k, x) + \frac{c(x)}{\Delta t_k} T_{k-1}(x),$$

$$x \in [0, l],$$

описывающее искомое распределение температуры $T_k(x)$. Это ОДУ может быть решено с учетом граничных условий в фиксированный момент времени t_k способами, рассмотренными выше (см. 8.1).

Второй подход основан на использовании аппроксимации правой части (8.49) на *одномерной сетке* с N узлами $x_n \in [0, l]$, $n = \overline{0, N}$. Это соответствует переходу от непрерывного во времени и пространстве процесса теплопроводности к его дискретной в пространственном отношении *математической модели*. В итоге с учетом заданных граничных условий получаем нормальную систему ОДУ относительно искомых *узловых значений* $T_n(t)$ температуры, зависящих от времени. В задачу Коши для этой системы в качестве начальных условий входят узловые значения $T_n(0) = T^\circ(x_n) = T_n^\circ$ в момент времени $t = 0$.

Наконец, третий подход объединяет первые два и связан с переходом к дискретной математической модели нестационарного процесса теплопроводности как в пространстве, так и во времени. Эта модель в каждый фиксированный момент времени на одномерной сетке соответствует системе линейных алгебраических уравнений (СЛАУ). Последовательно рассмотрим пути решения нестационарной краевой задачи для уравнения (8.49) на основе второго и третьего подходов.

При аппроксимации правой части (8.49) на одномерной неравномерной сетке с N узлами $x_n \in [0, l]$, $n = \overline{0, N}$, нормальная система ОДУ принимает вид

$$s_n \frac{dT_n(t)}{dt} = a_n T_{n-1}(t) - b_n T_n(t) + c_n T_{n+1}(t) + f_n(t). \quad (8.53)$$

Здесь в соответствии с *методом баланса*

$$\frac{1}{c_n} = \frac{1}{a_{n+1}} = \int_{x_n}^{x_{n+1}} \frac{dx}{\lambda(x)}, \quad s_n = \int_{x_{n-1/2}}^{x_{n+1/2}} c(x) dx, \quad (8.54)$$

$n = \overline{0, N}$, причем $a_0 = c_N = a_{N+1} = 0$, $x_{n \pm 1/2} = \frac{1}{2}(x_n + x_{n \pm 1})$, но $x_{-1/2} = 0$ и $x_{N+1/2} = x_N = l$. Кроме того,

$$b_n = a_n + c_n, \quad f_n(t) = \int_{x_{n-1/2}}^{x_{n+1/2}} I_V^{(q)}(t, x) dx, \quad (8.55)$$

$n = \overline{1, N-1}$. По аналогии с (8.8) при аппроксимации (8.50), (8.51) получим $b_0 = c_0 + \alpha$, $b_N = a_N + \beta$ и

$$f_0(t) = \tilde{\alpha}(t) + \int_0^{x_{1/2}} I_V^{(q)}(t, x) dx, \quad f_N(t) = \tilde{\beta}(t) + \int_{x_{N-1/2}}^{x_N} I_V^{(q)}(t, x) dx.$$

В матричной записи (8.53) имеет вид

$$S \frac{dT(t)}{dt} + AT(t) = f(t) \quad (8.56)$$

с начальным условием $T(0) = T^0$, где векторы $T(t)$, T^0 и $f(t)$ размерности $N + 1$ имеют координаты $T_n(t)$, T_n^0 и $f_n(t)$, $n = \overline{0, N}$, соответственно; диагональная матрица S содержит на своей диагонали элементы s_n ; квадратная матрица A порядка $N + 1$ является *трехдиагональной* вида (8.11). Используя (8.39), нетрудно установить, что $\det A = 0$ при $\alpha = \beta = 0$, т.е. в этом случае матрица A будет вырожденной. В дальнейшем предполагаем, что значения α и β неотрицательны и хотя бы одно из них положительно. При этом A будет *матрицей с частичным диагональным преобладанием*.

Из (8.54) следует, что $a_{n+1} = c_n \neq 0$, $n = \overline{0, N-1}$, т.е. матрица A *неразложимая* и симметрическая. В силу замечания 8.3 она имеет $N + 1$ простых (т.е. попарно различных) положительных *собственных значений* λ_m , $m = \overline{1, N+1}$, а ее собственные векторы образуют ортонормированный базис [IV]. В этом базисе матрица A имеет диагональный вид Λ , в котором диагональными элементами являются собственные значения. При этом $\Lambda = C^T A C$, где C — *ортогональная матрица* порядка $N + 1$, столбцами которой являются *собственные векторы матрицы* A . Учитывая, что $CC^T = C^T C = I_{N+1}$, где I_{N+1} — единичная матрица порядка $N + 1$, получаем $A = C \Lambda C^T$.

В линейном пространстве $M_{N+1}(\mathbb{R})$ матриц порядка $N + 1$ введем *спектральную норму*. В этом случае $M_{N+1}(\mathbb{R})$ будет конечномерным $((N + 1)$ -мерным) нормированным пространством. Определим в этом пространстве функцию $e^B = \exp(B)$, $B \in M_{N+1}(\mathbb{R})$:

$$\exp(B) = I_{N+1} + B + \frac{1}{2!} B^2 + \frac{1}{3!} B^3 + \dots,$$

где ряд в правой части равенства сходится абсолютно для любой матрицы $B \in M_{N+1}(\mathbb{R})$ по спектральной норме к некото-

рой матрице $\exp(B) \in M_{N+1}(\mathbb{R})$ [IX]. Нетрудно доказать, что для любых $t, s \in \mathbb{R}$ и $A \in M_{N+1}(\mathbb{R})$ справедливо $\exp((t+s)A) = \exp(tA)\exp(sA)$ и

$$\frac{d}{dt} \exp(tA) = \lim_{\Delta t \rightarrow 0} \frac{\exp((t+\Delta t)A) - \exp(tA)}{\Delta t} = A \exp(tA).$$

Из последнего равенства следует, что общее решение системы линейных однородных ОДУ $\frac{T(t)}{dt} = -A_* T(t)$, определяемой матрицей A_* порядка $N+1$, имеет вид $T(t) = \exp(-A_* t) \tilde{T}$, где $\tilde{T}$ — произвольный вектор из $\mathbb{R}^{N+1}$. Применяя *метод вариации постоянных*, получаем решение системы (8.56) в матричной форме

$$T(t) = \exp(-A_* t) T^0 + S^{-1} \int_0^t \exp(-A_*(t-\tau)) f(\tau) d\tau, \quad (8.57)$$

где $A_* = S^{-1}A$, причем экспоненциальную функцию от квадратной матрицы $A_* t$ можно представить сходящимся рядом

$$\exp(-A_* t) = I + \sum_{j=1}^{\infty} \frac{(-1)^j t^j}{j!} A_*^j. \quad (8.58)$$

В частном случае $s_n = s = \text{const} > 0$, $n = \overline{0, N}$, имеем $A_* = A/s$, т.е. все собственные значения $\lambda_m^* = \lambda_m/s$, $m = \overline{1, N+1}$, матрицы A_* положительные и простые. Поэтому

$$\exp(-A_* t) = CDC^T, \quad (8.59)$$

где D — диагональная матрица порядка $N+1$ с элементами $\exp\left(-\frac{\lambda_m t}{s}\right)$, $m = \overline{1, N+1}$.

При $N \gg 1$ практическое использование решения непосредственно в виде (8.57) затруднено. Рассмотрим приближенные

способы нахождения вектора $T(t)$. Для этого введем совокупность интервалов времени Δt_k , $k \in \mathbb{N}$. Тогда в пределах k -го интервала $\Delta t_k = t_k - t_{k-1}$ можно представить (8.57) в виде

$$T^{(k)}(t) = \exp(-A_*(t - t_{k-1}))T_{k-1} + \int_{t_{k-1}}^t \exp(-A_*(t - \tau))f(\tau) d\tau,$$

где $T_{k-1} = T(t_{k-1})$ — вектор узловых значений температуры в момент времени t_{k-1} . В частности, при $t = t_k$ для конца k -го интервала отсюда получим

$$T_k = T(t_k) = \exp(-A_*\Delta t_k)T_{k-1} + \int_0^{\Delta t_k} \exp(-A_*(\Delta t_k - \tau))f(t_{k-1} + \tau) d\tau. \quad (8.60)$$

Для малых значений Δt_k в (8.58) под знаком суммы можно ограничиться лишь первым слагаемым и вместо (8.60) записать

$$T_k = T_{k-1} - A_*T_{k-1}\Delta t_k + \int_0^{\Delta t_k} f(t_{k-1} + \tau) d\tau - A_* \int_0^{\Delta t_k} f(t_{k-1} + \tau)(\Delta t_k - \tau) d\tau. \quad (8.61)$$

При известной зависимости координат вектора $f(t)$ от t интегралы в (8.61) нетрудно вычислить при помощи *квадратурных формул*.

Требование малости интервала Δt_k количественно можно охарактеризовать при помощи какой-либо нормы матрицы A_* , например ее спектральной нормы $\|A_*\|$. Тогда справедливость (8.61) связана с предположением $\|A_*\|\Delta t_k \ll 1$. Но алгоритм вычислений по (8.61) будет устойчивым по отношению к накоплению погрешностей при выполнении более слабого ограничения $\|I - A_*\Delta t_k\| \leq 1$, приводящего к неравенству $\|A_*\|\Delta t_k \leq 2$.

Спектральная норма квадратной матрицы является кольцевой [IV], т.е. для спектральной нормы произведения квадратных матриц B и C справедливо неравенство $\|BC\| \leq \|B\|\|C\|$. Так как матрица S^{-1} диагональная с элементами $\frac{1}{s_n} > 0$ на диагонали, то все ее собственные значения положительны (как и для матрицы A). Поэтому спектральные нормы матриц S^{-1} и A будут равны их наибольшим собственным значениям [IV]. Тогда с учетом (8.38) можно написать

$$\|A_*\| = \|S^{-1}A\| \leq \left(\max_{n=0, N} \frac{1}{s_n} \right) \max_{n=0, N} \lambda_n \leq 2 \frac{b_{\max}}{s_{\min}},$$

где

$$b_{\max} = \max_{n=0, N} b_n, \quad s_{\min} = \min_{n=0, N} s_n.$$

Таким образом, из условия $\|A_*\|\Delta t_k \leq 2$ *вычислительной устойчивости* рассматриваемого алгоритма следует ограничение на выбор интервала времени

$$\Delta t_k \leq \frac{2}{\|A_*\|} \leq \frac{s_{\min}}{b_{\max}}. \quad (8.62)$$

В том случае, когда собственные значения каждой из матриц S^{-1} и A сильно отличаются друг от друга, т.е. их *числа обусловленности* велики, ограничение (8.62) заставляет выбирать Δt_k слишком малым и использование (8.61) становится нерациональным. **Системы ОДУ** вида (8.56) с такими матрицами принято называть **жесткими**. В рассматриваемой задаче теплопроводности такая ситуация может возникнуть вследствие сильной неоднородности материала тела, т.е. резкого изменения функций $c(x)$ и $\lambda(x)$, и при значительной неравномерности сетки. Аналитическое решение для жестких систем содержит обычно экспоненциальные функции с сильно отличающимися друг от друга показателями, что и вызывает указанные вычислительные трудности и требует применения специальных методов*.

*См.: Амосов А. А., Дубинский Ю. А., Копченова Н. В.

Для численного решения системы (8.56) можно использовать один из вариантов *метода Рунге — Кутты*. Эти варианты можно построить на основе представления решения этой системы интегралом в пределах интервала времени Δt_k :

$$T_k = T_{k-1} + S^{-1} \int_0^{\Delta t_k} (f(t_{k-1} + \tau) - AT(t_{k-1} + \tau)) d\tau. \quad (8.63)$$

Если интеграл в (8.63) заменить с погрешностью, пропорциональной $(\Delta t_k)^2$, выражением $(f_{k-1} - AT_{k-1})\Delta t_k$, где $f_{k-1} = f(t_{k-1})$, то получим рекуррентную формулу *метода Эйлера*

$$T_k = (I - A_* \Delta t_k) T_{k-1} + S^{-1} f_{k-1} \Delta t_k. \quad (8.64)$$

Метод Эйлера является простейшим вариантом метода Рунге — Кутты. В этом случае условием устойчивости алгоритма остается неравенство (8.62).

Если для представления интеграла в (8.63) использовать квадратурную *формулу трапеции*, то придем к выражению

$$T_k = T_{k-1} - A_*(T_{k-1} + T_k) \frac{\Delta t_k}{2} + S^{-1} (f_{k-1} + f_k) \frac{\Delta t_k}{2} \quad (8.65)$$

с погрешностью, пропорциональной $(\Delta t_k)^3$. Отсюда следует, что

$$T_k = \left(I + \frac{1}{2} A_* \Delta t_k\right)^{-1} \left(\left(I - \frac{1}{2} A_* \Delta t_k\right) T_{k-1} + \frac{1}{2} S^{-1} (f_{k-1} + f_k) \Delta t_k \right),$$

где $f_k = f(t_k)$. Условие устойчивости алгоритма вычислений теперь принимает вид

$$\left\| \left(I + \frac{1}{2} A_* \Delta t_k\right)^{-1} \left(I - \frac{1}{2} A_* \Delta t_k\right) \right\| \leq 1. \quad (8.66)$$

Можно показать, что для матрицы A с частичным диагональным преобладанием и диагональными элементами $b_n > 0$ это условие выполняется для любых значений Δt_k .

Чтобы избежать обращения матрицы $I + A\Delta t_k/2$ и сохранить третий порядок погрешности аппроксимации, алгоритм вычислений на k -м интервале времени можно построить следующим образом: сначала по (8.64) вычислить вектор-прогноз $\tilde{T}_k$, который затем использовать в правой части (8.65) вместо вектора T_k . Условием устойчивости такого алгоритма с прогнозом и последующей коррекцией будет неравенство

$$\left\| I - A_*\Delta t_k + \frac{1}{2}(A_*\Delta t_k)^2 \right\| \leq 1, \quad (8.67)$$

которое обычно в меньшей мере ограничивает выбор Δt_k , чем условие (8.62).

Другой путь построения алгоритма, обеспечивающего третий порядок погрешности аппроксимации на k -м интервале времени, связан с представлением интеграла в (8.63) выражением $(f_{k-1/2} - AT_{k-1/2})\Delta t_{k-1}$, где индекс $k - 1/2$ указывает на то, что значения зависящих от времени величин соответствуют моменту $t_{k-1/2} = t_{k-1} + \frac{1}{2}\Delta t_k$ в середине этого интервала. Тогда по методу Эйлера найдем вектор

$$T_{k-1/2} = \left(I - \frac{1}{2}A_*\Delta t_k \right) T_{k-1} + \frac{1}{2}S^{-1}f_{k-1}\Delta t_k, \quad (8.68)$$

а затем — искомый вектор

$$T_k = T_{k-1} + S^{-1}(f_{k-1/2} - AT_{k-1/2})\Delta t_{k-1}. \quad (8.69)$$

Устойчивость двухступенчатого алгоритма (8.68), (8.69) гарантирована при выполнении неравенства (8.67). Отметим, что в соответствии с (8.58) левые части в (8.62) и (8.67) являются нормами усеченных представлений выражения $\exp(-A_*\Delta t_k)$ с остаточными членами $O((\Delta t_k)^2)$ и $O((\Delta t_k)^3)$ соответственно.

Чтобы найти вектор T_k с погрешностью порядка $r = 4$ необходимо в пределах k -го интервала времени использовать

трехступенчатый алгоритм, например такой:

$$\begin{aligned}\Delta T_k^{(1)} &= S^{-1}(f_{k-1} - AT_{k-1})\Delta t_k, \\ \Delta T_k^{(2)} &= S^{-1}\left(f_{k-1/2} - A\left(T_{k-1} + \frac{1}{2}\Delta T_k^{(1)}\right)\right)\Delta t_k, \\ \Delta T_k^{(3)} &= S^{-1}\left(f_k - A(T_{k-1} - \Delta T_k^{(1)} + 2\Delta T_k^{(2)})\right)\Delta t_k, \\ T_k &= T_{k-1} + \frac{1}{6}(\Delta T_k^{(1)} + 4\Delta T_k^{(2)} + \Delta T_k^{(3)}).\end{aligned}$$

Нетрудно проверить, что условием устойчивости этого алгоритма будет неравенство

$$\left\| I - A_*\Delta t_k + \frac{(A_*\Delta t_k)^2}{2} - \frac{(A_*\Delta t_k)^3}{6} \right\| \leq 1.$$

Наконец, приведем пример четырехступенчатого алгоритма с порядком погрешности $r = 5$, в котором для $\Delta T_k^{(1)}$ и $\Delta T_k^{(2)}$ сохраняются прежние выражения, а далее вычисления проводят по формулам

$$\begin{aligned}\Delta T_k^{(3)} &= S^{-1}\left(f_{k-1/2} - A\left(T_{k-1} + \frac{1}{2}\Delta T_k^{(2)}\right)\right)\Delta t_k, \\ \Delta T_k^{(4)} &= S^{-1}\left(f_k - A(T_{k-1} + \Delta T_k^{(3)})\right)\Delta t_k, \\ T_k &= T_{k-1} + \frac{1}{6}(\Delta T_k^{(1)} + 2\Delta T_k^{(2)} + 2\Delta T_k^{(3)} + \Delta T_k^{(4)}).\end{aligned}$$

Все перечисленные алгоритмы принадлежат семейству метода Рунге — Кутты. В этом семействе можно построить алгоритмы с еще более высоким порядком погрешности r . Способы их построения удобно пояснить на примере решения системы (8.56) с нулевым вектором $f(t)$. В этом случае на k -м интервале времени из (8.60) получим

$$T_k = T_{k-1} \exp(-A_*\Delta t_k). \quad (8.70)$$

Учитывая свойство экспоненциальной функции

$$\exp(-A_* \Delta t_k) = \exp\left(-\frac{1}{2} A_* \Delta t_k\right) \exp\left(-\frac{1}{2} A_* \Delta t_k\right),$$

можно в соответствии с (8.58) сначала вычислить

$$\exp\left(-\frac{1}{2^r} A_* \Delta t_k\right) \approx 1 + \sum_{j=1}^{r-1} \frac{(-1)^j (\Delta t_k/2)^j}{j!} A_*^j$$

с порядком погрешности r , а затем последовательно за r шагов с тем же порядком погрешности найти

$$\exp\left(-\frac{1}{2^{r-1}} A_* \Delta t_k\right), \quad \exp\left(-\frac{1}{2^{r-2}} A_* \Delta t_k\right), \quad \dots, \quad \exp\left(-\frac{1}{2} A_* \Delta t_k\right)$$

и, наконец, $\exp(-A_* \Delta t_k)$, используя этот результат в (8.70).

Наиболее употребительными являются алгоритмы метода Рунге — Кутты при $r \leq 5$. Их можно использовать и для решения нелинейных задач (например, в случае, когда в (8.49) параметры c , λ и $I_V^{(g)}$ произвольным образом зависят от искомой функции — температуры).

Реализация третьего подхода к решению нестационарной краевой задачи при помощи МКР связана с предположением о постоянстве скорости изменения искомой функции в пределах интервала времени в каждом узле одномерной сетки по пространственной координате и равносильна использованию в (8.56) аппроксимации производной по t конечно-разностным соотношением вида (8.52). Тогда из (8.56) получим разностную схему

$$S \frac{T_k - T_{k-1}}{\Delta t_k} = (1 - \eta)(f_{k-1} - AT_{k-1}) + \eta(f_k - AT_k) \quad (8.71)$$

с весовыми коэффициентами $1 - \eta$ и $\eta \in [0, 1]$, называемую **двухслойной разностной схемой с весами**, поскольку в нее входят узловые значения температуры на двух *слоях пространственно-временной сетки*.

Значению $\eta = 0$ соответствует **явная двухслойная разностная схема**, совпадающая с (8.64). В этом случае (8.71) можно явно разрешить относительно искомого вектора T_k , а каждое узловое значение T_n^k , $n = \overline{0, N}$, температуры в конце k -го интервала времени Δt_k вычислить независимо от остальных из выражения

$$T_n^k = T_n^{k-1} \left(1 - \frac{b_n}{s_n} \Delta t_k \right) + \frac{a_n T_{n-1}^{k-1} + c_n T_{n+1}^{k-1} + f_n^{k-1}}{s_n} \Delta t_k, \quad (8.72)$$

которое получается из (8.53) при аппроксимации производной $\frac{dT_n(t)}{dt}$ *правой конечной разностью* (разностью „вперед“).

Если в (8.49) $c(x) \equiv c = \text{const}$, $\lambda(x) \equiv \lambda = \text{const}$, $I_V^{(q)}(t, x) \equiv 0$, то в случае одномерной равномерной сетки с шагом h разностная схема, соответствующая (8.72), принимает вид

$$\frac{T_n^k - T_n^{k-1}}{\Delta t_k} = \tilde{a} \frac{T_{n-1}^{k-1} - 2T_n^{k-1} + T_{n+1}^{k-1}}{h^2}, \quad \tilde{a} = \frac{\lambda}{c}.$$

Сюда входят узловые значения температуры в четырех узлах пространственно-временной сетки, составляющих **шаблон** этой **разностной схемы**. При графическом изображении шаблона его узлы принято соединять отрезками прямых (рис. 8.3). В данном случае шаблон и разностную схему называют **четыреугольными**. Разностная схема (8.71) и ее шаблон являются **шестигранными**. Для одномерной равномерной сетки он изображен на рис. 8.4.

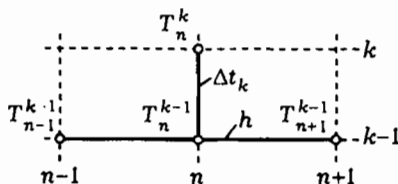


Рис. 8.3

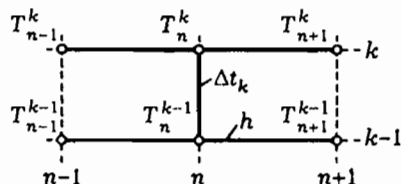


Рис. 8.4

Процесс вычислений по (8.72) будет устойчивым, т.е. погрешность вычисления температур T_n^k , $n = \overline{0, N}$, по абсолютной величине не будет превышать наибольшей из погрешностей, возникших при вычислениях на предыдущем интервале времени, если сумма модулей коэффициентов при значениях температуры в правой части (8.72) не превышает единицы:

$$\left| 1 - \frac{b_n}{s_n} \Delta t_k \right| + \frac{|a_n| + |c_n|}{s_n} \Delta t_k \leq 1, \quad n = \overline{0, N}.$$

Отсюда в силу частичного диагонального преобладания матрицы A в (8.71) и положительности ее диагональных элементов имеем

$$\Delta t_k \max_{n=\overline{0, N}} \frac{|a_n| + b_n + |c_n|}{s_n} \leq \max_{n=\overline{0, N}} \frac{2b_n}{s_n} \leq 2,$$

что не противоречит условию (8.62), но приводит в общем случае к несколько менее жесткому ограничению

$$\Delta t_k \leq \min_{n=\overline{0, N}} \frac{s_n}{b_n} \quad (8.73)$$

выбора интервала времени Δt_k .

Если в (8.71) принять $\eta = 1$, то получим *нелвную двух-слойную разностную схему* в виде СЛАУ

$$(S + A\Delta t_k)T_k = ST_{k-1} + f_k\Delta t_k \quad (8.74)$$

с трехдиагональной симметрической матрицей $S_* = S + A\Delta t_k$, имеющей положительные диагональные элементы и обладающей частичным диагональным преобладанием. Четырехточечный шаблон этой схемы для одномерной равномерной сетки изображен на рис. 8.5. Из (8.74) не удастся получить явного выражения вида (8.72) для вычисления узловых значений T_n^k , но такую СЛАУ можно решить обычным *методом прогонки*, причем *разностная схема* (8.74) в сочетании с алгоритмом метода прогонки будет *корректной*. Отметим, что при $\Delta t_k = \Delta t = \text{const}$ коэффициенты μ_n в (8.86) не будут зависеть от номера

k интервала времени. Это существенно уменьшает число арифметических операций при решении нестационарной краевой задачи.

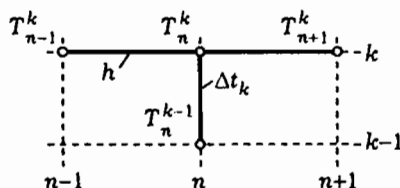


Рис. 8.5

Выбор в (8.71) $\eta = 1/2$ приводит к так называемой **двухслойной симметричной разностной схеме**. Она также является неявной и соответствует СЛАУ

$$\left(S + A \frac{\Delta t_k}{2}\right) T_k = \left(S - A \frac{\Delta t_k}{2}\right) T_{k-1} + (f_{k-1} + f_k) \frac{\Delta t_k}{2}$$

относительно вектора T_k узловых значений температуры в конце интервала времени Δt_k . Трехдиагональная симметрическая матрица $S'_k = S + \frac{1}{2} A \Delta t_k$ этой СЛАУ также имеет положительные диагональные элементы и частичное диагональное преобладание, что обеспечивает корректность разностной схемы при $\eta = 1/2$ в сочетании с методом прогонки. Поскольку эта схема эквивалентна (8.65), то она устойчива при любых значениях Δt_k , а погрешность аппроксимации этой схемой производных по времени пропорциональна $(\Delta t_k)^2$. Шеститочечный шаблон этой схемы аналогичен изображенному на рис. 8.4.

Замечание 8.4. Отметим, что формально устойчивая разностная схема может подавлять возникающие вычислительные погрешности двумя путями: затухающими колебаниями и монотонно. Первый путь не отвечает физическому содержанию задачи теплопроводности, так как при отсутствии внешнего теплового воздействия температура любой точки тела стремится к установившемуся значению монотонно, без колебаний

около этого значения. Поэтому для получения при численном решении задачи нестационарной теплопроводности с применением (8.71) физически оправданных результатов следует выбирать Δt_k из условия монотонного затухания погрешности. Можно показать*, что схема (8.71) в случае равномерной пространственной сетки с шагом h обладает таким свойством при $\Delta t_k \leq \frac{h^2}{\bar{a}(4-4\eta)}$. #

Производную по времени в (8.56) можно аппроксимировать не только при помощи разности узловых значений T_n в начале и конце k -го интервала времени Δt_k , но использовать для этого известные значения T_n в предшествующие моменты времени. Такая аппроксимация позволяет повысить порядок погрешности по Δt_k , но требует в процессе решения задачи хранить в памяти ЭВМ результаты вычислений на нескольких предыдущих интервалах. Если использовать центральную конечную разность и скорость изменения температуры в пределах удвоенного интервала времени $2\Delta t = t_k - t_{k-2}$ вычислить в момент времени t_{k-1} , то из (8.56) с учетом (8.53) получим **явную трехслойную разностную схему**

$$T_n^k = T_n^{k-2} + 2(a_n T_{n-1}^{k-1} - b_n T_n^{k-1} + c_n T_{n+1}^{k-1} + f_n^{k-1})\Delta t. \quad (8.75)$$

Эта схема неустойчива при любых значениях Δt , так как сумма модулей коэффициентов при значениях температур в правой части (8.75) превышает единицу.

Замена $T_n^{k-1} \approx \frac{1}{2}(T_n^k + T_n^{k-2})$ в (8.75) приводит к явной трехслойной схеме

$$T_n^k = \frac{(1 - 2b_n\Delta t)T_n^{k-2} + 2(a_n T_{n-1}^{k-1} + c_n T_{n+1}^{k-1} + f_n^{k-1})\Delta t}{1 + 2b_n\Delta t}, \quad (8.76)$$

устойчивой в силу (8.54) и (8.55) при любых значениях Δt , поскольку теперь сумма модулей коэффициентов при значениях

*См., например: Зарубин В.С., Селиванов В.В.

температур в правой части не превышает единицы. Однако эта схема при больших значениях Δt становится неточной, узловые значения температуры от интервала к интервалу изменяются немонотонно даже при отсутствии внешнего теплового воздействия, и результаты расчета теряют физический смысл (см. замечание 8.4).

С учетом физических соображений выбор больших значений Δt допускает **невязная трехслойная разностная схема**, в матричной записи имеющая вид

$$\frac{T_k - T_{k-2}}{2\Delta t} = \frac{1}{3}A(T_k + T_{k-1} + T_{k-2}) + f_{k-1}. \quad (8.77)$$

Она удобна, в частности, при решении нелинейных задач, когда c , λ и $I_V^{(q)}$ в (8.49) произвольным образом зависят от температуры. В этом случае элементы матрицы A и координаты вектора f_{k-1} вычисляют по известным узловым значениям температуры в момент времени t_{k-1} и, используя метод прогонки, находят узловые значения температуры в момент времени t_k . Однако при расчете на первом интервале по времени ($k = 1$) вектор T_{k-2} неизвестен. Поэтому применение трехслойной разностной схемы при решении задачи теплопроводности возможно лишь со второго интервала, а расчет на первом интервале следует провести по какой-либо двухслойной разностной схеме.

8.4. Некоторые динамические задачи

Трехслойные разностные схемы применяют при решении *методом конечных разностей* (МКР) динамических краевых задач, описываемых дифференциальными уравнениями со второй производной искомой функции по времени t . Например, при распространении в идеальной (невязкой) среде волн малой амплитуды функция $u(t, x)$ перемещения частиц среды удовлетворяет волновому уравнению

$$\frac{\partial^2 u(t, x)}{\partial t^2} = a^2 \frac{\partial^2 u(t, x)}{\partial x^2}, \quad t > 0, \quad x \in (0, l), \quad (8.78)$$

где a — скорость звука в этой среде. Примем для определенности на концах отрезка $[0, l]$ граничные условия в виде $u(t, 0) = f_0(t)$ и $u(t, l) = f_l(t)$, а начальные условия

$$u(0, x) = u^0(x), \quad \left. \frac{\partial u(t, x)}{\partial t} \right|_{t=0} = v^0(x) \quad (8.79)$$

согласуем с граничными:

$$\begin{cases} u^0(0) = f_0(0), & u^0(l) = f_l(0), \\ v^0(0) = f'_0(t)|_{t=0}, & v^0(l) = f'_l(t)|_{t=0}. \end{cases} \quad (8.80)$$

На равномерной одномерной сетке с шагом h и узлами $x_n = nh$, $n = \overline{0, N}$, при постоянном интервале времени $\Delta t = \tau = \text{const}$ можно аппроксимировать (8.78) **трехслойной симметричной разностной схемой**

$$\begin{aligned} \frac{u_n^k - 2u_n^{k-1} + u_n^{k-2}}{a^2 \tau^2} = \eta \frac{u_{n-1}^k - 2u_n^k + u_{n+1}^k}{h^2} + \\ + \eta \frac{u_{n-1}^{k-2} - 2u_n^{k-2} + u_{n+1}^{k-2}}{h^2} + (1 - 2\eta) \frac{u_{n-1}^{k-1} - 2u_n^{k-1} + u_{n+1}^{k-1}}{h^2}, \end{aligned} \quad (8.81)$$

$n = \overline{1, N-1}$, с весовыми коэффициентами η и $1 - 2\eta$. Кроме того, из граничных условий имеем $u_0^k = f_0(t_k)$ и $u_N^k = f_l(t_k)$, где $t_k = k\tau$, $k \in \mathbb{N}$, а из первого условия (8.79) — $u_0^0 = u^0(x_n)$, $n = \overline{0, N}$.

Так как в (8.81) аппроксимация вторых частных производных искомой функции $u(t, x)$, входящих в (8.78), имеет второй порядок погрешности (см. 7.2), то целесообразно, чтобы аппроксимация производной в (8.79) имела бы такой же порядок погрешности. Для этого используем *центральную конечную разность* и запишем

$$v^0(x_n) = v_n^0 = \left. \frac{\partial u(t, x_n)}{\partial t} \right|_{t=0} \approx \frac{u_n^1 - u_n^{-1}}{2\tau}, \quad n = \overline{0, N},$$

где верхним индексом -1 отмечены узловые значения u_n функции $u(t, x)$ на фиктивном слое, соответствующем моменту вре-

мени $t = -\tau$. Отсюда находим $u_n^{-1} = u_n^1 - 2v_n^0\tau$ и после подстановки в (8.81) при $k = 1$ получаем СЛАУ

$$\begin{aligned} & -\eta\gamma^2 u_{n-1}^1 + (2\eta\gamma^2 + 1)u_n^1 - \eta\gamma^2 u_{n+1}^1 = \\ & = u_n^0 + v_n^0\tau + \frac{1-2\eta}{2}\gamma^2(u_{n-1}^0 - 2u_n^0 + u_{n+1}^0) - \\ & \quad - \eta\gamma^2(v_{n-1}^0 - 2v_n^0 + v_{n+1}^0)\tau \quad (8.82) \end{aligned}$$

относительно неизвестных значений u_n^1 , $n = \overline{1, N-1}$.

СЛАУ (8.82) имеет *трехдиагональную* симметрическую матрицу с одинаковыми диагональными элементами $b_n = 2\eta\gamma^2 + 1$, где $\gamma = \frac{a\tau}{h}$, а все ненулевые внедиагональные элементы также одинаковы и равны $-\eta\gamma^2$, т.е. при $\eta > 0$ и любых значениях γ она является матрицей с диагональным преобладанием. Следовательно, СЛАУ (8.82) имеет единственное решение, которое нетрудно получить с учетом известных из граничных условий значений $u_0^1 = f_0(\tau)$ и $u_N^1 = f_1(\tau)$ обычным методом прогонки. Затем тем же методом можно при $k \geq 2$ решить СЛАУ

$$\begin{aligned} & -\eta\gamma^2 u_{n-1}^k + (2\eta\gamma^2 + 1)u_n^k - \eta\gamma^2 u_{n+1}^k = \\ & = \eta\gamma^2(u_{n-1}^{k-2} - 2u_n^{k-2} + u_{n+1}^{k-2}) + \\ & + (1 - 2\eta)\gamma^2(u_{n-1}^{k-1} - 2u_n^{k-1} + u_{n+1}^{k-1}) + 2u_n^{k-1} - u_n^{k-2}, \quad (8.83) \end{aligned}$$

которая следует из (8.81). СЛАУ (8.83) также имеет матрицу с диагональным преобладанием, что обеспечивает вычислительную устойчивость алгоритма метода прогонки.

Можно показать*, что разностная схема (8.81) в сочетании с методом прогонки при условии $\gamma^2 = \frac{a\tau^2}{h^2} < \frac{1}{1-4\eta}$ обладает и устойчивостью по входным данным. При $\eta = 1/4$ это условие выполняется для любых значений интервала времени τ . В частном случае $\eta = 0$ (8.82) и (8.83) явно разрешимы относительно

*См. Самарский А А

искомых *узловых значений* u_n^k , $n = \overline{1, N-1}$, $k \in \mathbb{N}$, но при этом возникает ограничение $\tau < \frac{h}{a}$ на выбор интервала времени.

При невыполнении равенств (8.80), выражающих согласование начальных и граничных условий, волновое уравнение (8.78) может иметь негладкие решения $u(t, x)$ с разрывами частных производных на линиях $x \pm at = \text{const}$, называемых характеристиками волнового уравнения [XII]. Негладкие решения, имеющие разрывы не только производных, но и самих искомых функций, характерны для задач газовой динамики. Для построения таких решений используют специальные разностные схемы*.

Рассмотрим особенности построения разностной схемы для динамической задачи, описываемой системой уравнений с несколькими неизвестными функциями. В качестве примера обратимся к одномерной математической модели магнитной гидродинамики (см. пример 3.3), содержащей систему нелинейных уравнений

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho v}{\partial x_1} = 0, \quad \rho \frac{dv}{dt} + \frac{\partial}{\partial x_1} \left(p(\rho) + \frac{\mu_0}{2} H^2 \right) = 0, \quad \frac{\partial H}{\partial t} + \frac{\partial v H}{\partial x_1} = 0$$

с тремя неизвестными функциями: плотностью ρ , скоростью среды v и напряженностью магнитного поля H , зависящими от времени t и координаты x_1 (μ_0 — магнитная постоянная, зависимость $p(\rho)$ давления среды от плотности считаем заданной).

Используя первое уравнение, преобразуем первое слагаемое во втором уравнении:

$$\rho \frac{dv}{dt} = \rho \frac{\partial v}{\partial t} + \rho v \frac{\partial v}{\partial x_1} = \frac{\partial \rho v}{\partial t} + \frac{\partial \rho v^2}{\partial x_1}.$$

Для сокращения записи введем матрицу-столбец $U = (\rho \quad \rho v \quad H)^T$ размера 3×1 и представим систему уравнений в матричном

*См., например: Самарский А.А., Попов Ю.П., а также: Пирумов У.Г., Росляков Г.С. или Физтмайер Р., Мортон К

виде

$$\frac{\partial U}{\partial t} + \frac{\partial v U}{\partial x_1} = -\frac{\partial W}{\partial x_1}, \quad (8.84)$$

где $W = \left(0 \quad p(\rho) + \mu_0 \frac{H^2}{2} \quad 0 \right)^T$ — матрица-столбец.

Для решения системы (8.84) при помощи МКР можно использовать как явные, так и неявные разностные схемы. Однако в случае неявной схемы неизвестные величины в конце каждого интервала по времени в силу нелинейности системы приходится находить последовательными приближениями. Отметим, что каждую из итераций можно выполнить при помощи *матричной прогонки*.

Здесь ограничимся описанием явной **разностной схемы типа предиктор-корректор**, алгоритм которой на каждом интервале времени состоит из двух этапов. Сначала по известным в момент времени t_{k-1} , т.е. в начале k -го интервала времени Δt_k , значениям элементов матриц U и W в узлах одномерной равномерной сетки с шагом h находят значения элементов матрицы U в так называемых полуцелых узлах пространственно-временной сетки (рис. 8.6):

$$U_{n\pm 1/2}^{k-1/2} = \frac{U_n^{k-1} + U_{n\pm 1}^{k-1}}{2} - \frac{\Delta t_k}{2h} (v_{n\pm 1}^{k-1} U_{n\pm 1}^{k-1} - v_n^{k-1} U_n^{k-1} + W_{n\pm 1}^{k-1} - W_n^{k-1}).$$

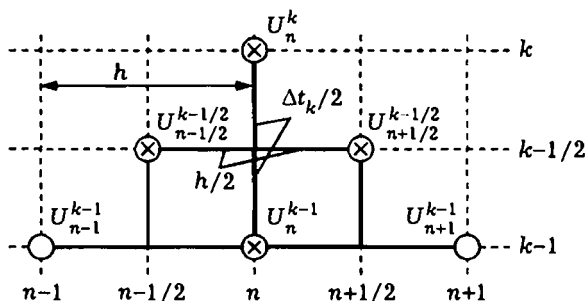


Рис. 8.6

Эта явная схема („предиктор“), в которой для аппроксимации производных использованы разности вперед по времени t и центральные разности по координате x_1 , позволяет с погрешностью $O(\Delta t_k, h^2)$ получить прогноз для значений элементов матрицы U в момент времени $t_{k-1/2}$ в середине интервала Δt_k в *промежуточных узлах* сетки по координате x_1 . По этим значениям в тех же полущелых узлах можно вычислить значения скорости v и единственного отличного от нуля элемента $p(\rho) + \frac{1}{2}\mu_0 H^2$ матрицы W и затем перейти ко второму этапу („корректор“):

$$U_n^k = U_n^{k-1} - \frac{\Delta t_k}{h} (v_{n+1/2}^{k-1/2} U_{n+1/2}^{k-1/2} - v_{n-1/2}^{k-1/2} U_{n-1/2}^{k-1/2} + W_{n+1/2}^{k-1/2} - W_{n-1/2}^{k-1/2}).$$

Шаблон этой явной схемы носит название „крест“ (узлы пространственно-временной сетки, входящие в этот шаблон, отмечены на рис. 8.6 крестиком в кружке).

Можно показать*, что несмотря на погрешность $O(\Delta t_k, h^2)$ первого этапа после выполнения второго этапа итоговая погрешность имеет второй порядок как по времени, так и по пространственной координате, поскольку главные члены погрешности, возникшие при аппроксимации производной по времени на этапе „предиктор“, компенсируются затем на этапе „корректор“. Наряду с удобством вычислений по явным схемам описанная разностная схема имеет ограничение на выбор допустимого интервала времени: $\Delta t_k \leq \frac{h}{a + |v|}$, где $a = \sqrt{\frac{dp(\rho)}{d\rho}}$.

Дополнение 8.1. Модификации метода прогонки

Выше (см. 7.4) показано, что выполнение условия (7.26) достаточно для существования и единственности решения системы линейных алгебраических уравнений (СЛАУ) с *трех-*

*См., например Пирумов У.Г., Росляков Г.С.

диагональной матрицей A . При выполнении этого условия A является матрицей с частичным диагональным преобладанием, а алгоритм метода прогонки — устойчивым. Рассмотрим эти вопросы применительно к решению СЛАУ (8.10) с трехдиагональной матрицей (8.11) при более слабых ограничениях

$$|b_0| \geq |c_0|, \quad |b_N| \geq |a_N|, \quad |b_n| \geq |a_n| + |c_n|, \quad n = \overline{1, N-1}, \quad (8.85)$$

причем $b_0 b_N \neq 0$ и в последнем неравенстве $a_n c_n \neq 0$, а хотя бы в одном из них должно быть соблюдено строгое неравенство. В частных случаях задания узловых значений u_0 и (или) u_N искомой функции $u(x)$ имеем в (8.8), (8.11) и (8.85) $c_0 = 0$, $u_0 = y_0/b_0$ и (или) $a_N = 0$, $u_N = y_N/b_N$.

Для построения алгоритма решения СЛАУ (8.10) с матрицей (8.11) используем представление (7.27) в виде

$$u_n = \mu_n u_{n+1} + \nu_n, \quad n = \overline{0, N-1}. \quad (8.86)$$

Из первого равенства (8.8) имеем $\mu_0 = c_0/b_0$ и $\nu_0 = y_0/b_0$, а далее из (8.9) и (8.86) находим

$$\mu_n = \frac{c_n}{b_n - a_n \mu_{n-1}}, \quad \nu_n = \frac{y_n + a_n \nu_{n-1}}{b_n - a_n \mu_{n-1}}, \quad n = \overline{1, N-1}. \quad (8.87)$$

Используя уравнения (8.86) и (8.87) для $n = N-1$ и подставляя во второе равенство (8.8) представление $u_{N-1} = \mu_{N-1} u_N + \nu_{N-1}$, получаем

$$u_N = \frac{y_N + a_N \nu_{N-1}}{b_N - a_N \mu_{N-1}}, \quad (8.88)$$

что позволяет затем при помощи (8.86) и предварительно вычисленных по (8.87) коэффициентов μ_n и ν_n последовательно найти остальные значения u_n для $n = N-1, N-2, \dots, 1, 0$.

Непосредственным перемножением нижней

$$A_1 = \begin{pmatrix} b_0 & 0 & 0 & \dots & 0 & 0 \\ -a_1 & \Delta_1 & 0 & \dots & 0 & 0 \\ 0 & -a_2 & \Delta_2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \Delta_{N-1} & 0 \\ 0 & 0 & 0 & \dots & -a_N & \Delta_N \end{pmatrix}$$

и верхней

$$A_2 = \begin{pmatrix} 1 & -\mu_0 & 0 & \dots & 0 & 0 \\ 0 & 1 & -\mu_1 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 & -\mu_{N-1} \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

треугольных матриц, где $\Delta_n = b_n - a_n \mu_{n-1}$, $n = \overline{1, N}$, — знаменатели в соотношениях (8.87), можно показать, что $A = A_1 A_2$. Поскольку $\det A = \det A_1 \det A_2$ и $\det A_2 = 1$, получим

$$\det A = \det A_1 = b_0 \prod_{n=1}^N \Delta_n. \quad (8.89)$$

Таким образом, условие $\det A \neq 0$ существования обратной матрицы A^{-1} будет выполнено, если при $b_0 \neq 0$ выполнены условия $\Delta_n \neq 0$, $n = \overline{1, N}$.

Алгоритм метода прогонки можно построить и в обратном направлении, положив

$$u_n = \mu'_n u_{n-1} + \nu'_n, \quad n = N, N-1, \dots, 2, 1, \quad (8.90)$$

причем $\mu'_N = a_N/b_N$, $\nu'_N = y_N/b_N$ и

$$\mu'_n = \frac{a_n}{b_n - c_n \mu'_{n+1}}, \quad \nu'_n = \frac{y_n + c_n \nu'_{n+1}}{b_n - c_n \mu'_{n+1}}, \quad n = N-1, \dots, 1. \quad (8.91)$$

Тогда получим

$$u_0 = \frac{y_0 + c_0 \nu'_1}{b_0 - c_0 \mu'_1}, \quad (8.92)$$

и вместо (8.89) будем иметь

$$\det A = b_N \prod_{n=0}^{N-1} \Delta'_n, \quad \Delta'_n = b_n - c_n \mu'_{n+1}.$$

Отсюда следует, что для существования матрицы A^{-1} необходимо, чтобы $\Delta'_n \neq 0$, $n = \overline{0, N-1}$, при $b_N \neq 0$.

Покажем, что для обеспечения $\Delta_n \neq 0$ или $\Delta'_n \neq 0$ достаточно выполнения неравенств (8.85). Действительно, поскольку $|\mu_0| = \left| \frac{c_0}{b_0} \right| \leq 1$, с учетом третьего неравенства (8.85) при $n = 1$ имеем

$$|\Delta_1| = |b_1 - a_1 \mu_0| \geq |b_1| - |a_1| |\mu_0| \geq |c_1| + |a_1| (1 - |\mu_0|) > 0, \quad (8.93)$$

т.е. $|\Delta_1| \geq |c_1|$ и $\Delta_1 \neq 0$. Отсюда, используя (8.87), находим

$$|\mu_1| = \frac{|c_1|}{|\Delta_1|} \leq \frac{|c_1|}{|c_1|} = 1. \quad (8.94)$$

Увеличивая номер n , последовательно устанавливаем, что $\Delta_n \neq 0$ и $|\mu_n| \leq 1$ при $n < N$. Для выполнения условия $\Delta_N \neq 0$ требуется, чтобы в (8.85) имело место хотя бы одно строгое неравенство. Если $|b_N| > |a_N|$, то при $|\mu_{N-1}| \leq 1$ получаем

$$|\Delta_N| = |b_N - a_N \mu_{N-1}| \geq |b_N| - |a_N| |\mu_{N-1}| > 0 \quad (8.95)$$

и $\Delta_N \neq 0$. Если $|b_0| > |c_0|$, то $|\mu_0| < 1$ и в соответствии с (8.93) и (8.94) $|\Delta_1| > |c_1|$, $|\mu_1| < 1$, а значит, и $|\mu_n| < 1$, $n = \overline{2, N-1}$, в том числе $|\mu_{N-1}| < 1$. Теперь (8.95) справедливо даже при $|b_N| = |a_N|$. Наконец, в случае $|b_m| > |a_m| + |c_m|$ при $1 \leq m \leq N-1$ с учетом $|\mu_{m-1}| \leq 1$ получим

$$\begin{aligned} |\Delta_m| &= |b_m - a_m \mu_{m-1}| \geq |b_m| - |a_m| |\mu_{m-1}| > \\ &> |c_m| + |a_m| (1 - |\mu_{m-1}|) \geq |c_m|, \end{aligned}$$

т.е. $|\Delta_m| > |c_m|$, $|\Delta_m| \neq 0$ и $|\mu_m| = \frac{|c_m|}{|\Delta_m|} < \frac{|c_m|}{|c_m|} = 1$, а затем, увеличивая номер $n = \overline{m+1, N-1}$, будем иметь $|\mu_n| < 1$, в том числе $|\mu_{N-1}| < 1$, так что неравенство (8.95) справедливо даже при $|b_N| = |a_N|$.

Таким образом, условия (8.85) с учетом $b_0 b_N \neq 0$ гарантируют существование и единственность решения СЛАУ (8.10) с матрицей (8.11) и попутно являются достаточными для выполнения неравенств $|\mu_n| \leq 1$, $n = \overline{0, N-1}$, что формально обеспечивает устойчивость процесса вычислений при использовании рекуррентного соотношения (8.86). При выполнении этих условий говорят, что **разностная схема** (8.8), (8.9) в сочетании с алгоритмом метода прогонки является **корректной**.

Отметим, что условия (8.85) можно ослабить, допустив $a_n = 0$ или $c_n = 0$ для некоторых $n = \overline{1, N-1}$. В самом деле, пусть, например, $a_m = 0$ при таком m , что $1 \leq m \leq N-1$. Тогда (8.8), (8.9) можно представить в виде двух СЛАУ:

$$\begin{aligned} b_m u_m - c_m u_{m+1} &= y_m, & -a_N u_{N-1} + b_N u_N &= y_N, \\ -a_n u_{n-1} + b_n u_n - c_n u_{n+1} &= y_n, & n &= \overline{m+1, N-1}, \end{aligned}$$

относительно неизвестных u_n , $n = \overline{m, N}$, и

$$-a_n u_{n-1} + b_n u_n - c_n u_{n+1} = y_n, \quad n = \overline{1, m-2},$$

$$b_0 u_0 - c_0 u_1 = y_0, \quad -a_{m-1} u_{m-2} + b_{m-1} u_{m-1} = y_{m-1} + c_{m-1} u_m$$

относительно неизвестных u_n , $n = \overline{0, m-1}$. Каждая из этих СЛАУ соответствует корректной разностной схеме, но обе могут быть решены сквозной прогонкой путем вычисления коэффициентов (8.87) и нахождения значений u_N из (8.88) и u_n , $n = \overline{N-1, 0}$, из (8.86).

Подсчет количества арифметических операций в соотношениях (8.86)–(8.88) показывает, что в процессе вычислений следует выполнить $3N$ умножений, $2N+1$ делений и $3N$ сложений и вычитаний. При реализации алгоритма на ЭВМ наиболее медленной операцией является деление. Вместо $2N+1$ делений

можно выполнить $N + 1$ обращений знаменателя Δ_n коэффициентов μ_n и ν_n , но тогда возрастет на $2N + 1$ число операций умножения. Если же не делать различия между арифметическими операциями, то общее их количество составит $8N + 1$, причем из них $3N - 2$ операций уходит на вычисление коэффициентов μ_n , $n = \overline{0, N-1}$.

Отметим, что возможность построения алгоритма метода прогонки как на основе (8.86), (8.87), так и на основе (8.90), (8.91) позволяет уменьшить суммарное количество арифметических операций, если при решении задачи представляют интерес не все значения u_n , $n = \overline{0, N}$, а лишь одно значение u_m . Пусть $m \approx N/2$. Тогда, используя (8.87), находим μ_m и ν_m , при помощи (8.91) вычисляем μ'_{m+1} и ν'_{m+1} , а затем из (8.86) при $n = m$ и из (8.90) при $n = m + 1$ получаем систему из двух алгебраических уравнений

$$u_m = \mu_m u_{m+1} + \nu_m, \quad u_{m+1} = \mu'_{m+1} u_m + \nu'_{m+1},$$

откуда следует

$$u_m = \frac{\mu_m \nu'_{m+1} + \nu_m}{1 - \mu_m \mu'_{m+1}}.$$

В этом алгоритме, называемом **встречной прогонкой**, удастся сэкономить примерно N арифметических операций по сравнению с обычным методом прогонки. Экономия будет более существенной, если коэффициенты a_n , b_n и c_n в (8.9) постоянны, а в (8.8) $c_0 = a_N = 0$, поскольку тогда $\mu'_{N-n} = \mu_n$, $n = \overline{0, m}$.

В ряде задач математической физики помимо значений u_n в узлах разностной схемы представляют интерес и значения

$$q_{n-1/2} = -a_n(u_n - u_{n-1}), \quad n = \overline{1, N}, \quad (8.96)$$

пропорциональные приближенным значениям производной $u'(x)$ искомой функции $u(x)$. Если $u(x)$ имеет смысл перемещения в твердом теле, то $u'(x)$ является деформацией. Тогда $q_{n-1/2}$ в (8.96) имеет смысл приближенного значения напряжения в точ-

ке $x_{n-1/2}$. В теплофизических задачах $u(x)$ — распределение температуры, а $q_{n-1/2}$ в (8.96) по абсолютной величине соответствует плотности теплового потока при $x = x_{n-1/2}$. В задачах электрофизики $u(x)$ может иметь смысл электрического потенциала. Тогда $u'(x)$ по абсолютной величине соответствует напряженности электрического поля, а $q_{n-1/2}$ — плотности электрического тока в точке $x_{n-1/2}$.

Для неоднородной среды коэффициенты в (8.9) и в том числе a_n могут резко изменять свои значения. В таком случае использование (8.96) для вычисления $q_{n-1/2}$ по предварительно найденным значениям u_n , $n = \overline{0, N}$, приводит к значительной потере точности. Этого можно избежать, если перейти к описанной ниже **потокковой прогонке**.

Рассмотрим случай, когда матрица (8.11) является симметрической, т.е. $c_n = a_{n+1}$, $n = \overline{0, N-1}$, и обозначим $b_n = a_n + a_{n+1} + d_n$, причем $a_0 = a_{N+1} = 0$, $d_n > 0$, $n = \overline{0, N}$, и $a_n > 0$, $n = \overline{1, N}$. Тогда вместо (8.8) и (8.9) с учетом (8.96) получим

$$\begin{cases} q_{n+1/2} - q_{n-1/2} + d_n u_n = y_n, & n = \overline{1, N-1}, \\ u_0 = \frac{a_1 u_1 + y_0}{a_1 + d_0}, & -q_{N-1/2} + d_N u_N = y_N. \end{cases} \quad (8.97)$$

Из (8.96) найдем $u_n = u_{n+1} + \frac{q_{n+1/2}}{a_{n+1}}$, $n = \overline{0, N-1}$, и подставим в (8.86), где теперь

$$\mu_n = \frac{a_{n+1}}{a_{n+1} + d_n + (1 - \mu_{n-1})a_n}, \quad \nu_n = \mu_n \frac{y_n + a_n \nu_{n-1}}{a_{n+1}}, \quad (8.98)$$

$n = \overline{0, N-1}$. В итоге будем иметь

$$q_{n+1/2} + a_{n+1}(1 - \mu_n)u_{n+1} = a_{n+1}\nu_n, \quad n = \overline{0, N-1},$$

или

$$q_{n-1/2} + \tilde{\mu}_n u_n = \tilde{\nu}_n, \quad n = \overline{1, N}, \quad (8.99)$$

где $\tilde{\mu}_n = a_n(1 - \mu_{n-1})$ и $\tilde{\nu}_n = a_n \nu_{n-1}$.

СЛАУ (8.97), (8.99) содержит $2N + 1$ уравнений с $2N + 1$ неизвестными и распадается на две независимые СЛАУ. Первую получим, если представление $u_n = \frac{\tilde{\nu}_n - q_{n-1/2}}{\tilde{\mu}_n}$, вытекающее из (8.99), подставим в (8.97):

$$q_{n-1/2} = \frac{\tilde{\mu}_n q_{n+1/2} + \tilde{\nu}_n d_n - \tilde{\mu}_n y_n}{\tilde{\mu}_n + d_n}, \quad n = \overline{1, N}. \quad (8.100)$$

Здесь $q_{N+1/2} = 0$. Из этой системы для $n = N$ находим $q_{N-1/2}$, а затем последовательно для $n = N - 1, N - 2, \dots, 1$ вычисляем и все остальные значения $q_{n-1/2}$. Вторую систему получим, если выражение $q_{n-1/2} = \tilde{\nu}_n - \tilde{\mu}_n u_n$, вытекающее из (8.99), подставим снова в (8.97):

$$u_n = \frac{\tilde{\mu}_{n+1} u_{n+1} + y_n - \tilde{\nu}_{n+1} + \tilde{\nu}_n}{\tilde{\mu}_n + d_n}, \quad n = \overline{1, N}, \quad (8.101)$$

причем $\mu_N = \mu_{N+1} = \nu_{N+1} = 0$. При $n = N$ из (8.101) вычисляем u_N , затем последовательно для $n = N - 1, N - 2, \dots, 1$ находим значения u_n , а из (8.97) по найденному значению u_1 вычисляем u_0 .

Из (8.98) и (8.99) имеем $\tilde{\mu}_1 = a_1 d_0 / (a_1 + d_0)$ и

$$\tilde{\mu}_{n+1} = \frac{a_{n+1} (d_n + a_n (1 - \mu_{n-1}))}{a_{n+1} + d_n + a_n (1 - \mu_{n-1})} = \frac{a_{n+1} (d_n + \tilde{\mu}_n)}{a_{n+1} + d_n + \tilde{\mu}_n}, \quad n = \overline{1, N}.$$

Отсюда при $a_{n+1} > 0$ и $d_n > 0$, $n = \overline{0, N-1}$, следует, что $\tilde{\mu}_n > 0$. Тогда в (8.100) при $q_{n+1/2}$ получаем коэффициент $\frac{\tilde{\mu}_n}{\tilde{\mu}_n + d_n} < 1$ что обеспечивает устойчивость вычислений по этой формуле. Но и в (8.101) при u_{n+1} коэффициент

$$\frac{\tilde{\mu}_{n+1}}{\tilde{\mu}_n + d_n} = \frac{a_{n+1}}{a_{n+1} + d_n + \tilde{\mu}_n} < 1,$$

что также гарантирует устойчивость вычислений по (8.101).

Таким образом, алгоритм потоковой прогонки обладает устойчивостью по отношению к погрешностям вычислений. Для его реализации требуется $21N + 1$ арифметических действий*, что примерно вдвое превышает число операций, необходимых для вычисления значений u_n обычным методом прогонки с последующим нахождением значений $q_{n-1/2}$ из (8.96).

В том случае, когда краевая задача для обыкновенного дифференциального уравнения (ОДУ) (8.1) имеет периодическое решение, матрица A в СЛАУ (8.10) имеет вид (8.17), т.е. не является трехдиагональной. Для решения такой СЛАУ можно применить модифицированный вариант метода прогонки, называемый **циклической прогонкой**.

Этот вариант основан на представлении узловых значений искомой функции в виде

$$u_n = v_n + u_0 w_n, \quad n = \overline{0, N}, \quad (8.102)$$

где значения v_n находят из решения системы

$$-a_n v_{n-1} + b_n v_n - c_n v_{n+1} = y_n, \quad n = \overline{1, N-1}, \quad (8.103)$$

и условий $v_0 = v_N = 0$, а значения w_n — из системы

$$-a_n w_{n-1} + b_n w_n - c_n w_{n+1} = 0, \quad n = \overline{1, N-1}, \quad (8.104)$$

и условий $w_0 = w_N = 1$. Умножением (8.104) на u_0 и почленным сложением результата с (8.103) при том же номере n можно убедиться, что с учетом (8.102) будет удовлетворено (8.9). Аналогично удастся удовлетворить и второе равенство в (8.16). Чтобы удовлетворить первое равенство в (8.16), подставим в него (8.102) при $n = 0, 1$ и $N - 1$. Тогда с учетом $v_0 = 0$ и $w_0 = 1$ запишем

$$-a_0 v_{N-1} - a_0 u_0 w_{N-1} + b_0 u_0 - c_0 v_1 - c_0 u_0 w_1 = y_0,$$

*См.. Самарский А.А., Николаев Е.С.

или

$$u_0 = \frac{y_0 + a_0 v_{N-1} + c_0 v_1}{b_0 - a_0 w_{N-1} - c_0 w_1}. \quad (8.105)$$

Решение систем (8.103) и (8.104) получают обычным методом прогонки, причем для каждого номера n коэффициент μ_n будет одинаков для обеих систем, что позволяет уменьшить общее число арифметических операций, которое при полной реализации алгоритма циклической прогонки составляет $14N - 8$. Если выполнены достаточные условия (8.85), то, как было показано ранее, гарантирована устойчивость алгоритма решения СЛАУ (8.103) и (8.104). Кроме того, можно показать*, что из этих условий следует $b_0 - a_0 w_{N-1} - c_0 w_1 \neq 0$, т.е. u_0 в соответствии с (8.105) имеет конечное значение.

Более сложной модификации приходится подвергать метод прогонки при решении класса задач, которым соответствует СЛАУ (8.10) с матрицей (8.25). Решение такой СЛАУ можно представить в виде

$$u_n = u_n^0 + u_0 v_n + u_N w_n, \quad n = \overline{0, N}, \quad (8.106)$$

где u_n^0 находят при $u_0^0 = u_N^0 = 0$ из СЛАУ

$$-a_n u_{n-1}^0 + b_n u_n^0 - c_n u_{n+1}^0 = y_n, \quad n = \overline{1, N-1}, \quad (8.107)$$

v_n при $v_0 = 1$ и $v_N = 0$ из СЛАУ

$$-a_n v_{n-1} + b_n v_n - c_n v_{n+1} = \varphi_n, \quad n = \overline{1, N-1}, \quad (8.108)$$

а w_n при $w_0 = 0$ и $w_N = 1$ из СЛАУ

$$-a_n w_{n-1} + b_n w_n - c_n w_{n+1} = \psi_n, \quad n = \overline{1, N-1}. \quad (8.109)$$

Из (8.106)–(8.109) ясно, что значения u_n , $n = \overline{1, N-1}$, удовлетворяют разностным уравнениям (8.23). Для нахождения u_0 и u_N используем оставшиеся два разностных уравнения в (8.22)

*См.: Самарский А.А., Николаев Е.С.

и (8.24), подставив в них (8.106) при $n = 0$ и $n = N$. В итоге получим систему двух уравнений относительно u_0 и u_N :

$$\begin{aligned}(b_0 - (g^\circ, v))u_0 - (\psi_0 + (g^\circ, w))u_N &= y_0 + (g^\circ, u^\circ), \\ -(\varphi_N + (g, v))u_0 - (b_N - (g, w))u_N &= y_N + (g, u^\circ).\end{aligned}$$

Здесь $(\cdot, \cdot)$ обозначает сумму $N - 1$ попарных произведений. Например,

$$(g^\circ, v) = \sum_{n=1}^{N-1} g_n^\circ v_n.$$

Если определитель этой системы

$$\begin{aligned}\Delta &= (b_0 - (g^\circ, v))(b_N - (g, w)) - \\ &\quad - (\psi_0 + (g^\circ, w))(\psi_N + (g, v)) \neq 0, \quad (8.110)\end{aligned}$$

то она имеет единственное решение

$$\begin{aligned}u_0 &= \frac{(b_N - (g, w))(y_0 + (g^\circ, u^\circ)) + (\psi_0 + (g^\circ, w))(y_N + (g, u^\circ))}{\Delta}, \\ u_N &= \frac{(\psi_N + (g, v))(y_0 + (g^\circ, u^\circ)) + (b_0 - (g^\circ, v))(y_N + (g, u^\circ))}{\Delta}.\end{aligned}$$

Выполнение условий (8.85) достаточно для устойчивости алгоритма обычного метода прогонки при решении систем (8.107)–(8.109). Можно показать* справедливость (8.110) при более ограничительных по сравнению с (8.85) неравенствах

$$\begin{cases} |b_0| \geq |\psi_0| + \sum_{n=1}^{N-1} |g_n^\circ|, & |b_N| \geq |\varphi_N| + \sum_{n=1}^{N-1} |g_n|, & b_0 b_N \neq 0, \\ |b_n| \geq |a_n| + |c_n| + |\varphi_n| + |\psi_n|, & a_n c_n \neq 0, & n = \overline{1, N-1}, \end{cases} \quad (8.111)$$

причем хотя бы в одном из них должно быть выполнено строгое неравенство. Вместе с тем, если известно, что СЛАУ

*См. Самарский А. А., Николаев Е. С.

(8.22)–(8.24) имеет единственное решение, то справедливо соотношение (8.110), и u_0 и u_N имеют единственные значения даже при нарушении условий (8.111).

Рассмотренные варианты метода прогонки для решения СЛАУ (8.10) можно применять, когда A является *матрицей с частичным диагональным преобладанием*, т.е. выполнены условия (8.85). При их нарушении нельзя гарантировать отличие от нуля знаменателя коэффициентов μ_n и ν_n в (8.87) и справедливость неравенства $\mu_n \leq 1$, обеспечивающего устойчивость описанного алгоритма метода прогонки по отношению к вычислительным погрешностям. Удобство такого алгоритма состоит в использовании простых рекуррентных формул, реализующих для СЛАУ с трехдиагональной матрицей метод Гаусса без выбора главного элемента по строкам матрицы, поскольку при частичном диагональном преобладании главный элемент расположен на главной диагонали. Этот алгоритм связан с монотонным изменением номера узла сначала в одном направлении, а затем — в противоположном. Если СЛАУ (8.10) имеет единственное решение, но условие (8.85) нарушено, то можно построить алгоритм **немонотонной прогонки***, который основан на *методе Гаусса с выбором главного элемента*. В этом случае может не соблюдаться монотонное изменение номера при нахождении узловых значений искомой функции.

Метод прогонки можно использовать для решения последовательными приближениями нелинейных краевых задач. Рассмотрим нелинейное ОДУ второго порядка

$$u''(x) + f(x, u) = 0, \quad x \in [0, l], \quad (8.112)$$

где $f(x, u)$ — непрерывно дифференцируемая по u и непрерывная по x действительная функция, при заданных значениях

$$u(0) = u_0 \quad \text{и} \quad u(l) = u_l \quad (8.113)$$

искомой функции $u(x)$ на концах отрезка $[0, l]$.

*См.: Самарский А.А., Николаев Е.С.

Аппроксимируя при помощи (7.3) вторую производную в (8.112) на равномерной одномерной сетке с шагом h и узлами $x_n = nh$, $n = \overline{0, N}$, получаем систему разностных уравнений

$$-u_{n-1} + 2u_n - u_{n+1} = h^2 f_n, \quad n = \overline{1, N-1}, \quad (8.114)$$

где $f_n = f(x_n, u_n)$, $x_n = nh$. Эта система включает $N - 1$ нелинейных конечных уравнений относительно $N - 1$ неизвестных узловых значений u_n . Алгоритм решения системы (8.114) с помощью последовательных приближений можно построить различным образом, но он должен обеспечивать сходимость итераций к искомому решению.

Пусть узловые значения $u_n^{(k-1)}$ получены на $(k - 1)$ -й итерации (при $k = 1$ они соответствуют некоторому начальному приближению). Тогда на k -й итерации (8.114) перейдет в систему уравнений

$$-u_{n-1}^{(k)} + 2u_n^{(k)} - u_{n+1}^{(k)} = h^2 f_n^{(k-1)}, \quad n = \overline{1, N-1}, \quad (8.115)$$

где $f_n^{(k-1)} = f(x_n, u_n^{(k-1)})$, образующих СЛАУ с квадратной трехдиагональной матрицей

$$A = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 \\ 0 & -1 & 2 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & 1 & 2 \end{pmatrix}$$

порядка $N - 1$. Решая эту СЛАУ методом прогонки и используя (8.86), с учетом граничного условия $u_N^{(k)} = u_l$ последовательно находим

$$u_n^{(k)} = \mu_n u_{n+1}^{(k)} + \nu_n^{(k)}, \quad n = N - 1, N - 2, \dots, 2, 1, \quad (8.116)$$

где в соответствии с (8.87) при $a_n = c_n = 1$, $b_n = 2$, $n = \overline{1, N-1}$, $a_0 = \mu_0 = 0$ и $\nu_0^{(k)} = u_0$

$$\mu_n = \frac{1}{2 - \mu_{n-1}} = \frac{n}{n+1} < 1, \quad \nu_n^{(k)} = \mu_n (h^2 f_n^{(k-1)} + \nu_{n-1}^{(k)}). \quad (8.117)$$

Обозначим $z_n^{(k)} = u_n^{(k)} - u_n^*$ погрешность на k -й итерации, где u_n^* — узловые значения, соответствующие истинному решению системы (8.114), т.е. удовлетворяющие равенствам

$$-u_{n-1}^* + 2u_n^* - u_{n+1}^* = h^2 f_n^* = h^2 f(x_n, u_n^*), \quad n = \overline{1, N-1}.$$

Вычитая эти равенства почленно из (8.115), получаем СЛАУ

$$-z_{n-1}^{(k)} + 2z_n^{(k)} - z_{n+1}^{(k)} = d_n^{(k-1)}, \quad n = \overline{1, N-1},$$

где

$$d_n^{(k)} = h^2 (f_n^{(k-1)} - f_n^*) \approx h^2 \frac{\partial f(x_n, u)}{\partial u} \Big|_{u=u_n^*} z_n^{(k-1)}. \quad (8.118)$$

Решая эту СЛАУ также методом прогонки и учитывая (8.116) и (8.117), при $z_0^{(k)} = z_N^{(k)} = 0$ находим

$$z_n^{(k)} = \mu_n z_{n+1}^{(k)} + \nu_n^{(k)} = n \sum_{j=n+1}^N \frac{1}{j(j-1)} \sum_{m=1}^{j-1} m d_m^{(k)}. \quad (8.119)$$

Используя (8.118), введем оценку

$$|d_n^{(k)}| \leq h^2 |z_n^{(k-1)}| \max_{n=\overline{1, N-1}} \left| \frac{\partial f(x_n, u)}{\partial u} \right| = h^2 M_1 |z_n^{(k-1)}|. \quad (8.120)$$

Тогда из (8.119) будем иметь

$$|z_n^{(k)}| \leq h^2 M_1 |z_n^{(k-1)}| n \sum_{j=n+1}^N \frac{1}{j(j-1)} \sum_{m=1}^{j-1} m = n \frac{N-n}{2} h^2 M_1 |z_n^{(k-1)}|.$$

Отсюда, учитывая, что $\max_{n=1, N-1} n(N-n) \leq \frac{N^2}{4}$ и $Nh = l$, получаем

$$|z_n^{(k)}| \leq \frac{N^2 h^2}{8} M_1 |z_n^{(k-1)}| = \frac{l^2 M_1}{8} |z_n^{(k-1)}|.$$

Таким образом, достаточным условием сходимости алгоритма последовательных приближений является выполнение неравенства $\frac{1}{8} l^2 M_1 < 1$, не зависящего от числа узлов сетки.

Пример 8.4. Рассмотрим прямой стержень длиной L с круговым поперечным сечением радиуса R , излучающий с боковой поверхности. Температура T_0 торцов стержня (при $x = 0$ и $x = L$) поддерживается постоянной. Если принять, что температура в поперечном сечении стержня однородна и изменяется лишь вдоль координаты $x \in [0, L]$, то установившееся распределение температуры $T(x)$ будет описываться нелинейным ОДУ

$$\pi R^2 \lambda \frac{d^2 T(x)}{dx^2} - 2\pi R \varepsilon \sigma_0 T^4(x) = 0, \quad x \in (0, L), \quad (8.121)$$

где λ — коэффициент теплопроводности материала стержня; ε — коэффициент излучения его боковой поверхности; $\sigma_0 = 5,67 \cdot 10^{-8} \frac{\text{Вт}}{\text{м}^2 \cdot \text{К}^4}$ — постоянная Стефана — Больцмана (см. пример 3.5). Из физического смысла задачи ясно, что $T(x) \leq T_0$ при $x \in (0, 1)$.

Положив $u = T/T_0$, приведем (8.121) к безразмерному виду

$$\frac{d^2 u(\xi)}{d\xi^2} - C u^4(\xi) = 0, \quad \xi = \frac{x}{L} \in (0, 1), \quad C = 2L^2 \frac{\varepsilon \sigma_0}{R\lambda} T_0^3, \quad (8.122)$$

и запишем граничные условия: $u(0) = u(1) = 1$. Из сравнения (8.48) с (8.122) следует, что в данном случае $f(x, u) = -Cu^4$, так что $\frac{\partial f(x, u)}{\partial u} = -4Cu^3$ и в соответствии с (8.120) при $l = 1$ и $u \leq 1$ имеем

$$\frac{l^2}{8} M_1 = \frac{1}{8} \max_{n=1, N-1} \left| \frac{\partial f(x_n, u)}{\partial u} \right| = \frac{1}{8} \max_{n=1, N-1} |-4Cu^3| = \frac{C}{2}.$$

Следовательно, для сходимости изложенного выше алгоритма последовательных приближений достаточно выполнения неравенства $C/2 < 1$. В табл. 8.1 приведены значения $u_n \cdot 10^3$, $n = \overline{1, N/2}$, рассчитанные при $N = 20$ и $C = 2$ (при этом учтена симметрия распределения температуры относительно середины стержня). Отметим, что положение, когда значения u_n на двух последовательных итерациях в каждом узле отличаются меньше чем на 10^{-6} , было достигнуто за 20 итераций.

Таблица 8.1

C	n									
	1	2	3	4	5	6	7	8	9	10
2	969	943	921	902	886	874	864	858	854	852
5,12	944	897	859	828	803	784	769	758	752	750
100	779	650	566	507	464	434	412	397	389	386
1000	573	417	335	286	253	230	215	205	199	197

Непосредственными вычислениями при $N = 20$ сходимость изложенного алгоритма была установлена при $C \leq 5,12$. Чтобы значения u_n на двух последовательных итерациях отличались меньше чем на 10^{-6} , при $C = 5,12$ потребовалось свыше 1000 итераций (см. табл. 8.1).

Обеспечить сходимость последовательных приближений при больших значениях параметра C можно за счет некоторой модификации изложенного алгоритма. Для этого (8.122) запишем в виде $\frac{d^2 u}{d\xi^2} + C(1 - u^4) = C$, или

$$\frac{d^2 u(\xi)}{d\xi^2} + C(1 - u)(1 + u + u^2 + u^3) = C.$$

Если при выполнении k -й итерации множитель $1 + u + u^2 + u^3$ вычислять по узловым значениям $u_n^{(k-1)}$, найденным на предыдущей итерации, то вместо (8.115), полагая $h = 1/N$, получим

$$-u_{n-1}^{(k)} + (2 + b_n^{(k-1)})u_n^{(k)} - u_{n+1}^{(k)} = b_n^{(k-1)} - \frac{C}{N^2}, \quad (8.123)$$

$n = \overline{1, N-1}$, где

$$b_n^{(k-1)} = \frac{C}{N^2} (1 + u_n^{k-1} + (u_n^{(k-1)})^2 + (u_n^{(k-1)})^3).$$

Решение системы (8.123) методом прогонки при $N = 20$ для значений $C = 2$ и $C = 5,12$ привело к указанным в табл. 8.1 результатам, но за 6 и 8 итераций соответственно. Там же приведены значения $u_n \cdot 10^3$, $n = \overline{1, N/2}$, при $C = 100$ и $C = 1000$. Положение, когда значения на двух последовательных итерациях отличаются менее чем на 10^{-6} , было достигнуто за 44 и 205 итераций соответственно. #

Рассмотрим модификацию метода прогонки для решения СЛАУ (8.27), (8.28) с *пятидиагональной матрицей* (см. пример 8.2). Такая матрица позволяет построить алгоритм прогонки, положив

$$u_n = \varepsilon_n u_{n+1} - \xi_n u_{n+2} + \eta_n. \quad (8.124)$$

Формулы для коэффициентов ε_n , ξ_n и η_n найдем, выразив при помощи (8.124) u_{n-1} и u_{n-2} через u_n и u_{n+1} и подставив их в (8.28). В итоге получим

$$\begin{cases} \varepsilon_n = \frac{d_n + (a_n \varepsilon_{n-2} - b_n) \xi_{n-1}}{\tilde{\Delta}_n}, & \xi_n = \frac{e_n}{\tilde{\Delta}_n}, \\ \eta_n = \frac{f_n - a_n \eta_{n-2} - (a_n \varepsilon_{n-2} - b_n) \eta_{n-1}}{\tilde{\Delta}_n}, \\ \tilde{\Delta}_n = c_n - a_n \xi_{n-2} + (a_n \varepsilon_{n-2} - b_n) \varepsilon_{n-1}. \end{cases} \quad (8.125)$$

Если считать, что $a_0 = a_1 = b_0 = 0$, то из (8.125) следует

$$\begin{aligned} \varepsilon_0 &= \frac{d_0}{c_0}, & \xi_0 &= \frac{e_0}{c_0}, & \eta_0 &= \frac{f_0}{c_0}, & \varepsilon_1 &= \frac{d_1 - b_1 \xi_0}{\tilde{\Delta}_1}, \\ \xi_1 &= \frac{e_1}{\tilde{\Delta}_1}, & \eta_1 &= \frac{f_1 + b_1 \eta_0}{\tilde{\Delta}_1}, & \tilde{\Delta}_1 &= c_1 - b_1 A_0 = \frac{c_0 c_1 - b_1 d_0}{c_0}. \end{aligned}$$

Далее для $n = \overline{2, N}$ используем непосредственно (8.125), учитывая, что $\epsilon_{N-1} = \epsilon_N = d_N = 0$. Тогда при $n = N - 1$ получим $\xi_{N-1} = 0$, а при $n = N$ имеем $\epsilon_N = \xi_N = 0$ и

$$u_N = \eta_N = \frac{f_N - a_N \eta_{N-1} - (a_N \epsilon_{N-2} - b_N) \eta_{N-1}}{c_N - a_N \xi_{N-2} + (a_N \epsilon_{N-2} - b_N) \epsilon_{N-1}}. \quad (8.126)$$

После вычисления u_N из (8.124) при $n = N - 1$ с учетом $\xi_{N-1} = 0$ находим $u_{N-1} = \epsilon_{N-1} u_n + \eta_{N-1}$ и затем по этой же формуле для $n = N - 2, N - 3, \dots, 0$ последовательно вычисляем все остальные узловые значения u_n . Для реализации описанного алгоритма необходимо выполнить $19N - 10$ арифметических операций.

Можно показать*, что разностная схема (8.27), (8.28) в сочетании с алгоритмом (8.124)–(8.126) является корректной, (однозначно разрешима и гарантирована от накопления вычислительных погрешностей), если отличны от нуля все выписанные в левых частях (8.27) и (8.28) коэффициенты и выполнены условия частичного диагонального преобладания приведенной в примере 8.2 пятидиагональной матрицы, т.е.

$$\begin{aligned} |c_0| &\geq |d_0| + |e_0|, \quad |c_1| \geq |b_1| + |d_1| + |e_1|, \\ |c_N| &\geq |a_N| + |b_N|, \quad |c_{N-1}| \geq |a_{N-1}| + |b_{N-1}| + |d_{N-1}|, \\ |c_n| &\geq |a_n| + |b_n| + |d_n| + |e_n|, \quad n = \overline{2, N-2}, \end{aligned}$$

и хотя бы в одном из этих условий достигается строгое неравенство.

Подобно обычному методу прогонки (для решения СЛАУ с трехдиагональной матрицей) и в случае пятидиагональной матрицы также можно построить алгоритмы встречной, поточковой, циклической и немонотонной прогонок, а также прогонки для случая задания дополнительных условий интегрального типа. Возможно применение алгоритма (8.124)–(8.126) для решения последовательными приближениями нелинейных ОДУ четвертого порядка.

*См.: Самарский А.А., Николаев Е.С.

Вопросы и задачи

8.1. Показать, что погрешность аппроксимации в (7.8) имеет первый порядок, а в (7.9) — второй порядок.

8.2. Методом прогонки найти установившееся распределение

$$\Theta\left(\frac{x}{h}\right) = \frac{T(x/h) - T_0}{T_h - T_0}$$

безразмерной температуры по толщине стенки (см. задачу 6.1) при $R_1 = -2R_2$ и $h = 0,2R_2$. Используя разностную схему (8.71) при $\eta = 0, 1$ и $1/2$, решить нестационарную часть этой задачи, приняв в качестве безразмерного времени $\tau = \frac{\tilde{a}t}{R_2^2}$.

8.3. Методом прогонки найти установившееся распределение

$$\Theta\left(\frac{r}{R_2}\right) = \lambda \frac{T(r/R_2) - T_0}{I_V^{(q)} R_2^2}$$

безразмерной температуры в стенке трубы (см. задачу 6.3) при $R_1 = \frac{1}{2} R_2$.

8.4. Решить задачу 6.2 методом прогонки для случая пятидиагональной матрицы.

8.5. Сформулировать теорему Гершгорина.

8.6. Решить задачу 6.10 методом стрельбы.

8.7. Показать, что условие (8.66) выполняется при любых значениях Δt_k , если трехдиагональная матрица A порядка N обладает диагональным преобладанием и ее диагональные элементы $b_n > 0$, $n = \overline{1, N}$.

8.8. Построить шаблоны разностных схем (8.75)–(8.77) и (8.81).

9. МНОГОМЕРНЫЕ ЗАДАЧИ

9.1. Особенности решения многомерных задач

Изложенные в предыдущей главе (см. 8.3) подходы к решению одномерной нестационарной *краевой задачи* применимы и к многомерным задачам. Если время протекания физического процесса, описываемого многомерной нестационарной задачей, разбить на последовательность интервалов Δt_k , $k \in \mathbb{N}$, и провести *аппроксимацию производных* искомых функций по времени t , то в соответствии с *методом прямых* можно перейти к многомерной стационарной краевой задаче относительно распределений этих функций в момент t_k в конце k -го интервала Δt_k . Эту задачу можно решить приближенными аналитическими методами (см. 6), а для численного решения применимы *метод конечных разностей* (МКР), *метод конечных элементов* и *метод граничных элементов*.

Можно также в области решения многомерной нестационарной задачи ввести *пространственную сетку* и на этой сетке аппроксимировать производные искомых функций по пространственным координатам. В результате получим систему обыкновенных дифференциальных уравнений (ОДУ) относительно изменяющихся во времени *узловых значений* этих функций. Такая система в сочетании с заданными начальными условиями составит математическую формулировку задачи Коши, для численного решения которой можно использовать один из вариантов *метода Рунге — Кутты*.

Как и в случае одномерных нестационарных задач, третий подход к решению многомерных задач объединяет первые два и связан с переходом к дискретной *математической модели* рассматриваемого физического процесса как во времени, так и в пространстве. Эта модель на каждом k -м интервале приведет

к системе конечных уравнений относительно узловых значений искомых функций в момент времени t_k в конце этого интервала. Если исходная краевая задача является линейной, то конечные уравнения будут также линейными.

9.2. Двумерная и трехмерная задачи теплопроводности

Третий подход к решению многомерных задач (см. 9.1) рассмотрим сначала на достаточно простом примере двумерной нестационарной задачи теплопроводности в твердом теле, описываемой уравнением

$$\frac{\partial T(t, M)}{\partial t} = \tilde{a} \left(\frac{\partial^2 T(t, M)}{\partial x_1^2} + \frac{\partial^2 T(t, M)}{\partial x_2^2} \right), \quad t > 0, \quad M \in F, \quad (9.1)$$

где $\tilde{a}$ — коэффициент температуропроводности, а температура $T(t, M)$ зависит от времени t и декартовых координат x_1, x_2 точки $M \in F$ двумерной области F решения задачи. Граничные и начальные условия примем в виде

$$T(t, P) = f(t, P), \quad P \in \Gamma; \quad T(0, M) = T^0(M), \quad M \in F, \quad (9.2)$$

где Γ — граница области F (рис. 9.1).

Введем в F двумерную пространственную сетку F_h с шагами $h_i, i = 1, 2$, постоянными вдоль каждой из координатных осей (на рис. 9.2 принято $h_1 = h_2 = h = \text{const}$). При этом в узлы, вообще говоря, не принадлежащие Γ , но расположенные наиболее близко к границе и составляющие множество Γ_h , перенесем из ближайшей к каждому узлу точки $P \in \Gamma$ заданные граничные значения (9.2) температуры. Отметим, что возникающая при этом погрешность является основной причиной, ограничивающей применение МКР к решению задач в многомерных областях произвольной конфигурации.

Равенством

$$\Lambda_{ii} T_n = \tilde{a} \frac{T_{n-1} - 2T_n + T_{n+1}}{h_i^2}, \quad i = 1, 2, \quad (9.3)$$

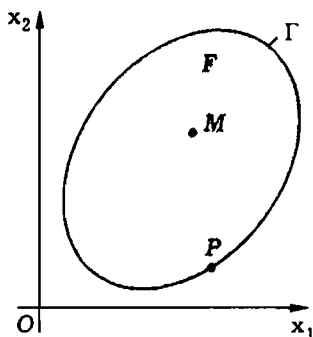


Рис. 9.1

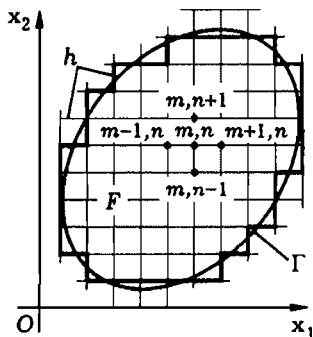


Рис. 9.2

где n — номер внутреннего узла двумерной сетки F_h , отсчитываемый вдоль координатной оси Ox_1 , введем **разностный оператор** Λ_{ii} . Тогда для любого внутреннего узла $M_n \in F_h$ этой сетки, имеющего температуру $T_*(t)$, с учетом (7.3) запишем

$$\left. \frac{\partial^2 T(t, M)}{\partial x_i^2} \right|_{M=M_n} = \Lambda_{ii} T_n(t) + O(h_i^2).$$

Опуская нижний индекс в обозначении узлового значения температуры и аппроксимируя в (9.1) производную по времени на k -м интервале, для каждого внутреннего узла сетки F_h получаем

$$\frac{T^k - T^{k-1}}{\Delta t_k} = \sum_{i=1,2} \Lambda_{ii} (\eta_i T^k + (1 - \eta_i) T^{k-1}), \quad \eta_i \in [0, 1]. \quad (9.4)$$

В более подробной записи (9.4) имеет вид

$$\begin{aligned} \frac{T_{mn}^k - T_{mn}^{k-1}}{\Delta t_k} = & \eta_1 \frac{T_{m-1,n}^k - 2T_{mn}^k + T_{m+1,n}^k}{h_1^2} + \\ & + (1 - \eta_1) \frac{T_{m-1,n}^{k-1} - 2T_{mn}^{k-1} + T_{m+1,n}^{k-1}}{h_1^2} + \\ & + \eta_2 \frac{T_{m,n-1}^k - 2T_{mn}^k + T_{m,n+1}^k}{h_2^2} + (1 - \eta_2) \frac{T_{m,n-1}^{k-1} - 2T_{mn}^{k-1} + T_{m,n+1}^{k-1}}{h_2^2}. \end{aligned}$$

Она основана на переходе от разностной схемы (9.4) к разностной схеме

$$\begin{cases} \frac{T^{k-1/2} - T^{k-1}}{\Delta t_k} = \eta_1 \Lambda_{11} T^{k-1/2} + (1 - \eta_2) \Lambda_{22} T^{k-1}, \\ \frac{T^k - T^{k-1/2}}{\Delta t_k} = (1 - \eta_1) \Lambda_{11} T^{k-1/2} + \eta_2 \Lambda_{22} T^k. \end{cases} \quad (9.5)$$

Каждое из двух соотношений этой схемы представляет совокупность одномерных *нелинейных двухслойных разностных схем* по осям x_1 и x_2 соответственно, причем первое из них содержит неизвестные узловые значения $T^{k-1/2}$ температуры в момент времени $t_{k-1/2}$ в середине k -го интервала, а второе — значения T^k в тех же узлах в момент t_k в конце этого интервала. Отдельно взятая одномерная схема, следующая из первого соотношения (9.5), включает значения $T^{k-1/2}$ в цепочке узлов при $x_2 = \text{const}$, а следующая из второго соотношения — искомые значения T^k в цепочке узлов при $x_1 = \text{const}$. Все эти узловые значения могут быть найдены при помощи обычного *метода прогонки*.

Выясним условия эквивалентности (9.4) и (9.5). Для этого исключим $T^{k-1/2}$ из (9.5), введя разностные операторы

$$B_i = I - \eta_i \Lambda_{ii} \Delta t_k, \quad C_i = I + (1 - \eta_i) \Lambda_{ii} \Delta t_k, \quad i = 1, 2, \quad (9.6)$$

где I — *тождественный оператор*. Тогда (9.5) примет вид

$$B_1 T^{k-1/2} = C_2 T^{k-1}, \quad B_2 T^k = C_1 T^{k-1/2},$$

откуда получаем $B_1 B_2 T^k = C_1 C_2 T^{k-1}$, т.е. с учетом (9.6) приходим к разностной схеме

$$\begin{aligned} (I + \eta_1 \eta_2 \Lambda_{11} \Lambda_{22} (\Delta t_k)^2) \frac{T^k - T^{k-1}}{\Delta t_k} = \\ = (1 - \eta_1 - \eta_2) \Lambda_{11} \Lambda_{22} T^{k-1} \Delta t_k + \\ + \sum_{i=1,2} \Lambda_{ii} (\eta_i T^k + (1 - \eta_i) T^{k-1}). \end{aligned} \quad (9.7)$$

Сравнивая уравнение (9.7) с (9.4), заключаем, что при $\eta_1 = \eta_2 = 1/2$ разностная схема (9.7), а значит, и (9.5) аппроксимируют дифференциальное уравнение (9.1) с погрешностью $O((\Delta t_k)^2, h_1^2 + h_2^2)$. Отметим, что несмотря на совпадение порядков погрешности разностных схем (9.4) и (9.5) результаты расчетов по ним будут различны, что объясняется наличием в (9.7) дополнительных слагаемых.

В частном случае $\eta_1 = \eta_2 = 1$ каждое из соотношений (9.5) приводит к *неявной двухслойной разностной схеме*, обеспечивающей *вычислительную устойчивость алгоритма* метода прогонки при любых значениях $\Delta t_k > 0$. В этом случае (9.5) называют *локально-одномерной разностной схемой*. Ясно, что переход в (9.5) при выборе $\eta_1 = \eta_2 = 0$ к явной схеме не имеет смысла, поскольку такой выбор приводит к *явной двухслойной разностной схеме* непосредственно из (9.4). Можно показать, что при этом алгоритм вычислений устойчив при выполнении условия

$$\tilde{a} \Delta t_k \leq \frac{h_1^2 h_2^2}{h_1^2 + h_2^2}.$$

Рассмотренные свойства разностной схемы (9.5) можно распространить на случаи переменных (в том числе разрывных) коэффициентов в уравнении теплопроводности, наличия в теле источников энергоснабжения и двумерной сетки с переменным шагом в направлении каждой из координатных осей, а также — на случаи двумерных задач в цилиндрической системе координат и осесимметричных задач в сферических координатах. Однако идею продольно-поперечной прогонки не удастся в общем виде перенести на трехмерные задачи нестационарной теплопроводности. Но для задачи в произвольной области с размерностью d , описываемой дифференциальным уравнением

$$\frac{\partial T}{\partial t} = \tilde{a} \sum_{i=1}^d \frac{\partial^2 T}{\partial x_i^2}, \quad (9.8)$$

можно построить локально-одномерную разностную схему

$$\frac{T^{k-1+i/d} - T^{k-1+(i-1)/d}}{\Delta t_k} = \Lambda_{ii}(\eta_i T^{k-1+i/d} + (1 - \eta_i) T^{k-1+(i-1)/d}), \quad (9.9)$$

которая при каждом значении $i = \overline{1, d}$ позволяет использовать обычный метод прогонки в отличие от разностной схемы

$$\frac{T^k - T^{k-1}}{\Delta t_k} = \sum_{i=1}^d \Lambda_{ii}(\eta_i T^k + (1 - \eta_i) T^{k-1}), \quad (9.10)$$

обобщающей (9.4) и аппроксимирующей (9.8) непосредственно.

При $\eta_i = 1/2$ для всех значений $i = \overline{1, d}$ разностная схема (9.10) аппроксимирует (9.8) со вторым порядком погрешности как по Δt_k , так и по шагам h_i пространственной сетки, а в остальных случаях порядок погрешности по Δt_k уменьшается до единицы. Выясним порядок погрешности схемы (9.9). Для этого исключим из (9.9) промежуточные значения $T^{k-1+i/d}$, $i = \overline{1, d-1}$, и с учетом (9.6) запишем

$$\left(\prod_{i=1}^d B_i\right) T^k = \left(\prod_{i=1}^d C_i\right) T^{k-1}.$$

Отсюда, например, для случая $d = 3$ получим

$$\begin{aligned} \frac{T^k - T^{k-1}}{\Delta t_k} = & \sum_{i=1}^3 \Lambda_{ii}(\eta_i T^k + (1 - \eta_i) T^{k-1}) + \\ & + \Lambda_{11} \Lambda_{22} \Lambda_{33} (\eta_1 \eta_2 \eta_3 T^k + \chi_1 \chi_2 \chi_3 T^{k-1}) (\Delta t_k)^2 - \\ & - (\eta_1 \eta_2 \Lambda_{11} \Lambda_{22} + \eta_2 \eta_3 \Lambda_{22} \Lambda_{33} + \eta_3 \eta_1 \Lambda_{33} \Lambda_{11}) T^k (\Delta t_k)^2 + \\ & + (\chi_1 \chi_2 \Lambda_{11} \Lambda_{22} + \chi_2 \chi_3 \Lambda_{22} \Lambda_{33} + \chi_3 \chi_1 \Lambda_{33} \Lambda_{11}) T^{k-1} \Delta t_k, \end{aligned}$$

где $\chi_i = 1 - \eta_i$, $i = \overline{1, 3}$. В случае $\eta_1 = \eta_2 = \eta_3 = 1/2$ имеем

$$\begin{aligned} & \left(I + (\Lambda_{11}\Lambda_{22} + \Lambda_{22}\Lambda_{33} + \Lambda_{33}\Lambda_{11}) \left(\frac{\Delta t_k}{2} \right)^2 \right) \frac{T^k - T^{k-1}}{\Delta t_k} = \\ & = \left(\Lambda_{11} + \Lambda_{22} + \Lambda_{33} + \Lambda_{11}\Lambda_{22}\Lambda_{33} \left(\frac{\Delta t_k}{2} \right)^2 \right) \frac{T^k + T^{k-1}}{2} \end{aligned}$$

и, сравнивая с (9.10) при $d = 3$ и $\eta_i = 1/2$, $i = \overline{1, 3}$, приходим к выводу, что (9.9) аппроксимирует (9.8) с погрешностью $O((\Delta t_k)^2, h_1^2 + h_2^2 + h_3^2)$. Иное сочетание значений η_i снижает порядок погрешности по Δt_k до 1.

При $d = 2$ и $\eta_i = 1/2$, $i = 1, 2$, разностная схема (9.9) по порядку погрешности эквивалентна (9.5), но приводит к более простой структуре алгоритма, поскольку при прогонке вдоль каждой цепочки узлов сетки, соответствующих $x_i = \text{const}$, используются значения температуры только в этих узлах. При $d = 3$ таким же преимуществом обладает *разностная схема типа предиктор-корректор*

$$\begin{aligned} \frac{T^{k-5/6} - T^{k-1}}{\Delta t_k/2} &= \Lambda_{11}T^{k-5/6}, \quad \frac{T^{k-2/3} - T^{k-5/6}}{\Delta t_k/2} = \Lambda_{22}T^{k-2/3}, \\ \frac{T^{k-1/2} - T^{k-2/3}}{\Delta t_k/2} &= \Lambda_{33}T^{k-2/3}, \\ \frac{T^k - T^{k-1}}{\Delta t_k} &= (\Lambda_{11} + \Lambda_{22} + \Lambda_{33})T^{k-1/2}, \end{aligned}$$

которая сочетает локально-одномерную схему („предиктор“) для предсказания узловых значений в момент времени $t_{k-1/2}$ в середине интервала Δt_k и явную схему („корректор“) для завершения расчета на этом интервале. Эта схема также имеет второй порядок погрешности как по Δt_k , так и по шагам пространственной сетки, а алгоритм вычислений по каждому из соотношений этой схемы устойчив при любых значениях Δt_k .

Для решения многомерных нелинейных задач теплопроводности наряду с явной схемой, которая следует из (9.10) при $\eta_i = 0$, $i = \overline{1, d}$, и имеет ограничение на выбор значения Δt_k , можно использовать не имеющую такого ограничения схему, обобщающую *явную трехслойную разностную схему* вида (8.77). Для многомерных задач эту схему можно представить в виде

$$\frac{T^k - T^{k-2}}{2\tau} = \frac{1}{3}\Lambda^{k-1}(T^k + T^{k-1} + T^{k-2}) + f^{k-1}, \quad (9.11)$$

где Λ^{k-1} — разностный оператор на ортогональной многомерной сетке с переменными (в общем случае) шагами h_i вдоль каждой из координатных осей, аппроксимирующий дифференциальный оператор $\frac{1}{c}\nabla(\lambda\nabla T)$ в нелинейном *уравнении теплопроводности* вида (2.53). Зависящие от температуры узловые значения f^{k-1} источников члена $\frac{I_V^{(q)}}{c}$, входящего в (2.53), и коэффициенты c и λ вычисляются по известным при расчете на k -м интервале узловым значениям T^{k-1} в момент времени t_{k-1} в середине удвоенного интервала $2\tau = t_k - t_{k-2}$.

Из (9.11) следует разностная схема

$$\left(I - \frac{2\tau}{3}\Lambda^{k-1}\right)T^k = T^{k-2} + 2\tau\left(\Lambda^{k-1}\frac{T^{k-1} + T^{k-2}}{3} + f^{k-1}\right). \quad (9.12)$$

В d -мерной декартовой прямоугольной системе координат

$$\Lambda^{k-1} = \sum_{i=1}^d \Lambda_{ii}^{k-1}, \quad (9.13)$$

где Λ_{ii}^{k-1} — разностный оператор, аппроксимирующий дифференциальный оператор $\frac{1}{c}\frac{\partial}{\partial x_i}\left(\lambda\frac{\partial}{\partial x_i}\right)$, входящий в правую часть (2.53). Отметим, что (9.13) верно и для анизотропного тела

при совпадении *главных осей тензора теплопроводности* с координатными осями (см. замечание 2.1).

Так как (9.11) аппроксимирует (2.53) по τ с первым порядком погрешности, то (9.12) имеет по τ погрешность $O(\tau^2)$. Поскольку с учетом (9.13)

$$I - \frac{2\tau}{3} \Lambda^{k-1} = I - \frac{2\tau}{3} \sum_{i=1}^d \Lambda_{ii}^{k-1} = \prod_{i=1}^d \left(I - \frac{2\tau}{3} \Lambda_{ii}^{k-1} \right) + O(\tau^2),$$

то порядок погрешности по τ не уменьшится, если от (9.12) перейти к разностной схеме

$$\left(\prod_{i=1}^d \left(I - \frac{2\tau}{3} \Lambda_{ii}^{k-1} \right) \right) T^k = T^{k-2} + \frac{2\tau}{3} \Lambda^{k-1} (T^{k-1} + T^{k-2}) + 2f^{k-1}\tau,$$

которая эквивалентна схеме

$$\begin{aligned} \left(I - \frac{2\tau}{3} \Lambda_{11}^{k-1} \right) T^{k-1+1/d} &= T^{k-2} + \frac{2\tau}{3} \Lambda^{k-1} (T^{k-1} + T^{k-2}) + 2f^{k-1}\tau, \\ \left(I - \frac{2\tau}{3} \Lambda_{ii}^{k-1} \right) T^{k-1+1/d} &= T^{k-1+(i-1)/d}, \quad i = \overline{2, d}. \end{aligned}$$

Первое соотношение явно разрешено относительно узловых значений $T^{k-1+1/d}$, а второе (при $d=2$) или два других (при $d=3$) приводят к СЛАУ с трехдиагональными матрицами, решаемым обычным методом прогонки.

9.3. Различные многомерные задачи

Изложенные выше (см. 9.2) способы решения многомерных задач нестационарной теплопроводности можно применять в более широком классе нестационарных краевых задач математической физики, описываемых дифференциальными уравнениями параболического типа или системами таких уравнений. Эти же способы применимы и для решения стационарных

(статических) задач, описываемых дифференциальными уравнениями эллиптического типа. Искомое решение таких задач можно рассматривать как предельное, установившееся состояние в условной дискретной системе, в которой происходит нестационарный физический процесс при заданных в стационарной задаче постоянных во времени граничных условиях. Если стационарная задача имеет единственное решение, то при произвольно выбранном начальном условии решение нестационарной задачи для условной дискретной системы в пределе приводит к искомому установившемуся состоянию. Такой прием получения решения стационарной задачи называют **методом установления**.

В случае неединственности решения рассматриваемой стационарной задачи установившееся состояние будет связано с задаваемым начальным условием. Тогда начальное условие приобретает смысл нулевого приближения, от близости которого к искомому решению зависит объем вычислений при использовании метода установления.

Помимо этого метода для решения многомерных линейных стационарных задач математической физики при помощи МКР можно указать еще ряд способов, которые приводят к СЛАУ, вытекающей из дискретной *математической модели* рассматриваемого физического процесса. К ним относятся вычислительные методы линейной алгебры, связанные с последовательным исключением неизвестных или *мультипликативным разложением матрицы* СЛАУ, а также большая группа итерационных методов решения СЛАУ [IV]. В случае нелинейных задач дискретная модель приводит к системе конечных уравнений, решаемой также итерационными методами [V].

Некоторые из итерационных методов* в определенном смысле можно трактовать и как варианты метода установления, поскольку получаемые в процессе последовательных приближений к искомому решению стационарной задачи промежуточные

*См. Самарский А.А., Николаев Е.С., а также: Трауб Дж.

состояния соответствуют условному нестационарному процессу в дискретной системе. Если решаемая задача имеет вариационную формулировку, включающую *функционал* с известными экстремальными свойствами, то используемый итерационный метод можно соотнести с *методом локальных вариаций* и контролировать сходимость процесса последовательных приближений по изменению значения функционала от итерации к итерации.

Все эти способы направлены на вычисление значений искомой функции во всех узлах сетки, соответствующей области решения задачи. Если же цель решения задачи состоит в нахождении не всех узловых значений, а лишь в одном или нескольких характерных узлах, то эффективным может оказаться вероятностный подход, который рассмотрим на простом примере.

Пример 9.1. Пусть в двумерной области F , ограниченной замкнутой кривой Γ и представленной сеткой F_h с шагом $h = \text{const}$ (сетка имеет квадратные ячейки), необходимо найти в некотором внутреннем узле $M_m^l \in F_h$ с номерами l и m значение $u(M_m^l)$ искомой функции $u(M)$, описываемой уравнением Лапласа

$$\frac{\partial^2 u(M)}{\partial x_1^2} + \frac{\partial^2 u(M)}{\partial x_2^2} = 0, \quad M \in F, \quad (9.14)$$

с граничными условиями

$$u(R) = \tilde{g}(R), \quad R \in \Gamma. \quad (9.15)$$

Если для аппроксимации вторых частных производных использовать (7.3), то задаче Дирихле (9.14), (9.15) будет соответствовать разностная схема

$$\begin{cases} \frac{u(M_{q-1}^p) - 2u(M_q^p) + u(M_{q+1}^p)}{h^2} + \\ + \frac{u(M_q^{p-1}) - 2u(M_q^p) + u(M_q^{p+1})}{h^2} = 0, \\ u(R_s^r) = g(R_s^r), \quad R_s^r \in \Gamma_h, \end{cases} \quad (9.16)$$

где $M_q^p \in F_h$ — любой внутренний узел сетки F_h , а Γ_h — множество узлов этой сетки, вообще говоря, не принадлежащих границе Γ , но наименее удаленных от нее. В каждый такой узел R_s^r перенесено заданное в (9.15) граничное значение $g(R_s^r) = \tilde{g}(R)$ из точки $R \in \Gamma$, наиболее близкой к $R_s^r \in \Gamma_h$.

Поскольку разностная схема (9.16) линейная, то справедливо равенство

$$u(M_q^p) = \sum_{R_s^r \in \Gamma_h} P(M_q^p, R_s^r) g(R_s^r), \quad M_q^p \in F_h,$$

где $P(M_q^p, R_s^r) g(R_s^r)$ — коэффициент влияния граничного узла $R_s^r \in \Gamma$ на значение искомой функции во внутреннем узле M_q^p . По физическому смыслу задачи Дирихле в частном случае $\tilde{g}(R) = g_0 = \text{const}$, $R \in \Gamma$, должно быть $u(M) = g_0$, $M \in F$. В таком случае $g(R_s^r) = g_0$, $R_s^r \in \Gamma_h$, и $u(M_q^p) = g_0$, $M_q^p \in F_h$. Отсюда следует, что

$$\sum_{R_s^r \in \Gamma_h} P(M_q^p, R_s^r) = 1.$$

Коэффициент $P(M_q^p, R_s^r)$ можно истолковать как *вероятность события*, состоящего в том, что частица, выпущенная в узле $M_q^p \in F_h$, после случайного блуждания по сетке F_h попадет в граничный узел $R_s^r \in \Gamma_h$. Записав первое уравнение в (9.16) в виде

$$u(M_q^p) = 0,25u(M_{q-1}^p) + 0,25u(M_{q+1}^p) + \\ + 0,25u(M_q^{p-1}) + 0,25u(M_q^{p+1}),$$

коэффициенты в правой части интерпретируем как вероятность перехода блуждающей частицы из произвольного внутреннего узла $M_q^p \in F_h$ в конкретный соседний с ним узел. Ясно, что для квадратной сетки эти вероятности одинаковы и равны 0,25. Направление перехода можно выбрать в зависимости

от того, в какой промежуток попадет число z , выбранное как *случайная величина* в полуинтервале $(0, 1]$ (для генерирования случайных чисел существуют специальные алгоритмы). Например, если $z \in (0, 0,25]$, то блуждающая частица переходит из узла M_q^p в узел M_{q-1}^p , если $z \in (0,25, 0,5]$, то — в узел M_{q+1}^p и т.п.

Блуждающая частица, выпущенная из интересующего нас узла M_m^l , после серии описанных переходов попадет в граничный узел $R_s^r \in \Gamma_h$ с известным значением $u_1 = g(R_s^r)$. Затем из того же узла M_m^l выпускают вторую частицу и фиксируют значение u_2 в том граничном узле, куда она попадет после случайного блуждания. Эту процедуру повторяют для N частиц, и в итоге искомое значение функции в узле $M_m^l \in F_h$ можно представить как *математическое ожидание* случайных значений u_n . $n = \overline{1, N}$, в достигнутых граничных узлах, т.е.

$$u(M_m^l) = \frac{1}{N} \sum_{n=1}^N u_n, \quad (9.17)$$

причем относительная погрешность (9.17) составит $1/\sqrt{N}$. #

Изложенный подход к решению стационарной задачи составляет существо *метода статистических испытаний*, называемого также методом Монте-Карло. Ясно, что для обеспечения приемлемой точности решения значение N в (9.17) должно быть достаточно велико, что связано со значительным временем счета на ЭВМ. Поэтому нахождение этим методом значений искомой функции во многих узлах может оказаться нереальным. Однако простота алгоритма и возможность вычислить искомые значения в одном или нескольких узлах без решения всей задачи в целом привлекают внимание к этому методу и стимулируют его дальнейшую разработку. Отметим, что вероятностную интерпретацию можно дать и разностным схемам, соответствующим нестационарным задачам.

Перейдем к рассмотрению способов решения при помощи МКР многомерных дифференциальных уравнений гиперболического типа. *Волновое уравнение*

$$\frac{\partial^2 u(t, M)}{\partial t^2} = a^2 \left(\frac{\partial^2 u(t, M)}{\partial x_1^2} + \frac{\partial^2 u(t, M)}{\partial x_2^2} \right), \quad t > 0, \quad M \in F, \quad (9.18)$$

описывает распространение звуковых волн в некоторой двумерной области F . Здесь a — скорость звука в среде, а функция $u(t, M)$ может соответствовать возмущенному распределению плотности или давления среды. Это же уравнение в сочетании с заданными на границе Γ области F значениями $u(t, P) \equiv 0$ описывает поперечные колебания закрепленной по контуру мембраны. В этом случае функция $u(t, M)$ в момент времени t характеризует отклонение точки $M \in F$ от положения ее равновесия в плоскости мембраны.

Используем для аппроксимации (9.18) аналогичную (8.81) *трехслойную симметричную разностную схему*

$$\begin{aligned} \frac{u^k - 2u^{k-1} + u^{k-2}}{(a\tau)^2} &= \\ &= \sum_{i=1,2} \Lambda_{ii} (\eta_i (u^k + u^{k-2}) + (1 - 2\eta_i) u^{k-1}), \end{aligned} \quad (9.19)$$

где τ — интервал времени, а Λ_{ii} соответствует обозначению (9.3). Эта схема при $\eta_1 = \eta_2 = 0$ становится явной, имеет погрешность $O(\tau^2, h_1^2 + h_2^2)$ и ограничение $(a\tau)^2 \leq \frac{h_1^2 h_2^2}{h_1^2 + h_2^2}$ на выбор значения τ , а при $\eta_i \in (0, 1]$ — неявной, так что для нахождения узловых значений u^k в конце k -го интервала времени приходится применять продольно-поперечную прогонку.

В случае $\eta_1 = \eta_2 = 1/2$ (9.19) аппроксимирует (9.18) также с погрешностью $O(\tau^2, h_1^2 + h_2^2)$. Поэтому следующее из (9.19) равенство

$$\left(I - (\Lambda_{11} + \Lambda_{22}) \frac{(a\tau)^2}{2} \right) \frac{u^k + u^{k-2}}{2} = u^{k-1} \quad (9.20)$$

имеет по τ четвертый порядок погрешности. Поскольку

$$\left(I - (\Lambda_{11} + \Lambda_{22}) \frac{(a\tau)^2}{2}\right) = \left(I - \Lambda_{11} \frac{(a\tau)^2}{2}\right) \left(I - \Lambda_{22} \frac{(a\tau)^2}{2}\right) + O(\tau^4),$$

то порядок погрешности по τ не уменьшится, если от (9.20) перейти к локально-одномерной разностной схеме

$$\begin{cases} \left(I - \Lambda_{11} \frac{(a\tau)^2}{2}\right) u^{k-1/2} = u^{k-1}, \\ \left(I - \Lambda_{22} \frac{(a\tau)^2}{2}\right) \frac{u^k + u^{k-2}}{2} = u^{k-1/2}. \end{cases} \quad (9.21)$$

Первое из этих соотношений позволяет найти узловые значения $u^{k-1/2}$ в момент времени $t_{k-1/2} = t_{k-1} + \frac{\tau}{2}$ обычным методом прогонки вдоль цепочек узлов при $x_2 = \text{const}$, а затем также прогонкой, но вдоль цепочек узлов при $x_1 = \text{const}$ решают СЛАУ с трехдиагональными матрицами, следующие из второго соотношения, и находят узловые значения u^k в конце k -го интервала времени.

Разностную схему (9.21) нетрудно распространить на случай трехмерной области V решения задачи. Функцию $u(t, M)$ в (9.18) можно рассматривать как перемещение или скорость среды, т.е. в многомерной задаче эта функция становится векторной, что требует соответствующего обобщения разностных схем (9.19)–(9.21). В качестве примера обратимся к двумерной динамической задаче теории упругости в перемещениях, описываемой системой дифференциальных уравнений со смешанными производными

$$\begin{cases} \rho \frac{\partial^2 u_1}{\partial t^2} = b_1 + (\tilde{\lambda} + 2\tilde{\mu}) \frac{\partial^2 u_1}{\partial x_1^2} + \tilde{\mu} \frac{\partial^2 u_1}{\partial x_2^2} + (\tilde{\lambda} + \tilde{\mu}) \frac{\partial^2 u_2}{\partial x_1 \partial x_2}, \\ \rho \frac{\partial^2 u_2}{\partial t^2} = b_2 + (\tilde{\lambda} + 2\tilde{\mu}) \frac{\partial^2 u_2}{\partial x_2^2} + \tilde{\mu} \frac{\partial^2 u_2}{\partial x_1^2} + (\tilde{\lambda} + \tilde{\mu}) \frac{\partial^2 u_1}{\partial x_1 \partial x_2}, \end{cases} \quad (9.22)$$

следующей после линеаризации уравнений Ламе в векторной форме (3.35). Здесь ρ — плотность среды, $\tilde{\lambda}$ и $\tilde{\mu}$ — константы

Ламе, u_i и b_i , $i = 1, 2$, — проекции векторов перемещения и плотности объемных сил на координатные оси Ox_i .

Аппроксимируем (9.22) трехслойной симметричной разностной схемой

$$\begin{aligned} \rho \frac{u_i^k - 2u_i^{k-1} + u_i^{k-2}}{\tau^2} = & b_i^{k-1} + ((\tilde{\lambda} + 2\tilde{\mu})\Delta_i \Delta_{-i} + \tilde{\mu}\Delta_{3-i} \Delta_{i-3}) \times \\ & \times (\eta_i(u_i^k + u_i^{k-2}) + (1 - 2\eta_i)u_i^{k-1}) + \frac{\tilde{\lambda} + \tilde{\mu}}{4}(\Delta_1 + \Delta_{-1})(\Delta_2 + \Delta_{-2}) \times \\ & \times (\eta_i(u_{3-i}^k + u_{3-i}^{k-2}) + (1 - 2\eta_i)u_{3-i}^{k-1}), \quad i = 1, 2, \quad (9.23) \end{aligned}$$

где операторы Δ_i и Δ_{-i} правой (вперед) и левой (назад) конечных разностей заданы равенствами

$$\Delta_i w_n = \frac{w_{n+1} - w_n}{h_i}, \quad \Delta_{-i} w_n = \frac{w_n - w_{n-1}}{h_i}, \quad i = 1, 2,$$

причем нижними индексами у скалярной величины w отмечены номера узлов сетки, равномерно расположенных с шагом h_i вдоль оси Ox_i при условии $x_{3-i} = \text{const}$. Ясно, что

$$(\Delta_i + \Delta_{-i}) \frac{w_n}{2} = \frac{w_{n+1} - w_{n-1}}{2h_i} \quad \text{и} \quad \Delta_i \Delta_{-i} w_n = \frac{w_{n-1} - 2w_n + w_{n+1}}{h_i^2}$$

дают аппроксимации первой и второй производных по x_i со вторым порядком погрешности (см. 7.2).

Разностная схема (9.23) при $\eta_1 = \eta_2 = 0$ явная, а при $\eta_1 = \eta_2 = 1/2$ неявная с погрешностью $O(\tau^2, h_1^2 + h_2^2)$. В последнем случае она имеет вид

$$\begin{aligned} \rho \frac{u_i^k - 2u_i^{k-1} + u_i^{k-2}}{\tau^2} = & \\ = & \frac{\tilde{\lambda} + \tilde{\mu}}{4}(\Delta_1 + \Delta_{-1})(\Delta_2 + \Delta_{-2}) \frac{u_{3-i}^k + u_{3-i}^{k-2}}{2} + b_i^{k-1} + \\ & + ((\tilde{\lambda} + 2\tilde{\mu})\Delta_i \Delta_{-i} + \tilde{\mu}\Delta_{3-i} \Delta_{i-3}) \frac{u_i^k + u_i^{k-2}}{2}, \quad i = 1, 2. \quad (9.24) \end{aligned}$$

Поскольку с учетом (7.3)

$$\begin{aligned}\frac{u_i^k + u_i^{k-2}}{2} &= u_i^{k-1} + \frac{u_i^k - 2u_i^{k-1} + u_i^{k-2}}{2} = \\ &= u_i^{k-1} + \frac{\partial^2 u_i}{\partial t^2} \Big|_{t=t_{k-1}} \tau^2 + O(\tau^4),\end{aligned}$$

то с погрешностью не ниже второго порядка по τ заменим в (9.24) $\frac{1}{2}(u_{3-i}^k + u_{3-i}^{k-2})$ на u_{3-i}^{k-1} и запишем

$$\begin{aligned}\left(I - ((\tilde{\lambda} + 2\tilde{\mu})\Delta_i \Delta_{-i} + \tilde{\mu}\Delta_{3-i} \Delta_{i-3}) \frac{\tau^2}{2\rho}\right)(u_i^k + u_i^{k-2}) &= b_i^{k-1} \frac{\tau^2}{\rho} + \\ + 2u_i^{k-1} + \frac{\tilde{\lambda} + \tilde{\mu}}{4\rho} \tau (\Delta_1 + \Delta_{-1})(\Delta_2 + \Delta_{-2}) u_{3-i}^{k-1}, \quad i = 1, 2. \quad (9.25)\end{aligned}$$

Разностный оператор в левой части (9.25) с погрешностью четвертого порядка заменим произведением операторов, входящих в левые части соотношений

$$\begin{aligned}\left(I - (\tilde{\lambda} + 2\tilde{\mu})\Delta_i \Delta_{-i} \frac{\tau^2}{2\rho}\right) u_i^{k-1/2} &= b_i^{k-1} \frac{\tau^2}{\rho} + 2u_i^{k-1} + \\ + \frac{\tilde{\lambda} + \tilde{\mu}}{4\rho} \tau (\Delta_1 + \Delta_{-1})(\Delta_2 + \Delta_{-2}) u_{3-i}^{k-1}, \quad i = 1, 2, \quad (9.26)\end{aligned}$$

и

$$\left(I - \tilde{\mu}\Delta_{3-i} \Delta_{i-3} \frac{\tau^2}{2\rho}\right)(u_i^k + u_i^{k-2}) = u_i^{k-1/2}, \quad i = 1, 2. \quad (9.27)$$

Эта разностная схема аппроксимирует (9.22) с погрешностью $O(\tau^2, h_1^2 + h_2^2)$ и применима без ограничения на выбор интервала τ . Аналогично схеме (9.21) соотношение (9.26) позволяет вычислить узловые значения $u_i^{k-1/2}$ в момент времени $t_{k-1/2} = t_{k-1} + \frac{\tau}{2}$ обычным методом прогонки вдоль цепочек узлов при $x_{3-i} = \text{const}$, а затем также прогонкой, но вдоль цепочек узлов при $x_i = \text{const}$ решают СЛАУ (9.27) с трехдиагональными ма-

трицами и находят узловые значения u_i^k в конце k -го интервала времени.

Локально-одномерную схему (9.26), (9.27) можно обобщить и на случай трехмерной динамической задачи теории упругости. Отметим, что разностная схема (9.26), (9.27) и ей подобные не применимы к решению статических задач теории упругости методом установления. Дело в том, что система уравнений (9.22) при постоянных во времени граничных условиях и проекциях вектора плотности объемных сил описывает незатухающие гармонические колебания в линейно упругом теле. Но разностную схему метода установления нетрудно построить, если в (9.22) левые части, характеризующие проекции инерционных сил, предварительно заменить на $r \frac{\partial u_1}{\partial t}$ и $r \frac{\partial u_2}{\partial t}$ соответственно, где r — параметр, имеющий смысл коэффициента сопротивления. Модифицированные таким способом дифференциальные уравнения можно аппроксимировать на интервале времени Δt_k разностной схемой*

$$r \frac{u_1^{k-1/2} - u_1^{k-1}}{\Delta t_k} = (\tilde{\lambda} + 2\tilde{\mu})\Delta_1\Delta_{-1}u_1^{k-1/2} + \tilde{\mu}\Delta_2\Delta_{-2}u_1^{k-1} + \\ + \frac{\tilde{\lambda} + \tilde{\mu}}{4}(\Delta_1 + \Delta_{-1})(\Delta_2 + \Delta_{-2})u_2^{k-1} + b_1^{k-1/2},$$

$$r \frac{u_2^{k-1/2} - u_2^{k-1}}{\Delta t_k} = \frac{\tilde{\lambda} + \tilde{\mu}}{4}(\Delta_1 + \Delta_{-1})(\Delta_2 + \Delta_{-2})u_1^{k-1/2} + \\ + \tilde{\mu}\Delta_1\Delta_{-1}u_2^{k-1/2} + (\tilde{\lambda} + 2\tilde{\mu})\Delta_2\Delta_{-2}u_2^{k-1} + b_2^{k-1/2},$$

$$r \frac{u_1^k - u_1^{k-1/2}}{\Delta t_k} = \tilde{\mu}\Delta_2\Delta_{-2}(u_1^k - u_1^{k-1}),$$

$$r \frac{u_2^k - u_2^{k-1/2}}{\Delta t_k} = (\tilde{\lambda} + 2\tilde{\mu})\Delta_2\Delta_{-2}(u_2^k - u_2^{k-1}),$$

*См.: Зарубин В.С., Селиванов В.В..

к которой можно применить метод установления в сочетании с обычным методом прогонки.

Рассмотрим особенности построения разностной схемы для дифференциального уравнения четвертого порядка (3.27)

$$\nabla^2(\nabla^2\psi(M)) = f(M), \quad M \in F, \quad (9.28)$$

где $\psi(M)$ и $f(M)$ - искомая и заданная функции соответственно в двумерной области F , ограниченной контуром Γ . Если $\psi(M)$ имеет физический смысл *функции тока*, то это уравнение описывает установившееся движение несжимаемой вязкой жидкости (см. пример 3.1). В декартовой прямоугольной системе координат Ox_1x_2 *бигармонический оператор* имеет вид

$$\begin{aligned} \nabla^2(\nabla^2\psi) &= \left(\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} \right) \left(\frac{\partial^2\psi}{\partial x_1^2} + \frac{\partial^2\psi}{\partial x_2^2} \right) = \\ &= \frac{\partial^4\psi}{\partial x_1^4} + 2 \frac{\partial^4\psi}{\partial x_1^2 \partial x_2^2} + \frac{\partial^4\psi}{\partial x_2^4}. \end{aligned}$$

Аппроксимация четвертой производной функции ψ по каждой из координат в соответствии с (7.8) требует использования значений этой функции в пяти соседних узлах конечно-разностной сетки, расположенных в направлении изменения координаты. Для аппроксимации смешанной производной потребуется девять узловых значений функции. В случае квадратной сетки с шагом h (рис. 9.4) с учетом (7.3) для узла с номерами m и n запишем

$$\begin{aligned} \left. \frac{\partial^4\psi}{\partial x_1^2 \partial x_2^2} \right|_{m,n} &\approx \frac{\psi_{m+1,n+1} - 2\psi_{m+1,n} + \psi_{m+1,n-1}}{h^4} - \\ &- 2 \frac{\psi_{m,n+1} - 2\psi_{m,n} + \psi_{m,n-1}}{h^4} + \frac{\psi_{m-1,n+1} - 2\psi_{m-1,n} + \psi_{m-1,n-1}}{h^4}. \end{aligned}$$

Можно показать, что погрешность такой аппроксимации имеет второй порядок. В итоге, учитывая (7.9), для (9.28) получаем

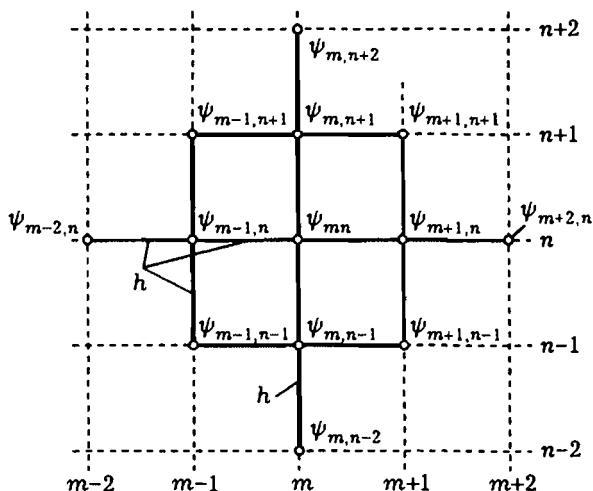


Рис. 9.4

разностную схему

$$\begin{aligned} & \frac{1}{h^4} (\psi_{m+2,n} + \psi_{m-2,n} + \psi_{m,n+2} + \psi_{m,n-2}) + \\ & + \frac{2}{h^4} (\psi_{m+1,n+1} + \psi_{m-1,n+1} + \psi_{m+1,n-1} + \psi_{m-1,n-1}) - \\ & - \frac{8}{h^4} (\psi_{m+1,n} + \psi_{m-1,n} + \psi_{m,n+1} + \psi_{m,n-1}) + \frac{20}{h^4} \psi_{mn} = f_{mn}, \end{aligned}$$

где f_{mn} — значение функции $f(M)$ в узле с номерами m и n .

Надо сказать, что эта схема обладает существенным недостатком. Если ее использовать для вычисления значений ψ_{mn} , переходя последовательно от одного узла сетки к другому, то такой процесс последовательных приближений сходится, но очень медленно. Причина заключается в том, что значение ψ_{mn} является разностью двух почти одинаковых сумм, заметно превосходящих это значение. Кроме того, эта схема применима лишь в узлах, отстоящих на два шага от границы. Так как помимо значений функции ψ на границе области обычно заданы значения ее первой или второй производной по направлению

нормали к границе (см. пример 3.1), то для аппроксимации этих производных в граничных узлах со вторым порядком погрешности приходится применять левые или правые конечные разности, включающие значения в трех и четырех узлах соответственно. Чтобы использовать для аппроксимации этих производных центральные разности, можно ввести дополнительный ряд фиктивных узлов за пределами границы. Такой прием допускает применение этой схемы во всех внутренних узлах, но тем не менее не устраняет отмеченного недостатка и затрудняет алгоритмизацию процесса последовательных приближений.

При решении прикладных задач удобнее (9.28) заменить системой двух уравнений второго порядка

$$\nabla^2 \psi(M) = -\zeta(M), \quad \nabla^2 \zeta(M) + f(M) = 0, \quad M \in F. \quad (9.29)$$

При двумерном течении несжимаемой вязкой жидкости по физическому смыслу функция $\zeta(M)$ является *завитренностью* (см. пример 3.1). Если задать в качестве нулевого приближения ожидаемое распределение функции $\zeta(M)$, то в соответствии с (3.29) граничных условий будет достаточно для однозначного решения при помощи МКР первого уравнения в (9.29). После этого недостающие граничные условия для функции ζ можно найти, рассматривая первое уравнение в граничных точках. Завершает итерацию решение второго уравнения в (9.29), дающее узловые значения функции ζ , используемые на следующей итерации при решении первого уравнения, и т.д. Вычислительная практика показывает, что такие итерации сходятся существенно быстрее, чем при решении исходного уравнения (9.28).

В некоторых случаях могут быть заданы граничные значения обеих искомых функций в (9.29). Так, поперечные прогибы $w(M)$ упругой пластинки площадью F в плане, нагруженной распределенной поперечной нагрузкой $q(M)$, описывает диф-

ференциальное уравнение

$$\nabla^2(D\nabla^2 w(M)) = q(M), \quad M \in F, \quad (9.30)$$

где $D = \frac{Eh_0^3}{12(1-\nu^2)}$ — жесткость пластинки толщиной h_0 с модулем упругости E и коэффициентом Пуассона ν . Это уравнение применимо, если прогибы малы по сравнению с h_0 . Погонные изгибающие моменты $\mathcal{M}_i$, $i = 1, 2$, приходящиеся на единицу длины поперечного сечения пластинки и перпендикулярные осям Ox_i , можно выразить через вторые производные от функции прогиба:

$$\mathcal{M}_1 = -D \left(\frac{\partial^2 w}{\partial x_1^2} + \nu \frac{\partial^2 w}{\partial x_2^2} \right), \quad \mathcal{M}_2 = -D \left(\frac{\partial^2 w}{\partial x_2^2} + \nu \frac{\partial^2 w}{\partial x_1^2} \right).$$

Отсюда

$$\mathcal{M} = \frac{\mathcal{M}_1 + \mathcal{M}_2}{1 + \nu} = -D \left(\frac{\partial^2 w}{\partial x_1^2} + \frac{\partial^2 w}{\partial x_2^2} \right) = -D \nabla^2 w.$$

Следовательно, (9.30) можно представить системой двух уравнений второго порядка

$$\nabla^2 w(M) = -\frac{\mathcal{M}}{D}, \quad \nabla^2 \mathcal{M}(M) + q(M) = 0, \quad M \in F. \quad (9.31)$$

Если пластинка по контуру Γ , ограничивающему область F , не закреплена жестко, а лишь оперта так, что касательная плоскость в граничных точках может свободно поворачиваться при изгибе пластинки, то изгибающие моменты на краях пластинки будут равны нулю. Тогда граничные условия для системы (9.31) примут вид $w(P) = 0$, $\mathcal{M}(P) = 0$, $P \in \Gamma$, что позволяет при заданной функции $q(M)$, $M \in F$, сначала при помощи МКР решить второе уравнение в (9.31), а затем по найденным узловым значениям функции $\mathcal{M}$ также при помощи МКР решить первое уравнение.

Отметим, что выше рассмотрены сравнительно простые примеры многомерных задач. Достаточно сложные разностные схемы приходится строить для решения при помощи МКР многомерных задач, которые описывают системы дифференциальных уравнений различных типов, причем среди этих уравнений могут быть и нелинейные. Характерные примеры таких задач можно найти в специальной литературе*.

Дополнение 9.1. Алгоритмы матричной и ортогональной прогонок

Любое линейное обыкновенное дифференциальное уравнение (ОДУ) высокого порядка или систему таких ОДУ можно свести к *нормальной системе ОДУ*. На *одномерной сетке* с узлами x_n , $n = \overline{0, N}$, *разностную схему* для такой системы можно представить в *матричной записи*:

$$\begin{cases} P_{n+1}v_{n+1} - Q_nv_n = f_{n+1}, & n = \overline{0, N-1}; \\ P_0v_0 = f_0, & Q_Nv_N = f_{N+1}, \end{cases} \quad (9.32)$$

где v_n — вектор размера K искомых неизвестных в узле с номером n ; векторы f_{n+1} , f_0 и f_{N+1} имеют размеры K , K_1 и K_2 соответственно (причем $K = K_1 + K_2$) и в качестве координат имеют *узловые значения* правых частей уравнений исходной нормальной системы ОДУ и ее граничных условий; элементы квадратных матриц P_{n+1} и Q_n порядка K зависят от узловых значений коэффициентов при неизвестных в исходной системе ОДУ, а элементы в общем случае прямоугольных матриц P_0 и Q_N размеров $K_1 \times K$ и $K_2 \times K$ соответственно зависят от коэффициентов в заданных граничных условиях рассматриваемой краевой задачи.

Разностную схему вида (9.32) можно получить не только непосредственной аппроксимацией производных в нормальной

*См.: Галинин М.П., Попов Ю.П., Самарский А.А., Попов Ю.П., Белоцерковский О.М.; Андерсон Д., Таннехилл Дж., Плетчер Р., Роуч П.

системе ОДУ. К подобному виду можно свести любую разностную схему, полученную в результате аппроксимации производных в ОДУ высокого порядка или в системе таких ОДУ. Например, обозначая применительно к разностной схеме (8.27), (8.28)

$$\begin{aligned} \mathbf{v}_n &= (u_{n+1} \ u_n \ u_{n-1} \ u_{n-2})^T, \quad n = \overline{2, N-1}; \quad \mathbf{f}_2 = (f_0 \ f_1)^T; \\ \mathbf{f}_{n+1} &= (f_n \ 0 \ 0 \ 0)^T, \quad n = \overline{2, N-2}; \quad \mathbf{f}_N = (f_{N-1} \ f_N)^T, \end{aligned}$$

вместо (8.27), (8.28) получаем

$$\begin{aligned} P_{n+1} \mathbf{v}_{n+1} - Q_n \mathbf{v}_n &= \mathbf{f}_{n+1}, \quad n = \overline{2, N-2}; \\ P_2 \mathbf{v}_2 &= \mathbf{f}_2, \quad Q_{N-1} \mathbf{v}_{N-1} = \mathbf{f}_N, \end{aligned}$$

где

$$\begin{aligned} P_{n+1} &= \begin{pmatrix} e_n & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad Q_n = \begin{pmatrix} d_n & -c_n & b_n & -a_n \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \\ P_2^T &= \begin{pmatrix} 0 & e_1 \\ e_0 & -d_1 \\ -d_0 & c_1 \\ c_0 & -b_1 \end{pmatrix}, \quad Q_{N-1}^T = \begin{pmatrix} -d_{N-1} & c_N \\ c_{N-1} & -b_N \\ -b_{N-1} & a_N \\ a_{N-1} & 0 \end{pmatrix}, \end{aligned}$$

т.е. в данном случае $K = 4$ и $K_1 = K_2 = 2$.

Краевую задачу, описываемую системой линейных ОДУ второго порядка, целесообразно аппроксимировать на одномерной сетке разностной схемой, сходной по структуре с (8.8), (8.9), но в матричной записи

$$\begin{cases} -A_n \mathbf{u}_{n-1} + B_n \mathbf{u}_n - C_n \mathbf{u}_{n+1} = \mathbf{f}_n, & n = \overline{1, N-1}; \\ B_0 \mathbf{u}_0 - C_0 \mathbf{u}_1 = \mathbf{f}_0, \quad -A_N \mathbf{u}_{N-1} + B_N \mathbf{u}_N = \mathbf{f}_N, \end{cases} \quad (9.33)$$

где $\mathbf{u}_n$ и $\mathbf{f}_n$ — в общем случае векторы размера K_n ; A_n и C_n — прямоугольные матрицы размеров $K_n \times K_{n-1}$ и $K_n \times K_{n+1}$

соответственно; B_n — квадратные матрицы порядка K_n . В частном случае $K_{n-1} = K_n = K_{n+1} = K$ матрицы A_n и C_n также становятся квадратными.

Для решения (9.33) можно построить алгоритм **матричной прогонки**, аналогичный алгоритму обычного **метода прогонки**:

$$\begin{cases} u_N = g_N; \\ u_n = E_n u_{n+1} + g_n, \quad n = N-1, N-2, \dots, 0, \end{cases} \quad (9.34)$$

где

$$\begin{aligned} E_0 &= B_0^{-1} C_0; \quad E_n = D_n C_n, \quad D_n = (B_n - A_n E_{n-1})^{-1}, \quad n = \overline{1, N-1}; \\ g_0 &= B_0^{-1} f_0; \quad g_n = D_n (f_n + A_n g_{n-1}), \quad n = \overline{1, N}. \end{aligned}$$

Здесь E_n — прямоугольная матрица размера $K_n \times K_{n+1}$, а g_n — вектор размера K_n .

Корректность разностной схемы (9.33) в сочетании с алгоритмом (9.34) матричной прогонки гарантирована*, если матрицы B_n , $n = \overline{0, N}$, невырожденные, а матрицы A_n и C_n , $n = \overline{1, N-1}$, ненулевые и, кроме того, выполнены неравенства $\|B_0^{-1} C_0\| \leq 1$, $\|B_N^{-1} A_N\| \leq 1$ и $\|B_n^{-1} A_n\| + \|B_n^{-1} C_n\| \leq 1$, $n = \overline{1, N-1}$, для норм матриц в соответствующих конечномерных векторных пространствах, причем хотя бы в одном из них имеет место строгое неравенство. Эти достаточные условия являются матричным аналогом достаточных условий (8.85) корректности обычного метода прогонки. При выполнении указанных условий матрицы $B_n - A_n E_{n-1}$, $n = \overline{1, N}$, будут невырожденными, т.е. существуют обратные к ним матрицы D_n , входящие в выражения для матриц E_n , и $\|E_n\| \leq 1$, что обеспечивает устойчивость алгоритма (9.34) относительно вычислительных погрешностей.

Если в (9.33) все векторы имеют одинаковый размер K , совпадающий с числом искомых действительных функций в

*См.: Самарский А.А., Николаев Е.С.

рассматриваемой задаче, а все матрицы — квадратные порядка K , то реализация алгоритма матричной прогонки при заранее найденных матрицах $D_n = (B_n - A_n E_{n-1})^{-1}$, $n = \overline{1, N}$, требует примерно $6K^2N$ арифметических операций. Однако для обращения каждой из матриц $B_n - A_n E_{n-1}$ необходимо порядка K^3 операций. Таким образом, при однократном решении задачи общее число арифметических операций можно оценить как $K^3N + 6K^2N$. Если данную задачу решают m раз для различных правых частей ОДУ или граничных условий, то матрицы D_n остаются неизменными и затраты на выполнение всей серии расчетов составят $K^3N + 6mK^2N$ операций, а на вычисление одного узлового значения — $\frac{K^2}{m} + 6K$ операций. Для сравнения напомним, что в алгоритме обычного метода прогонки для вычисления одного узлового значения при однократном решении задачи необходимо примерно восемь арифметических операций.

Разностную схему (9.32) также можно свести к схеме (9.33) и затем использовать для получения решения алгоритм (9.34)*. Непосредственное решение (9.32) можно получить при помощи алгоритма, называемого **ортогональной прогонкой** и связанного с обращением матриц P_{n+1} , $n = \overline{0, N-1}$, и ортогонализацией столбцов вспомогательных прямоугольных матриц.

Искомое решение (9.32) представим в виде

$$v_n = S_n w_n + y_n, \quad n = \overline{0, N}, \quad (9.35)$$

где S_n — прямоугольная матрица размера $K \times K_2$, а w_n и y_n — векторы размера K_2 и K соответственно. Чтобы при $n = 0$ решение (9.35) удовлетворяло равенству $P_0 v_0 = f_0$, матрицу S_0 и вектор y_0 находят из условий

$$P_0 y_0 = f_0 \quad \text{и} \quad P_0 S_0 = O^{(12)}, \quad (9.36)$$

где $O^{(12)}$ — прямоугольная нулевая матрица размера $K_1 \times K_2$. Способ нахождения S_0 и y_0 рассмотрим в дальнейшем.

*См.: Самарский А.А., Николаев Е.С.

Предполагая невырожденность матриц P_{n+1} в (9.32), после подстановки (9.35) в первое равенство (9.32) получаем

$$S_{n+1}w_{n+1} + y_{n+1} - P_{n+1}^{-1}Q_n S_n w_n = P_{n+1}^{-1}(f_{n+1} + Q_n y_n),$$

$n = \overline{0, N-1}$, или

$$S_{n+1}w_{n+1} + y_{n+1} - R_{n+1}w_n = r_{n+1}, \quad n = \overline{0, N-1}, \quad (9.37)$$

где

$$R_{n+1} = P_{n+1}^{-1}Q_n S_n \quad \text{и} \quad r_{n+1} = P_{n+1}^{-1}(f_{n+1} + Q_n y_n), \quad (9.38)$$

причем прямоугольная матрица R_{n+1} имеет размер $K \times K_2$, а вектор r_{n+1} — размер K .

Далее используем равенства

$$\begin{cases} y_{n+1} = r_{n+1} - S_{n+1}z_{n+1} = \\ \qquad \qquad \qquad = P_{n+1}^{-1}(f_{n+1} + Q_n y_n) - S_{n+1}z_{n+1}, \\ R_{n+1} = P_{n+1}^{-1}Q_n S_n = S_{n+1}\Omega_{n+1}, \end{cases} \quad (9.39)$$

$n = \overline{0, N-1}$, в которых неизвестны квадратная матрица Ω_{n+1} порядка K_2 и вектор z_{n+1} размера K_2 . Из (9.37) и (9.39) следует соотношение

$$S_{n+1}(w_{n+1} - \Omega_{n+1}w_n) = S_{n+1}z_{n+1}, \quad n = \overline{0, N-1},$$

которое становится тождеством, если положить

$$\Omega_{n+1}w_n = w_{n+1} - z_{n+1}, \quad n = \overline{0, N-1}, \quad (9.40)$$

Таким образом, если векторы z_{n+1} и невырожденные матрицы Ω_{n+1} , $n = \overline{0, N-1}$, а также матрица S_0 и вектор y_0 были бы известны, то, увеличивая номер n , по рекуррентным формулам

$$\begin{cases} y_{n+1} = P^{-1}(f_{n+1} + Q_n y_n) - S_{n+1}z_{n+1}, \\ S_{n+1} = \Omega_{n+1}^{-1}P_{n+1}^{-1}Q_n S_n, \\ n = \overline{0, N-1}, \end{cases} \quad (9.41)$$

которые следуют из (9.39), можно найти векторы y_{n+1} и матрицы S_{n+1} . Тогда в правой части (9.35) будут неизвестны лишь векторы w_n , $n = \overline{0, N}$.

Из последнего равенства (9.32), а также равенства (9.35) при $n = N$ получаем $Q_N v_N = Q_N (S_N w_N + y_N) = f_{N+1}$, или

$$Q_N S_N w_N = f_{N+1} - Q_N y_N.$$

Отсюда можно найти вектор w_N размера K_2 , если квадратная матрица $Q_N S_N$ порядка K_2 не является вырожденной. Затем, последовательно выбирая $n = N - 1, N - 2, \dots, 0$, из (9.40) находим все остальные векторы

$$w_n = \Omega_{n+1}^{-1} (w_{n+1} - z_{n+1}), \quad (9.42)$$

а из (9.35) — искомое решение для (9.32).

В выборе матриц Ω_n и векторов z_n есть некоторый произвол. Этот выбор целесообразно провести из условий

$$S_n^T S_n = I^{(2)} \quad \text{и} \quad S_n^T y_n = 0^{(2)}, \quad n = \overline{0, N}, \quad (9.43)$$

где $I^{(2)}$ — единичная матрица порядка K_2 , а $0^{(2)}$ — нулевой вектор размера K_2 . Выполнение этих условий гарантирует устойчивость рассматриваемого алгоритма относительно накопления вычислительных погрешностей, поскольку столбцы каждой из матриц S_n образуют систему ортонормированных векторов и норма каждого вектора и матрицы в целом не превышает единицы.

Каждую из матриц S_n , $n = \overline{0, N}$, получим из матрицы R_n ортонормированием столбцов последней так, чтобы $S_n = R_n \Omega_n$. Сначала для $n = 0$ представим в виде блочной матрицы

$$R_0 = \begin{pmatrix} (P_0^{(11)})^{-1} P_0^{(12)} \\ I^{(2)} \end{pmatrix},$$

где $P_0^{(11)}$ — невырожденная квадратная матрица порядка K_1 и $P_0^{(12)}$ — прямоугольная матрица размера $K_1 \times K_2$, являющиеся *блоками* прямоугольной матрицы P_0 размера $K_1 \times K$ ($K = K_1 + K_2$) при ее представлении в виде блочной: $P_0 = (P_0^{(11)} \quad P_0^{(12)})$. Обозначим через S_r и R_r столбцы матриц S_0 и R_0 с номером r . Предположим, что прямоугольная матрица R_0 размера $K \times K_2$ имеет ранг K_2 , т.е. столбцы R_r , $r = \overline{1, K_2}$, являются линейно независимыми. Тогда, используя процесс ортогонализации Грама — Шмидта, находим

$$S_r = \frac{1}{\Omega_{11}} R_r, \quad \Omega_{11} = (R_1^T R_1)^{1/2}, \quad S_r = \frac{1}{\Omega_{rr}} \left(R_r - \sum_{i=1}^{r-1} \Omega_{ir} S_i \right),$$

$$\Omega_{ir} = (S_i^T R_r)^{1/2}, \quad \Omega_{rr} = \left(R_r^T R_r - \sum_{i=1}^{r-1} \Omega_{ir}^2 \right)^{1/2}, \quad r = \overline{1, K_2},$$

где Ω_{ir} — элементы верхней треугольной матрицы Ω_0 порядка K_2 , т.е. $\Omega_{ir} = 0$ при $1 \leq r < i \leq K_2$.

Так как $S_0 = R_0 \Omega_0^{-1}$ и с учетом правил умножения блочных матриц

$$\begin{aligned} P_0 R_0 &= \begin{pmatrix} P_0^{(11)} & P_0^{(12)} \end{pmatrix} \begin{pmatrix} (P_0^{(11)})^{-1} P_0^{(12)} \\ I^{(2)} \end{pmatrix} = \\ &= P_0^{(11)} (P_0^{(11)})^{-1} P_0^{(12)} - P_0^{(12)} I^{(2)} = O^{(12)}, \end{aligned}$$

убеждаемся, что $P_0 S_0 = P_0 R_0 \Omega_0^{-1} = O^{(12)}$, т.е. матрица S_0 удовлетворяет второму условию (9.36). Чтобы удовлетворить первому условию (9.36), построим *матрицу-столбец*

$$X_0 = \begin{pmatrix} (P_0^{(11)})^{-1} f_0 \\ 0^{(2)} \end{pmatrix}$$

и выразим через нее вектор $y_0 = X_0 - S_0 S_0^T X_0$. Тогда, действительно, получим

$$\begin{aligned} P_0 y_0 &= P_0 X_0 - P_0 S_0 S_0^T X_0 = P_0 X_0 = \\ &= \begin{pmatrix} P_0^{(11)} & -P_0^{(12)} \end{pmatrix} \begin{pmatrix} (P_0^{(11)})^{-1} f_0 \\ \mathbf{0}^{(2)} \end{pmatrix} = \\ &= P_0^{(11)} (P_0^{(11)})^{-1} f_0 - P_0^{(12)} \mathbf{0}^{(2)} = f_0. \end{aligned}$$

По найденной матрице S_0 и известным матрицам Q_0 и P_1 из второго равенства (9.39) находим матрицу $R_1 = P_1^{-1} Q_0 S_0$ и, проводя процедуру ортонормирования ее столбцов, получаем матрицы S_1 и Ω_1 , а затем и все остальные матрицы S_{n+1} и Ω_{n+1} , $n = 1, N-1$.

Для нахождения векторов z_{n+1} , $n = 0, N-1$, подставим первое равенство (9.39) во второе условие (9.43) и, используя первое условие (9.43), запишем

$$S_{n+1}^T y_{n+1} = S_{n+1}^T r_{n+1} - S_{n+1}^T S_{n+1} z_{n+1} = S_{n+1}^T r_{n+1} - z_{n+1} = \mathbf{0}^{(2)}.$$

Отсюда следует, что $z_{n+1} = S_{n+1}^T r_{n+1}$, $n = 0, N-1$. При $n = 0$ из (9.38) находим $r_1 = P_1^{-1} (f_1 + Q_0 y_0)$ и затем $z_1 = S_1^T r_1$, что позволяет из первого равенства (9.39) получить вектор $y_1 = r_1 - S_1 z_1$, найти r_2 , z_2 и т.д.

Таким образом, установлены все соотношения, позволяющие получить решение системы (9.32) при помощи алгоритма ортогональной прогонки.

Вопросы и задачи

9.1. Показать, что при выборе в (9.4) и (9.19) $\eta_1 = \eta_2 = 0$ алгоритм вычислений будет устойчивым при выполнении условий $\tilde{a} \Delta t_k \leq \frac{h_1^2 h_2^2}{h_1^2 + h_2^2}$ и $(a\tau)^2 \leq \frac{h_1^2 h_2^2}{h_1^2 + h_2^2}$ соответственно.

9.2. Методом установления решить задачу 6.6, используя разностную схему (9.5) при $\eta_1 = \eta_2 = 1/2$.

9.3. Каков порядок погрешности аппроксимации в (9.16)? Как можно разбить полуинтервал $(0, 1]$ на промежутки, чтобы выбрать направление перехода блуждающей частицы в соседний узел (см. пример 9.1), если (9.14) аппроксимировать на прямоугольной сетке, для которой $h_1 = 2h_2$, т.е. шаг по оси Ox_1 в два раза больше шага по оси Ox_2 ?

9.4. Показать, что погрешность аппроксимации разностной схемы, построенной для (9.28), имеет второй порядок.

9.5. Построить разностную схему для (9.28) с переменными шагами сетки по обеим координатам.

9.6. Решить задачу 6.6 с помощью матричной и ортогональной прогонок.

ЧАСТЬ IV

Методы конечных и граничных элементов

10. ОСНОВЫ МЕТОДА КОНЕЧНЫХ ЭЛЕМЕНТОВ

При изложении *проекционных методов* (см. 6) решения задач математической физики отмечено, что в качестве *базисных функций* могут быть использованы кусочно непрерывные функции, отличные от нуля лишь в отдельных конечных подобластях той области, где рассматривают решение задачи. Такие подобласти в сочетании с выбранным типом базисных функций принято называть **конечными элементами (КЭ)**, что и дало название методу решения задач, опирающемуся на указанный подход к построению базисных функций. По существу *метод конечных элементов (МКЭ)* является *проекционно-сеточным методом* в том смысле, что процедуру, характерную для проекционных методов и базирующуюся на наличии *интегральной формулировки задачи*, реализуют на совокупности КЭ, заполняющей область решения задачи. Эту совокупность КЭ называют **сеткой конечных элементов**.

В пределах каждого КЭ искомое решение приближенно представляют многочленом. Коэффициенты этого многочлена выражают через заранее неизвестные значения искомой функции (в более общем случае — и значения ее производных) в определенным образом выбранных точках КЭ, называемых **узлами конечного элемента**. Как и в *методе конечных разностей* эти значения называют **узловыми**, причем узловые значения искомой функции и ее производных объединяют общим названием **узловые параметры**. Объединив отдельные КЭ в сетку, удастся выразить искомое решение через неизвестные узловые параметры, которые затем находят, используя интегральную формулировку задачи.

Сначала на достаточно простом примере одномерной краевой задачи для обыкновенного дифференциального уравнения

(ОДУ) второго порядка последовательно рассмотрим отдельные этапы применения КЭ к решению такой задачи. Эти этапы позволят сформировать более детальное представление о сущности МКЭ.

10.1. Одномерная краевая задача

Пусть на отрезке $[0, 1]$ определены ограниченные функции $f(x)$, $p(x)$ и $q(x)$, причем

$$p(x) \geq p_0 > 0, \quad q(x) \geq 0, \quad x \in [0, 1]. \quad (10.1)$$

На этом отрезке будем искать решение $u(x)$ линейного ОДУ

$$Au = f(x) \quad (10.2)$$

с дифференциальным оператором

$$A(u) = -(p(x)u'(x))' + q(x)u,$$

удовлетворяющее граничным условиям

$$u(0) = 0, \quad (10.3)$$

$$u'(1) = 0. \quad (10.4)$$

При выполнении условий (10.1) оператор $A(u)$ будет положительно определенным на множестве X дважды непрерывно дифференцируемых на $[0, 1]$ функций $u(x)$, удовлетворяющих условиям (10.3) и (10.4) (см. пример 5.10). Краевая задача (10.2)–(10.4) может иметь на множестве X классическое решение $u^*(x)$, если функции $f(x)$ и $q(x)$ непрерывны, а функция $p(x)$ непрерывно дифференцируема на отрезке $[0, 1]$. При этом функция $u^*(x)$ минимизирует функционал энергии

$$J[u] = \|u\|_A^2 - 2\langle f, u \rangle = \int_0^1 (p(u')^2 + qu^2 - 2fu) dx, \quad (10.5)$$

который допустимо рассматривать на более широком множестве функций $u(x)$ по сравнению с множеством X . Действительно, (10.5) сохраняет смысл на множестве $X^* \supset X$ функций, удовлетворяющих главному для (10.5) граничному условию (10.3) и имеющих на отрезке $[0, 1]$ лишь кусочно непрерывную производную (при условии, что множество точек разрыва производной имеет на $[0, 1]$ меру Лебега, равную нулю). Более того, (10.5) сохраняет смысл, если функции $p(x)$ и $q(x)$, удовлетворяющие условиям (10.1), и функция $f(x)$ имеют на отрезке $[0, 1]$ конечное число точек разрыва первого рода. Напомним, что функцию $u_*(x)$, минимизирующую при этом (10.5), называют *обобщенным решением* задачи (10.2)–(10.4).

Вариационная формулировка задачи (10.2)–(10.4), содержащая функционал (10.5), позволяет для поиска приближения к обобщенному решению использовать *метод Рунца*. Это приближение можно искать в виде

$$\tilde{u}_N(x) = \sum_{n=1}^N a_n u_n(x), \quad a_n \in \mathbb{R}, \quad (10.6)$$

где a_n — искомые коэффициенты, которые зависят от выбора системы *базисных функций* $u_n(x)$, $n = \overline{1, N}$. Подчеркнем, что (10.6), называемое при фиксированном N *приближенным решением операторного уравнения* (10.2), должно удовлетворять главному для функционала (10.5) граничному условию (10.3), но может не удовлетворять его *естественному граничному условию* (10.4). Действительно, (10.4) следует из условия $\delta J = 0$ стационарности функционала (10.5), получаемого приравливанием нулю его *первой вариации* [XV]

$$\begin{aligned} \delta J[u, \delta u] &= 2 \int_0^1 p u' \delta u' dx + 2 \int_0^1 (q u - f) \delta u dx = \\ &= 2 p(x) u'(x) \delta u(x) \Big|_0^1 + 2 \int_0^1 (-(p u')' + q u - f) \delta u dx = 0. \end{aligned}$$

Это равенство при произвольной на полуинтервале $(0, 1]$ вариации $\delta u(x)$ функции $u(x)$ и выполнении (10.3), т.е. $\delta u(0) = 0$, приводит к двум условиям в виде (10.2) и (10.4).

Если на отрезке $[0, 1]$ функции $p(x)$, $q(x)$, $f(x)$ имеют точки разрыва первого рода, то включим их в число точек $x_n \in [0, 1]$, $n = \overline{0, N}$, при помощи которых проведем разбиение этого отрезка на N частичных отрезков. В (10.6) в качестве базисных функций на отрезке $[0, 1]$ выберем

$$u_n(x) = \begin{cases} 0, & x < x_{n-1}; \\ \frac{x - x_{n-1}}{x_n - x_{n-1}}, & x \in [x_{n-1}, x_n]; \\ \frac{x_{n+1} - x}{x_{n+1} - x_n}, & x \in (x_n, x_{n+1}]; \\ 0, & x > x_{n+1}, \end{cases} \quad n = \overline{1, N-1}, \quad (10.7)$$

и, кроме того, $u_N = \frac{x - x_{N-1}}{x_N - x_{N-1}}$ при $x \in [x_{N-1}, x_N]$ и $u_N = 0$ при $x < x_{N-1}$, где x_n — координаты точек отрезка $[0, 1]$, являющихся узлами КЭ, расположенными на границах между элементами, причем $x_0 = 0$ и $x_N = 1$.

Таким образом, $u_n(x_n) = 1$ и функция $u_n(x)$ линейно изменяется в пределах частичных отрезков $[x_{n-1}, x_n]$ и $[x_n, x_{n+1}]$, примыкающих к точке x_n , $n = \overline{1, N-1}$, принимая на их противоположных концах нулевое значение (рис. 10.1). Это значение

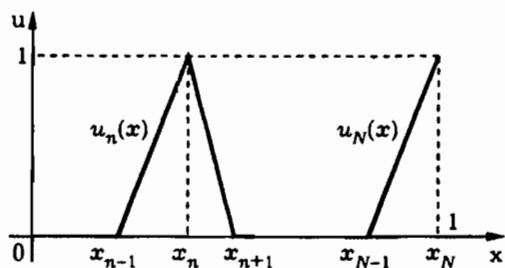


Рис. 10.1

остается неизменным на всех частичных отрезках, не содержащих точку x_n . При $n = N$ имеем $u_N(x_N) = 1$, а функция $u_N(x)$ линейно изменяется в пределах отрезка $[x_{N-1}, x_N]$ и равна нулю на остальных частичных отрезках.

На любом частичном отрезке $[x_{n-1}, x_n]$, $n = \overline{2, N}$, при помощи базисных функций $u_{n-1}(x)$ и $u_n(x)$ можно однозначно представить в виде $a_{n-1}u_{n-1} + a_nu_n(x)$ линейную функцию $u(x)$, принимающую в точках x_{n-1} и x_n значения $u(x_{n-1}) = a_{n-1}$ и $u(x_n) = a_n$ соответственно. В самом деле, учитывая (10.7), имеем

$$\begin{aligned} u(x) &= a_{n-1} \frac{x_n - x}{x_n - x_{n-1}} + a_n \frac{x - x_{n-1}}{x_n - x_{n-1}} = \\ &= a_{n-1} + (a_n - a_{n-1}) \frac{x - x_{n-1}}{x_n - x_{n-1}}. \end{aligned}$$

Ясно, что при $n = 1$ функция $u_1(x)$ однозначно представляет на частичном отрезке $[0, x_1]$ линейную функцию $u(x) = a_1u_1(x)$, имеющую на его концах значения $u_1(0) = 0$ и $u(x_1) = a_1$.

Каждый частичный отрезок в сочетании с определенными на нем линейными базисными функциями является простейшим примером *конечного элемента* (КЭ) с двумя узлами на его концах. Базисную функцию, в одном узле КЭ равную единице, а во всех остальных узлах (этого и других КЭ) равную нулю, называют **функцией формы** этого **конечного элемента**. Ее номер совпадает с номером узла, в котором она равна единице. В данном случае каждый из КЭ имеет две функции формы, определяемые (10.7). Геометрически линейная функция формы $u_n(x)$ в точке $x \in [x_{n-1}, x_n]$ равна отношению длин отрезков $[x_{n-1}, x]$ и $[x_{n-1}, x_n]$.

Если в (10.6) использовать систему базисных функций (10.7) с добавлением функции $u_N(x)$, то $\tilde{u}_N(x)$ будет на отрезке $[0, 1]$ непрерывной кусочно линейной функцией, принимающей в точках x_n , $n = \overline{1, N}$, значения $\tilde{u}_N(x_n) = a_n$ и $\tilde{u}_N(0) = 0$ (рис. 10.2). Чтобы не выделять особо частичный отрезок $[0, x_1]$, можно

на нем ввести функцию формы $u_0(x) = 1 - \frac{x}{x_1}$ при $x \in [0, x_1]$ и $u_0(x) = 0$ при $x > x_1$ и вместо (10.6) написать

$$\tilde{u}_N(x) = \sum_{n=0}^N a_n u_n(x), \quad a_n \in \mathbb{R}, \quad (10.8)$$

положив в соответствии с главным для функционала (10.5) граничным условием (10.3) $a_0 = \tilde{u}_N(0) = 0$.

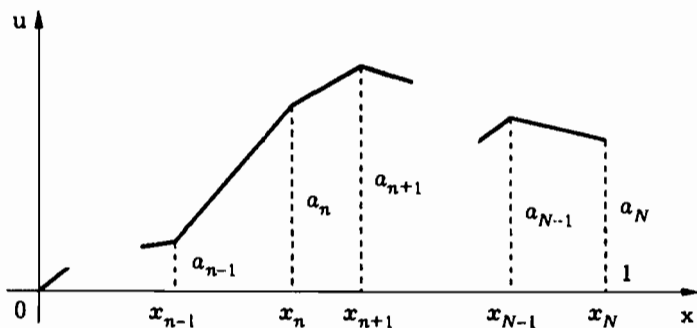


Рис. 10.2

Приближенное решение (10.8) принадлежит множеству X^* , на котором допустимо рассматривать функционал (10.5). Используя аддитивность определенного интеграла [VI], представим (10.5) в виде

$$J[u] = \sum_{n=1}^N \int_{x_{n-1}}^{x_n} (p(u')^2 + qu^2 - 2fu) dx. \quad (10.9)$$

На каждом частичном отрезке $[x_{n-1}, x_n]$, $n = \overline{1, N}$, запишем

$$\tilde{u}_N(x) = a_{n-1} u_n(x) + a_n u_n(x) = a_{n-1} + (a_n - a_{n-1}) \frac{x - x_{n-1}}{x_n - x_{n-1}}.$$

Отсюда дифференцированием находим

$$\tilde{u}'_N(x) = a_{n-1}u'_n(x) + a_n u'_n(x) = \frac{a_n - a_{n-1}}{x_n - x_{n-1}},$$

$x \in [x_{n-1}, x_n]$, причем $a_0 = 0$. Подставляя $\tilde{u}_N(x)$ и $\tilde{u}'_N(x)$ в (10.9) вместо u и $u'(x)$ соответственно, получаем

$$\begin{aligned} J[\tilde{u}_N] = & \sum_{n=1}^N \left(\frac{a_n - a_{n-1}}{x_n - x_{n-1}} \right)^2 \int_{x_{n-1}}^{x_n} p(x) dx + \\ & + \sum_{n=1}^N \int_{x_{n-1}}^{x_n} q(x) \left(a_{n-1} + (a_n - a_{n-1}) \frac{x - x_{n-1}}{x_n - x_{n-1}} \right)^2 dx - \\ & - 2 \sum_{n=1}^N \int_{x_{n-1}}^{x_n} f(x) \left(a_{n-1} + (a_n - a_{n-1}) \frac{x - x_{n-1}}{x_n - x_{n-1}} \right) dx. \quad (10.10) \end{aligned}$$

Теперь условие $\delta J = 0$ стационарности функционала (10.5) переходит в N необходимых условий $\frac{\partial J[\tilde{u}_N]}{\partial a_n} = 0$, $n = \overline{1, N}$, минимума квадратичной функции (10.10) переменных a_n [XV]. В итоге получаем систему линейных алгебраических уравнений (СЛАУ)

$$K\mathbf{a} = \mathbf{F}, \quad (10.11)$$

где $\mathbf{a}$ и $\mathbf{F}$ — N -мерные векторы с координатами a_n и

$$F_n = \int_{x_{n-1}}^{x_n} f(x) \frac{x - x_{n-1}}{x_n - x_{n-1}} dx + \int_{x_n}^{x_{n+1}} f(x) \frac{x_{n+1} - x}{x_{n+1} - x_n} dx, \quad (10.12)$$

а матрица K порядка N благодаря тому, что при фиксированном $n = \overline{1, N-1}$ в (10.10) отличны от нуля коэффициенты только

при произведениях $a_{n-1}a_n$ и a_na_{n+1} , является трехдиагональной с элементами

$$K_{nn} = \int_{x_{n-1}}^{x_n} \frac{p(x) + q(x)(x - x_{n-1})^2}{(x_n - x_{n-1})^2} dx + \\ + \int_{x_n}^{x_{n+1}} \frac{p(x) + q(x)(x_{n+1} - x)^2}{(x_n - x_{n-1})^2} dx, \quad (10.13)$$

$$K_{n,n-1} = - \int_{x_{n-1}}^{x_n} \frac{p(x) - q(x)(x_n - x)(x - x_{n-1})}{(x_n - x_{n-1})^2}, \quad (10.14)$$

$$K_{n,n+1} = - \int_{x_n}^{x_{n+1}} \frac{p(x) - q(x)(x_{n+1} - x)(x - x_n)}{(x_{n+1} - x_n)^2} dx \quad (10.15)$$

(при $n = N$ вторые интегралы в правых частях выражений для F_n и K_{nn} исчезают, а $K_{N,N+1}$ как и K_{10} не являются элементами матрицы K). При фиксированном $n = \overline{1, N-1}$ из равенств (10.14), (10.15) следует, что $K_{n,n+1} = K_{n+1,n}$, т.е. матрица K не только трехдиагональная, но и симметрическая.

Для произвольных функций $p(x)$, $q(x)$ и $f(x)$ интегралы в формулах (10.12) и (10.13)–(10.15) не удастся вычислить точно и приходится прибегать к методам численного интегрирования с применением *квадратурных формул*. Приближенное вычисление этих интегралов можно упростить, если на отрезке $[0, 1]$ использовать линейную аппроксимацию функций $p(x)$, $q(x)$ и $f(x)$ в виде

$$p(x) \approx \sum_{m=0}^N p_m u_m(x), \quad q(x) \approx \sum_{m=0}^N q_m u_m(x), \quad f(x) \approx \sum_{m=0}^N f_m u_m(x),$$

где $p_m = p(x_m)$, $q_m = q(x_m)$ и $f_m = f(x_m)$. Тогда, подставляя (10.6) в (10.5), получим

$$\begin{aligned} J[\tilde{u}_N] = & \int_0^1 \left(\sum_{m=0}^N p_m u_m(x) \right) \left(\sum_{n=1}^N a_n u'_n(x) \right)^2 dx + \\ & + \int_0^1 \left(\sum_{m=0}^N q_m u_m(x) \right) \left(\sum_{n=1}^N a_n u_n(x) \right)^2 dx - \\ & - 2 \int_0^1 \left(\sum_{m=0}^N f_m u_m(x) \right) \left(\sum_{n=1}^N a_n u_n(x) \right) dx. \end{aligned}$$

Из условий $\frac{\partial J[\tilde{u}_N]}{\partial a_n} = 0$ снова приходим к СЛАУ вида (10.11), но теперь с учетом

$$\int_0^1 u_{n-1}(x) u_{n+1}(x) dx = 0, \quad n = \overline{1, N-1},$$

координаты N -мерного вектора F равны

$$\begin{aligned} F_n = & \int_0^1 u_n(x) \sum_{m=0}^N f_m u_m(x) dx = \\ & = \int_{x_{n-1}}^{x_n} (f_{n-1} u_{n-1} + f_n u_n(x)) u_n(x) dx + \\ & + \int_{x_n}^{x_{n+1}} (f_n u_n(x) + f_{n+1} u_{n+1}(x)) u_n(x) dx, \quad n = \overline{1, N}, \quad (10.16) \end{aligned}$$

а элементы симметрической трехдиагональной матрицы K порядка N с учетом $u'_n(x) \equiv 0$, $x \notin [x_{n-1}, x_{n+1}]$, могут быть вы-

числены по следующим формулам (n принимает значения от 1 до N):

диагональные элементы

$$\begin{aligned}
 K_{nn} &= \int_0^1 \left((u'_n(x))^2 \sum_{m=0}^N p_m u_m(x) + u_n^2(x) \sum_{m=0}^N q_m u_m(x) \right) dx = \\
 &= \int_{x_{n-1}}^{x_n} \left(\frac{p_{n-1} u_{n-1}(x) + p_n u_n(x)}{(x_n - x_{n-1})^2} + \right. \\
 &\quad \left. + (q_{n-1} u_{n-1}(x) + q_n u_n(x)) u_n^2(x) \right) dx + \\
 &\quad + \int_{x_n}^{x_{n+1}} \left(\frac{p_n u_n(x) + p_{n+1} u_{n+1}(x)}{(x_{n+1} - x_n)^2} + \right. \\
 &\quad \left. + (q_n u_n(x) + q_{n+1} u_{n+1}(x)) u_n^2(x) \right) dx, \quad (10.17)
 \end{aligned}$$

недиагональные элементы

$$\begin{aligned}
 K_{n,n-1} &= \int_0^1 u'_n(x) u'_{n-1}(x) \sum_{m=0}^N p_m u_m(x) dx + \\
 &\quad + \int_0^1 u_n(x) u_{n-1}(x) \sum_{m=0}^N q_m u_m(x) dx = \\
 &= - \int_{x_{n-1}}^{x_n} \frac{p_{n-1} u_{n-1}(x) + p_n u_n(x)}{(x_n - x_{n-1})^2} dx + \\
 &\quad + \int_{x_{n-1}}^{x_n} (q_{n-1} u_{n-1}(x) + q_n u_n(x)) u_n(x) u_{n-1}(x) dx. \quad (10.18)
 \end{aligned}$$

При этом K_{10} не является элементом матрицы K . Снова при $n = N$ вторые интегралы в правых частях выражений для F_n и K_{nn} исчезают.

Интегралы от базисных функций в (10.16)–(10.18) можно вычислить по общей формуле*

$$\int_{x_{n-1}}^{x_n} u_{n-1}^r(x) u_n^s(x) dx = \frac{r!s!}{r+s+1} (x_n - x_{n-1}). \quad (10.19)$$

Тогда для $n = \overline{1, N}$ получим

$$F_n = \frac{f_{n-1} + 2f_n}{6} (x_n - x_{n-1}) + \frac{2f_n + f_{n+1}}{6} (x_{n+1} - x_n),$$

где в правой части при $n = N$ следует опустить второе слагаемое. Далее,

$$\begin{aligned} K_{nn} = & \frac{1}{2} \frac{p_{n-1} + p_n}{x_n - x_{n-1}} + \frac{1}{2} \frac{p_n + p_{n+1}}{x_{n+1} - x_n} + \\ & + \frac{q_{n-1} + 3q_n}{12} (x_n - x_{n-1}) + \frac{3q_n + q_{n+1}}{12} (x_{n+1} - x_n), \end{aligned} \quad (10.20)$$

где в правой части при $n = N$ следует опустить второе и четвертое слагаемые, и

$$K_{n,n-1} = -\frac{p_{n-1} + p_n}{2(x_n - x_{n-1})} + \frac{q_{n-1} + q_n}{12} (x_n - x_{n-1}). \quad (10.21)$$

Перечислим рассмотренные выше этапы:

1) переход от формулировки задачи в виде операторного уравнения к *интегральной формулировке* (в данном случае к вариационной формулировке, содержащей функционал (10.5));

2) разбиение области, в которой предстоит искать приближенное решение задачи (в данном случае отрезок $[0, 1]$), на

*См., например: Норри Д., де Фриз Ж.

подобласти и определение в них базисных функций, т.е. выбор типа КЭ;

3) формирование при помощи совокупности КЭ структуры приближенного решения и использование его в интегральной формулировке задачи (в данном случае приближенное решение в виде (10.6) принадлежит множеству функций, на котором допустимо рассматривать функционал (10.5));

4) получение СЛАУ $Ka = F$ для нахождения координат a_n вектора a , являющихся значениями искомой функции в узлах КЭ;

5) применение КЭ (при необходимости) для вычисления интегралов, входящих в выражения для элементов матрицы K и координат вектора F .

Эти этапы в целом и составляют процедуру получения приближенного решения задачи, называемую **методом конечных элементов (МКЭ)**. Ясно, что нахождение значений коэффициентов a_n требует применения тех или иных методов для решения СЛАУ. Однако этот этап неизбежен не только для завершения процедуры МКЭ и поэтому его обычно рассматривают как самостоятельный.

Замечание 10.1. Подчеркнем, что приближенное решение одномерной краевой задачи (10.2)–(10.4) при помощи МКЭ приводит к симметрической трехдиагональной матрице K в СЛАУ вида (10.11). Более того, если в (10.2) $p(x) \geq p_0 > 0$ и $q(x) \geq 0$ при $x \in [0, 1]$ и $q_n > 0$ хотя бы в одном из узлов с номером $n = \overline{1, N}$, то K является матрицей с **частичным диагональным преобладанием**, а для нахождения коэффициентов a_n в (10.6) эффективно применение **метода прогонки**. #

Из (10.12), (10.13)–(10.15) или (10.16)–(10.18) следует, что применение так называемых **финитных базисных функций** $u_n(x)$, $n = \overline{0, N}$, отличных от нуля только в пределах тех КЭ, которые содержат узел x_n , дает возможность независимо вычислять вклады отдельных элементов в матричное уравнение вида (10.11). Это очень удобно при алгоритмизации МКЭ по

сравнению с приближенными методами решения, в которых базисные функции могут быть отличны от нуля лишь в отдельных точках отрезка $[0, 1]$.

Чтобы получить представление о погрешности, которая может возникнуть при применении МКЭ, сравним n -е уравнение

$$K_{n(n-1)}\dot{u}_{n-1} + K_{nn}u_n + K_{n(n+1)}u_{n+1} = F_n$$

СЛАУ (10.11), которое с учетом (10.20), (10.21) принимает вид

$$\begin{aligned} & \frac{p_{n-1} + p_n}{2(x_n - x_{n-1})}(u_n - u_{n-1}) + \frac{p_{n+1} + p_n}{2(x_{n+1} - x_n)}(u_n - u_{n+1}) + \\ & + \frac{q_{n-1} + q_n}{12}(x_n - x_{n-1})(u_{n-1} + u_n) + \frac{q_n}{6}(x_{n+1} - x_{n-1})u_n + \\ & + \frac{q_n + q_{n+1}}{12}(x_{n+1} - x_n)(u_n + u_{n+1}) = \\ & = \frac{f_{n-1} + 2f_n}{6}(x_n - x_{n-1}) + \frac{2f_n + f_{n+1}}{6}(x_{n+1} - x_n), \quad (10.22) \end{aligned}$$

с конечно-разностной аппроксимацией (10.2), положив в (8.3) $p_{n\pm 1/2} = \frac{1}{2}(p_n + p_{n\pm 1})$. Тогда для узловой точки с номером $n = \overline{1, N-1}$ запишем

$$\begin{aligned} & \frac{p_{n-1} + p_n}{2(x_n - x_{n-1})}(u_n - u_{n-1}) + \frac{p_{n+1} + p_n}{2(x_{n+1} - x_n)}(u_n - u_{n+1}) + \\ & + q_n \frac{x_{n+1} - x_{n-1}}{2} u_n = f_n \frac{x_{n+1} - x_{n-1}}{2}. \quad (10.23) \end{aligned}$$

Для совпадения (10.22) с (10.23) достаточно, чтобы $q(x) \equiv 0$ и $f(x) \equiv \text{const}$ на $[0, 1]$. В этом случае полученные ранее (см. 8.1) оценки погрешности аппроксимации для уравнения (10.23) можно перенести на уравнение (10.22) и заключить, что последнее имеет первый *порядок аппроксимации*. В частном случае равномерного расположения узлов на отрезке $[0, 1]$, т.е. при $x_{n+1} - x_n = x_n - x_{n-1}$, порядок аппроксимации будет вторым.

10.2. Типы конечных элементов

Рассмотренный выше (см. 10.1) тип *конечного элемента* (КЭ) является одним из простейших и пригоден для приближенного решения *методом конечных элементов* (МКЭ) лишь одномерных задач математической физики. Существует достаточно много типов КЭ, в том числе позволяющих решать многомерные задачи. Можно выделить одномерные, двумерные и трехмерные КЭ. Если одномерный КЭ представляет собой отрезок с двумя или более *узлами*, то двумерные могут иметь форму треугольника, четырехугольника (в том числе прямоугольника) и вообще многоугольника с расположением узлов не только в вершинах. Столь же разнообразны по форме и трехмерные КЭ. Применительно к нестационарным (динамическим) задачам могут быть построены КЭ в пространственно-временной области, когда искомое приближенное решение зависит не только от пространственных координат, но и от времени.

При рассмотрении типов КЭ удобно принять, что в отдельно взятом КЭ надо решить задачу приближенного представления некоторой функции по заданным значениям ее *узловых параметров*, т.е. решить задачу *аппроксимации функции* в пределах этого КЭ. Если при построении КЭ в качестве узловых параметров используют лишь значения функции (в случае векторной функции — значения ее базисных функций), то КЭ называют *лагранжевыми*, поскольку в этом случае аппроксимация функции в пределах КЭ аналогична ее представлению *интерполяционным многочленом Лагранжа*. Если же наряду с этим в узлах КЭ используют и значения производных функции, то КЭ называют *эрмитовыми**.

Среди лагранжевых КЭ по виду аппроксимирующих многочленов различают *симплексные* (от латинского слова simplex — простой), *комплексные* (от латинского слова complexus — сочетание) и *мультиплéксные* (от латинского слова

*Ш. Эрмит (1822–1901) — французский математик.

multiplex — сложный). Интерполяционные многочлены, используемые в симплексных КЭ, являются линейными, т.е. первой степени, и содержат постоянное слагаемое и слагаемые, линейно зависящие от всех координат (и времени, если речь идет о КЭ в пространственно-временной области решения задачи). Число коэффициентов в таких многочленах равно числу узлов симплексного КЭ и на единицу больше числа независимых переменных. Одномерным симплексным КЭ является отрезок с узлами на концах (см. 10.1). Двумерным симплексным КЭ будет треугольник (рис. 10.3), а трехмерным — *тетраэдр* (рис. 10.4), причем узлы этих КЭ совпадают с вершинами треугольника и тетраэдра. Таким образом, d -мерный симплексный КЭ имеет $d + 1$ узлов.

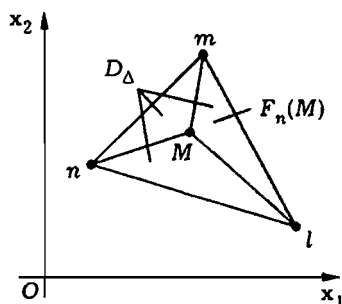


Рис. 10.3

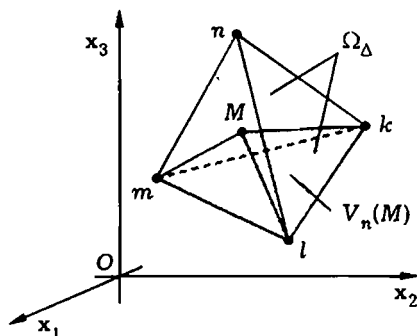


Рис. 10.4

Интерполяционные многочлены комплексных и мультиплексных КЭ имеют степень выше первой, т.е. являются нелинейными. Число узлов таких КЭ более чем на единицу превышает число независимых переменных. В комплексных КЭ используют *полные интерполяционные многочлены* степени $s \geq 2$, содержащие все возможные слагаемые, у которых сумма степеней не превышает s . Интерполяционные многочлены мультиплексных КЭ не являются полными, а границы таких КЭ совпадают с координатными поверхностями.

Симплексные конечные элементы. По-существу, было показано (см. 10.1), что при помощи *функций формы* одномерного симплексного КЭ можно однозначно представить линейную функцию, используя ее значения в двух узлах. Интерполяционный сплайн первой степени двух переменных можно однозначно построить по трем значениям функции в вершинах треугольника [V].

Это означает, что в двумерном симплексном КЭ D_Δ линейная функция $f(M)$ координат $x_1(M)$, $x_2(M)$ точки $M \in D_\Delta$ однозначно определена своими значениями f_l , f_m , f_n в трех его узлах с номерами l , m , n (см. рис. 10.3) при условии, что эти узлы не лежат на одной прямой (в этом случае говорят, что треугольный КЭ невырожденный). Приведем формулу для значения этой функции в точке $M \in D_\Delta$ [V]:

$$f(M) = \frac{F_l(M)}{F} f_l + \frac{F_m(M)}{F} f_m + \frac{F_n(M)}{F} f_n, \quad (10.24)$$

где $F > 0$ — площадь КЭ, а $F_l(M)$, $F_m(M)$, $F_n(M)$ — площади треугольников с вершиной в точке M и основанием, совпадающим со стороной КЭ, противолежащей узлам с номерами l , m , n соответственно. На рис. 10.3 выделен один из таких треугольников с площадью $F_n(M)$. Известно, что $F = \frac{1}{2} |\det A| [V]$, причем элементами первых двух строк квадратной матрицы

$$A = \begin{pmatrix} x_{1l} & x_{1m} & x_{1n} \\ x_{2l} & x_{2m} & x_{2n} \\ 1 & 1 & 1 \end{pmatrix}$$

третьего порядка являются координаты x_1 , x_2 узлов КЭ в прямоугольной системе координат Ox_1x_2 . Отметим, что определитель $\det A$ этой матрицы будет положительным, если КЭ невырожденный и очередность расположения координат его узлов в столбцах матрицы A соответствует обходу их против хода часовой стрелки, т.е. узлы КЭ составляют *правую тройку точек*.

Площади треугольников с вершиной в точке $M \in D$ можно вычислить аналогично. Например, $F_n(M) = \frac{1}{2} |\det A_n(M)|$, где

$$A_n(M) = \begin{pmatrix} x_{1l} & x_{1m} & x_1(M) \\ x_{2l} & x_{2m} & x_2(M) \\ 1 & 1 & 1 \end{pmatrix}.$$

Ясно, что $\frac{F_l(M)}{F} + \frac{F_m(M)}{F} + \frac{F_n(M)}{F} = 1$ для любой точки $M \in D_\Delta$. При совпадении этой точки, например, с узлом n имеем $\frac{F_n(M)}{F} = 1$ и $\frac{F_l(M)}{F} = \frac{F_m(M)}{F} = 0$. Таким образом, рассматривая отношения площадей в (10.24) как функции формы $\varphi_l(M) = \frac{F_l(M)}{F}$, $\varphi_m(M) = \frac{F_m(M)}{F}$ и $\varphi_n(M) = \frac{F_n(M)}{F}$ двумерного симплексного КЭ, запишем

$$f(M) = f_l \varphi_l(M) + f_m \varphi_m(M) + f_n \varphi_n(M), \quad M \in D_\Delta. \quad (10.25)$$

При помощи функций формы можно выразить координаты $x_1(M)$, $x_2(M)$ любой точки $M \in D_\Delta$. Действительно, полагая в (10.25) сначала $f(M) = x_1(M)$, а затем $f(M) = x_2(M)$, получаем

$$\begin{cases} x_1(M) = x_{1l} \varphi_l(M) + x_{1m} \varphi_m(M) + x_{1n} \varphi_n(M), \\ x_2(M) = x_{2l} \varphi_l(M) + x_{2m} \varphi_m(M) + x_{2n} \varphi_n(M). \end{cases} \quad (10.26)$$

Наоборот, каждую из функций формы такого КЭ можно выразить через координаты его узлов и координаты точки $M \in D_\Delta$. Например, для узла с номером n имеем

$$\begin{aligned} \varphi_n(M) = \frac{F_n}{F} = \frac{|\det A_n|}{2F} \begin{vmatrix} x_{1l} & x_{1m} & x_1(M) \\ x_{2l} & x_{2m} & x_2(M) \\ 1 & 1 & 1 \end{vmatrix} = \frac{1}{2F} \left(x_{1l} x_{2m} - \right. \\ \left. - x_{1m} x_{2l} + (x_{2l} - x_{2m}) x_1(M) + (x_{1m} - x_{1l}) x_2(M) \right). \end{aligned} \quad (10.27)$$

Так как узлы с номерами l и m и точка M образуют правую тройку точек, то $\det A_n \geq 0$ и $\varphi_n(M) \geq 0$, $M \in D_\Delta$. На рис. 10.5 представлены графики функций формы двумерного симплексного КЭ, являющиеся наклонными гранями пирамид с высотой 1 и основанием D_Δ .

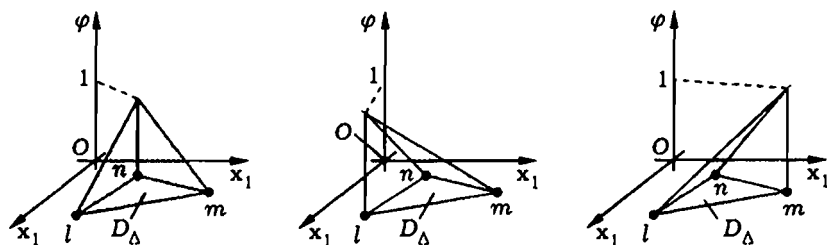


Рис. 10.5

Частные производные функций формы по координатам в пределах двумерного симплексного КЭ сохраняют постоянное значение. Например, из (10.27) находим $\frac{\partial \varphi_n}{\partial x_1} = \frac{x_{2l} - x_{2m}}{2F}$ и $\frac{\partial \varphi_n}{\partial x_2} = \frac{x_{1m} - x_{1l}}{2F}$. При вычислении интегралов от произведения натуральных степеней q, r, s функций формы удобно использовать формулу*

$$\int_F \varphi_l^q(M) \varphi_m^r(M) \varphi_n^s(M) dF(M) = \frac{q!r!s!}{(q+r+s+2)!} 2F. \quad (10.28)$$

С учетом $0! = 1$ (10.28) сохраняет смысл и при нулевых показателях степени.

В пределах двух треугольных КЭ с общей стороной можно точно представить непрерывную кусочно линейную функцию двух переменных, которая в пределах каждого из этих КЭ имеет постоянный градиент. В самом деле, их общая сторона является одномерным симплексным КЭ, так что зна-

*См., например: Норри Д., де Фриз Ж.

чение $f(M)$ этой функции в любой точке M такой стороны однозначно определено значениями функции в общих узлах соседних треугольных КЭ, а график такой функции состоит из двух треугольников с общей стороной (рис. 10.6). Произвольную границу (в том числе с криволинейными участками) плоской области $D^* \subset \mathbb{R}$ можно приближенно заменить ломаной и затем полученный в результате многоугольник заполнить

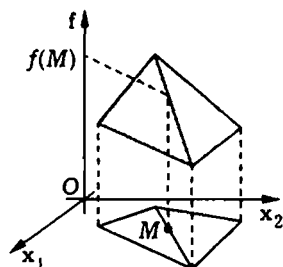


Рис. 10.6

треугольными КЭ, т.е. провести *триангуляцию области* и получить *сетку КЭ*. В данном случае такая сетка КЭ будет плоской. Рассматривая попарно соседние КЭ, приходим к выводу, что сетка КЭ позволяет ограниченную функцию $g(M)$, $M \in D^*$, определенную в области D^* , приближенно заменить непрерывной кусочно линейной функцией, принимающей в каждом узле значение, совпадающее со значением в этом узле функции $g(M)$.

Отметим, что функции формы, удовлетворяющие (10.26) в сочетании с условием $\varphi_l(M) + \varphi_m(M) + \varphi_n(M) = 1$, называют **барицентрическими координатами** точки M относительно вершин треугольника (по гречески *βαρος* — тяжесть). Они были введены в 1827 году немецким математиком А.Ф. Мебиусом (1790–1868) при решении задачи о таком размещении системы масс в вершинах треугольника, при котором данная точка была бы центром масс этой системы. Иногда эти функции формы называют также естественными узловыми координатами точки $M \in \Omega_\Delta$ в двумерном симплексном КЭ.

Перейдем к рассмотрению трехмерного симплексного КЭ Ω_Δ с узлами, имеющими номера k, l, m, n , не лежащими в одной плоскости и являющимися вершинами тетраэдра (см. рис. 10.4). Как и в треугольном КЭ положение точки $M \in \Omega_\Delta$ можно определить ее барицентрическими координатами, но теперь относительно четырех узлов, что приведет к системе линейных

алгебраических уравнений (СЛАУ)

$$\begin{cases} x_{1k}\varphi_k(M) + x_{1l}\varphi_l(M) + x_{1m}\varphi_m(M) + x_{1n}\varphi_n(M) = x_1(M), \\ x_{2k}\varphi_k(M) + x_{2l}\varphi_l(M) + x_{2m}\varphi_m(M) + x_{2n}\varphi_n(M) = x_2(M), \\ x_{3k}\varphi_k(M) + x_{3l}\varphi_l(M) + x_{3m}\varphi_m(M) + x_{3n}\varphi_n(M) = x_3(M), \\ \varphi_k(M) + \varphi_l(M) + \varphi_m(M) + \varphi_n(M) = 1 \end{cases} \quad (10.29)$$

относительно функций формы $\varphi(M)$ с соответствующими индексами. В первых трех строках квадратной матрицы

$$B = \begin{pmatrix} x_{1k} & x_{1l} & x_{1m} & x_{1n} \\ x_{2k} & x_{2l} & x_{2m} & x_{2n} \\ x_{3k} & x_{3l} & x_{3m} & x_{3n} \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

четвертого порядка этой СЛАУ стоят координаты x_1, x_2, x_3 узлов в прямоугольной системе координат $Ox_1x_2x_3$. Функции формы будут однозначно связаны с этими координатами, если матрица B невырожденная, т.е. $\det B \neq 0$.

Используя свойства определителей [III], из первых трех столбцов $\det B$ вычитаем четвертый столбец и получаем

$$\begin{vmatrix} x_{1k} & x_{1l} & x_{1m} & x_{1n} \\ x_{2k} & x_{2l} & x_{2m} & x_{2n} \\ x_{3k} & x_{3l} & x_{3m} & x_{3n} \\ 1 & 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} x_{1k} - x_{1n} & x_{1l} - x_{1n} & x_{1m} - x_{1n} & x_{1n} \\ x_{2k} - x_{2n} & x_{2l} - x_{2n} & x_{2m} - x_{2n} & x_{2n} \\ x_{3k} - x_{3n} & x_{3l} - x_{3n} & x_{3m} - x_{3n} & x_{3n} \\ 0 & 0 & 0 & 1 \end{vmatrix} =$$

$$= \begin{vmatrix} x_{1k} - x_{1n} & x_{1l} - x_{1n} & x_{1m} - x_{1n} \\ x_{2k} - x_{2n} & x_{2l} - x_{2n} & x_{2m} - x_{2n} \\ x_{3k} - x_{3n} & x_{3l} - x_{3n} & x_{3m} - x_{3n} \end{vmatrix}.$$

Определитель в правой части этого выражения равен смешанному произведению трех векторов, направленных вдоль ребер тетраэдра, выходящих из одной вершины (в данном случае, из

узла с номером n). Следовательно, $\det A_3 \neq 0$, если эти векторы некомпланарны [III], т.е. вершины тетраэдра не лежат в одной плоскости, а его объем V отличен от нуля. В этом случае говорят, что трехмерный симплексный КЭ невырожденный. Отметим, что $|\det B| = 6V$, причем $\det B = 6V$, если при наблюдении из n -го узла обход узлов с номерами k, l и m осуществляется по ходу часовой стрелки (см. рис. 10.4).

Согласно правилу Крамера, из (10.29) находим, например, для n -го узла

$$\varphi_n(M) = \frac{1}{6V} \begin{vmatrix} x_{1k} & x_{1l} & x_{1m} & x_1(M) \\ x_{2k} & x_{2l} & x_{2m} & x_2(M) \\ x_{3k} & x_{3l} & x_{3m} & x_3(M) \\ 1 & 1 & 1 & 1 \end{vmatrix} = \frac{V_n(M)}{V}, \quad (10.30)$$

поскольку определитель в (10.30) равен $6V_n(M)$, где $V_n(M)$ — объем тетраэдра с вершинами в точке M и в узлах с номерами k, l и m (см. рис. 10.4). Аналогично можно найти выражения для $\varphi_k(M)$, $\varphi_l(M)$ и $\varphi_m(M)$. Таким образом, барицентрическими координатами точки M относительно узлов трехмерного симплексного КЭ являются отношения объемов соответствующих тетраэдров.

Производные функций формы в пределах трехмерного симплексного КЭ также сохраняют постоянное значение. Так, например, из (10.30) с учетом изменения знака определителя при перестановке строк следует, что

$$\frac{\partial \varphi_n}{\partial x_1} = -\frac{1}{6V} \begin{vmatrix} x_{2k}x_{2l}x_{2m} \\ x_{3k}x_{3l}x_{3m} \\ 1 & 1 & 1 \end{vmatrix} = -\frac{1}{6V} \begin{vmatrix} 1 & 1 & 1 \\ x_{2k}x_{2l}x_{2m} \\ x_{3k}x_{3l}x_{3m} \end{vmatrix},$$

$$\frac{\partial \varphi_n}{\partial x_2} = -\frac{1}{6V} \begin{vmatrix} x_{1k}x_{1l}x_{1m} \\ 1 & 1 & 1 \\ x_{3k}x_{3l}x_{3m} \end{vmatrix}, \quad \frac{\partial \varphi_n}{\partial x_3} = -\frac{1}{6V} \begin{vmatrix} x_{1k}x_{1l}x_{1m} \\ x_{2k}x_{2l}x_{2m} \\ 1 & 1 & 1 \end{vmatrix}.$$

При вычислении интегралов от произведения натуральных степеней p, q, r, s функций формы справедлива формула*

$$\int_V \varphi_k^p(M) \varphi_l^q(M) \varphi_m^r(M) \varphi_n^s(M) dV(M) = \frac{p!q!r!s!}{(p+q+r+s+3)!} 6V. \quad (10.31)$$

С учетом $0! = 1$ (10.31) сохраняет смысл и в случае нулевых показателей степени.

Два соседних трехмерных симплексных КЭ с общей гранью позволяют точно представить непрерывную кусочно линейную функцию трех переменных, которая в пределах каждого из этих КЭ имеет постоянный градиент. Действительно, их общая треугольная грань является двумерным симплексным КЭ, так что значение этой функции в любой точке такой грани однозначно определено значениями функции в общих узлах соседних трехмерных КЭ. Аналогично плоской области произвольную пространственную область Ω^* можно приближенно заменить совокупностью таких КЭ, образующих пространственную сетку. Рассматривая попарно соседние КЭ, приходим к выводу, что эта сетка дает возможность ограничить функцию $h(M)$, $M \in \Omega^*$, определенную в области Ω^* , приближенно заменить непрерывной кусочно линейной функцией, принимающей в каждом узле значение, совпадающее со значением в этом узле функции $h(M)$.

Итак, при линейной аппроксимации на d -мерном симплексном КЭ Ω с числом узлов $d+1$ действительной функции $u(M)$, $M \in \Omega$, можно записать

$$u(M) \approx u_1 \varphi_1(M) + \dots + u_{d+1} \varphi_{d+1}(M) = U_e^T \Phi_e(M), \quad (10.32)$$

где U_e и $\Phi_e(M)$ — матрицы-столбцы размера $(d+1) \times 1$, элементами которых являются значения u_n этой функции в узлах

*См., например: Норри Д., де Фриз Ж.

и функции формы $\varphi_n(M)$, $M \in \Omega$, $n = \overline{1, d+1}$, этого КЭ. Однако линейная аппроксимация функции в пределах каждого симплексного КЭ может оказаться довольно грубой, особенно в случае больших по абсолютному значению градиентов этой функции, а повышение точности потребует использования сетки КЭ с весьма большим числом симплексных элементов. Общее число КЭ можно уменьшить, если в пределах каждого элемента повысить степень аппроксимирующего многочлена, т.е. перейти к комплексным или мультиплексным КЭ.

Комплексные конечные элементы. Одномерный комплексный лагранжев КЭ — это отрезок $X \subset \mathbb{R}$ с $N > 2$ узлами в точках x_n , $n = \overline{1, N}$. В пределах такого КЭ функцию $u(x)$ можно аппроксимировать многочленом степени $s = N - 1$:

$$u(x) \approx \sum_{n=1}^N u_n L_n(x), \quad (10.33)$$

где u_n — узловые значения этой функции, а

$$L_n(x) = \prod_{\substack{m=1, N \\ m \neq n}} \frac{x - x_m}{x_n - x_m} \quad —$$

интерполяционный многочлен степени $s = N - 1$, который и является функцией формы этого КЭ, соответствующей узлу x_n , так как $L_n(x_n) = 1$, $n = \overline{1, N}$, и $L_n(x_m) = 0$ при $x_m \neq x_n$, а

$$\sum_{n=1}^N L_n(x) = 1. \quad (10.34)$$

Последнее равенство следует из свойства правой части соотношения (10.33) точно представлять все многочлены до степени $s = N - 1$ включительно [II]. Функция $u(x) \equiv 1$ является многочленом нулевой степени и для нее $u_n = u(x_n) = 1$, $n = \overline{1, N}$, что доказывает справедливость (10.34).

Если на одном или обоих концах отрезка X узел отсутствует, то такой одномерный КЭ называют **сингулярным**. В этом случае (10.33) экстраполирует функцию $u(x)$ в окрестности конца отрезка, на котором отсутствует узел.

Полный многочлен двух переменных, имеющий степень s , включает $\frac{(s+1)(s+2)}{2}$ коэффициентов [VII]. Поэтому для построения двумерного комплексного лагранжева КЭ, функции формы $\varphi_n^{(s)}$ которого являются многочленами степени s , необходимо располагать значениями аппроксимируемой функции в $N = \frac{(s+1)(s+2)}{2}$ узлах. При $s = 2$ имеем $N = 6$. Если выбрать треугольный КЭ, то при его произвольном расположении в плоскости любая из сторон может оказаться параллельной одной из координатных осей, например Ox_1 , т.е. быть отрезком прямой $x_2 = \text{const}$. Вдоль этой стороны многочлен будет квадратным трехчленом по x_1 , и для однозначного определения его коэффициентов необходимо три узловых значения аппроксимируемой функции. Таким образом, на каждой стороне необходимо

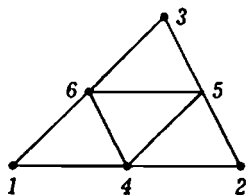


Рис. 10.7

иметь по три узла. Эти узлы можно расположить в вершинах и серединах сторон треугольника (рис. 10.7). Тогда функции формы $\varphi_n^{(2)}$, $n = \overline{1, 6}$, такого комплексного КЭ удастся выразить через три линейные функции формы $\varphi_1, \varphi_2, \varphi_3$ симплексного КЭ.

Так, в узле 4, находящемся в середине стороны треугольника, имеем $\varphi_1 = \varphi_2 = 1/2$. Если положить $\varphi_4^{(2)} = 4\varphi_1\varphi_2$, то такая функция формы будет многочленом второй степени, принимающим значение 1 в „своем“ узле и 0 во всех остальных. Функция формы $\varphi_1^{(2)}$ также должна быть произведением двух линейных сомножителей. Если одним из них будет φ_1 , то произведение обратится в нуль во всех точках стороны, противоположной узлу 1 (в частности, в узлах 2, 3 и 5). Чтобы произведение обращалось в нуль в узлах 4 и 6, оно должно содержать сомно-

житель $\varphi_1 - 1/2$. Но произведение $\varphi_1(\varphi_1 - 1/2)$ в узле 1 равно лишь $1/2$, т.е. его необходимо нормировать, умножив на два. В итоге получаем $\varphi_1^{(2)} = \varphi_1(2\varphi_1 - 1)$. Теперь можно записать

$$\varphi_1^{(2)} = \varphi_1(2\varphi_1 - 1), \quad \varphi_4^{(2)} = 4\varphi_1\varphi_2, \quad \varphi_2^{(2)} = \varphi_2(2\varphi_2 - 1) \quad (10.35)$$

и т.д. с учетом перестановки нижних индексов, указывающих номера узлов. Несложно проверить, что сумма всех функций формы $\varphi_n^{(2)}$, $n = \overline{1, 6}$, равна единице.

При $s = 3$ для построения полного многочлена необходимо $N = 10$ узлов. Если снова выбрать треугольный КЭ, то при его произвольном расположении на каждой стороне должно быть по четыре узла. Действительно, любая из сторон может оказаться параллельной одной из координатных осей, например Ox_1 . В этом случае многочлен будет кубичным по x_1 , и для однозначного определения его коэффициентов необходимо знать четыре значения аппроксимируемой функции. Если поместить узлы в вершины и каждую из сторон разделить узлами на три равные части, то останется десятый узел, который будет внутренним. Для упрощения построения функций формы его целесообразно поместить в точку пересечения медиан треугольника (рис. 10.8). Функции формы $\varphi_n^{(3)}$, $n = \overline{1, 10}$, этого комплексного КЭ также удастся выразить через три линейные функции формы $\varphi_1, \varphi_2, \varphi_3$ симплексного КЭ.

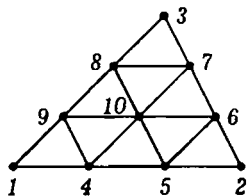


Рис. 10.8

В узле 10 $\varphi_1 = \varphi_2 = \varphi_3 = 1/3$. Положим $\varphi_{10}^{(3)} = 27\varphi_1\varphi_2\varphi_3$ и убедимся, что такая функция формы, являясь многочленом третьей степени, принимает значение 1 в „своем“ узле и 0 во всех остальных. Функции формы $\varphi_n^{(3)}$ остальных узлов с номерами $n = \overline{1, 9}$ также должны быть произведениями трех линейных сомножителей. Так, чтобы функция формы $\varphi_1^{(3)}$ обращалась в нуль в узлах 2, 3, 6 и 7, она должна содержать

сомножитель φ_1 , а сомножители $\varphi_1 - 1/3$ и $\varphi_1 - 2/3$ обеспечивают обращение в нуль в узлах 5, 8, 10 и 4, 9 соответственно. Однако произведение $\varphi_1(\varphi_1 - 1/3)(\varphi_1 - 2/3)$ в узле 1 равно $2/9$. Умножая его на $9/2$, получаем $\varphi_1^{(2)} = \frac{9}{2}\varphi_1(3\varphi_1 - 1)(3\varphi_1 - 2)$.

По аналогии с предыдущим КЭ можно заключить, что функция формы узла, принадлежащего только одной стороне треугольника, должна содержать произведение линейных функций формы прилегающих вершин. Так, для узла 4 наличие в его функции формы $\varphi_4^{(3)}$ сомножителя в виде произведения $\varphi_1\varphi_2$ обеспечивает ее обращение в нуль во всех „чужих“ узлах, кроме узлов 5 и 10. Обращение в нуль в этих двух узлах будет обеспечено, если третьим сомножителем будет $\varphi_1 - 1/3$. Но произведение $\varphi_1\varphi_2(\varphi_1 - 1/3)$ в узле 4 равно $2/27$. Умножив его на $27/2$, получим $\varphi_4^{(3)} = \frac{9}{2}\varphi_1\varphi_2(3\varphi_1 - 1)$. В итоге запишем

$$\begin{cases} \varphi_{10}^{(3)} = 27\varphi_1\varphi_4\varphi_7, & \varphi_1^{(3)} = \frac{1}{2}\varphi_1(3\varphi_1 - 1)(3\varphi_1 - 2), \\ \varphi_4^{(3)} = \frac{9}{2}\varphi_1\varphi_2(3\varphi_1 - 1), & \varphi_5^{(3)} = \frac{9}{2}\varphi_1\varphi_2(3\varphi_2 - 1), \\ \varphi_2^{(3)} = \frac{1}{2}\varphi_2(3\varphi_2 - 1)(3\varphi_2 - 2) \end{cases} \quad (10.36)$$

и т.д. также с учетом перестановки нижних индексов, указывающих номера узлов. Непосредственной проверкой можно убедиться, что сумма всех функций формы $\varphi_n^{(3)}$, $n = \overline{1, 10}$, снова равна единице.

Можно показать*, что треугольный комплексный элемент удастся построить и в случае $s > 3$, а четырехугольный — при $s > 4$. Полный многочлен степени s от трех координат x_1, x_2, x_3 имеет $\frac{(s+1)(s+2)(s+3)}{2}$ коэффициентов. Аналогично двумерному случаю можно установить, что при $s \leq 4$ такой многочлен удастся построить лишь для комплексных элементов в форме тетраэдра.

*См., например Зарубин В.С., Селиванов В.В.

Если при $s = 2$ каждую грань тетраэдра выбрать в виде треугольника, изображенного на рис. 10.7, то $N = 10$, а функции формы $\varphi^{(2)}$ с соответствующими индексами следуют из (10.35). Если же при $s = 3$ каждая грань тетраэдра имеет вид, показанный на рис. 10.8, то $N = 20$ и функции формы $\varphi^{(3)}$ с соответствующими индексами можно найти из (10.36).

Отметим, что представление функций формы $\varphi_n^{(s)}(M)$ комплексных КЭ в виде (10.35) или (10.36) позволяет при вычислении интегралов использовать (10.28) или (10.31), а производные этих функций по координатам выразить через производные функций формы симплексных КЭ. При аппроксимации действительной функции $u(M)$, $M \in \Omega$, в пределах комплексного КЭ Ω с числом узлов N остается в силе (10.32), но теперь размер матриц-столбцов U_e и Φ_e будет $N \times 1$, а их элементами — узловые значения u_n этой функции и функции формы $\varphi_n^{(s)}(M)$, $n = \overline{1, N}$, КЭ соответственно.

Мультиплексные конечные элементы. Мультиплексный лагранжев КЭ в прямоугольных декартовых координатах x_i , $i = \overline{1, d}$, имеет вид прямоугольника ($d = 2$) или прямоугольного параллелепипеда ($d = 3$). На каждой из параллельных координатной оси Ox_i сторон (или ребер) таких КЭ число узлов N_i и их расположение одинаково, но может отличаться от числа и расположения узлов на сторонах (или ребрах), параллельных другой оси (или другим осям). Наибольшее возможное число узлов в прямоугольнике (с учетом внутренних) равно $N_1 N_2$, а в параллелепипеде (с учетом внутренних и на его гранях) — $N_1 N_2 N_3$. Функция формы для n -го узла в точке M_n с координатами x_{in} является произведением интерполяционных многочленов по каждой из координат:

$$\varphi_n(M) = \prod_{i=1}^d L_n(x_i), \quad L_n(x_i) = \prod_{\substack{m=1, N_i \\ x_{im} \neq x_{in}}} \frac{x_i - x_{im}}{x_{in} - x_{im}}, \quad (10.37)$$

где M — точка с координатами x_i , $i = \overline{1, d}$, принадлежащая мультиплексному КЭ, а x_{im} — координаты узлов на стороне, параллельной оси Ox_i . При $x_i = x_{in}$, $i = \overline{1, d}$, в n -м узле $L_n(x_{in}) = 1$ и $\varphi_n(M_n) = 1$, а во всех остальных узлах $\varphi_n(M) = 0$.

Многочлены в (10.37) имеют в общем случае по каждой из координат x_i , $i = \overline{1, d}$, различную степень $s_i = N_i - 1$ и поэтому не являются полными. Вместе с тем многочлен, аппроксимирующий функцию $u(M)$ вдоль каждой стороны (или ребра) КЭ, является полным по соответствующей координате, обеспечивая непрерывность аппроксимации на границе между соседними элементами. В связи с этим мультиплексный КЭ с различными значениям s_i удобен для согласования между собой симплексных КЭ (или комплексных КЭ с низкой степенью полного интерполяционного многочлена) в области малых градиентов искомого решения и комплексных КЭ с высокой степенью многочлена в области, где ожидаются большие градиенты искомого решения.

Если на всех сторонах (или ребрах) мультиплексного КЭ число N_i узлов одинаково, то для прямоугольника или прямоугольного параллелепипеда при $s_i = N_i - 1 \leq 4$ многочлен остается неполным. Например, в прямоугольнике с узлами в вершинах (рис. 10.9) можно построить билинейный многочлен $a_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2$ с четырьмя коэффициентами. Эти коэффициенты однозначно определены узловыми значениями аппроксимируемой функции $[V]$, а на сетке таких прямоуголь-

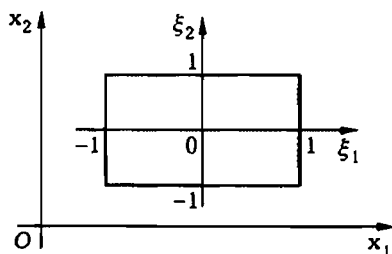


Рис. 10.9

ников билинейные многочлены представляют непрерывную кусочно линейную функцию.

Билинейный многочлен отличается от полного многочлена степени $s = 2$ отсутствием слагаемых, содержащих x_1^2 и x_2^2 . Если ввести дополнительные узлы в серединах сторон прямоугольника и не вводить внутренних узлов, то можно построить многочлен

$$a_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2 + c_1x_1^2x_2 + c_2x_1x_2^2 \quad (10.38)$$

с восемью коэффициентами. В отличие от полного многочлена степени $s = 3$ в нем отсутствуют слагаемые с x_1^3 и x_2^3 . Аналогичная ситуация возникает и для трехмерных мультиплексных КЭ в форме прямоугольного параллелепипеда.

В середину прямоугольного КЭ поместим начало прямоугольной системы координат $O\xi_1\xi_2$, оси которой параллельны его ребрам (см. рис. 10.9). Тогда функцию формы для любого n -го узла, находящегося в вершине прямоугольника, можно записать в виде

$$\varphi_n(M) = \frac{1}{4} \prod_{i=1}^2 (1 + \xi_{in}\xi_i), \quad n = \overline{1, 4}, \quad (10.39)$$

где ξ_i , $i = 1, 2$, — координаты принадлежащей КЭ точки M , а ξ_{in} — координаты n -го узла, принимающие значения ± 1 . В случае прямоугольника с дополнительными узлами в серединах сторон для узлов в вершинах получим

$$\varphi_n(M) = \frac{1}{4} (\xi_{1n}\xi_1 + \xi_{2n}\xi_2 - 1) \prod_{i=1}^2 (1 + \xi_{in}\xi_i), \quad n = \overline{1, 4}, \quad (10.40)$$

а для дополнительных узлов с координатами ξ_{im} , одна из которых равна нулю, —

$$\varphi_m(M) = \frac{1 - \xi_{2m}^2\xi_1^2 - \xi_{1m}^2\xi_2^2}{2} \prod_{i=1}^2 (1 + \xi_{im}\xi_i), \quad m = \overline{1, 4}. \quad (10.41)$$

Когда на каждой стороне прямоугольника равномерно расположено по четыре узла, имеем

$$\varphi_k(M) = \frac{9}{32}(1 + \xi_{2k}\xi_2)(1 - \xi_1^2)(1 + 9\xi_{1k}\xi_1), \quad \xi_{1k} = \pm 1, \quad \xi_{2k} = \pm \frac{1}{3};$$

$$\varphi_l(M) = \frac{9}{32}(1 + \xi_{1l}\xi_1)(1 - \xi_2^2)(1 + 9\xi_{2l}\xi_2), \quad \xi_{1l} = \pm \frac{1}{3}, \quad \xi_{2l} = \pm 1,$$

$$\varphi_n(M) = \frac{9(\xi_1^2 + \xi_2^2) - 10}{32} \prod_{i=1}^2 (1 + \xi_{in}\xi_i), \quad \xi_{in} = \pm 1.$$

В прямоугольной системе координат $O\xi_1\xi_2\xi_3$ с началом в середине КЭ в виде прямоугольного параллелепипеда (рис. 10.10), имеющего узлы только в вершинах с координатами $\xi_{in} = \pm 1$, $i = 1, 2, 3$, $n = \overline{1, 8}$, для функций формы получим

$$\varphi_n(M) = \frac{1}{8} \prod_{i=1}^3 (1 + \xi_{in}\xi_i), \quad n = \overline{1, 8}, \quad (10.42)$$

где ξ_i — координаты точки M , принадлежащей этому КЭ. При наличии дополнительных узлов в серединах ребер с координатами

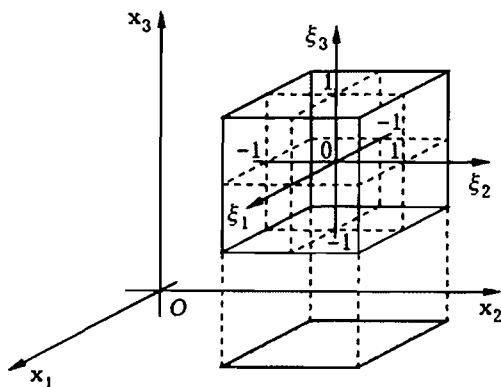


Рис. 10.10

тами ξ_{im} , $m = \overline{1, 12}$, одна из которых равна нулю, имеем

$$\varphi_m(M) = \frac{1 - (\xi_{1m}\xi_{2k}\xi_3)^2 - (\xi_{2m}\xi_{3m}\xi_1)^2 - (\xi_{3m}\xi_{1k}\xi_2)^2}{4} \prod_{i=1}^3 (1 + \xi_{im}\xi_i),$$

$$\varphi_n(M) = \frac{\xi_{1n}\xi_1 + \xi_{2n}\xi_2 + \xi_{3n}\xi_3 - 2}{8} \prod_{i=1}^3 (1 + \xi_{in}\xi_i)$$

для дополнительных узлов и узлов в вершинах соответственно. Когда на каждом ребре параллелепипеда равномерно расположено по четыре узла, для узлов в вершинах при $\xi_{in} = \pm 1$ получим

$$\varphi_n(M) = \frac{1}{64} \left(9 \sum_{i=1}^3 \xi_i^2 - 19 \right) \prod_{i=1}^3 (1 + \xi_{in}\xi_i), \quad n = \overline{1, 8}, \quad (10.43)$$

а для узлов при $\xi_{1k} = \pm 1/3$, $\xi_{2k} = \xi_{3k} = \pm 1$, $k = \overline{1, 8}$, —

$$\varphi_k(M) = \frac{9}{64} (1 - \xi_1^2) (1 + 9\xi_{1k}\xi_1) (1 + \xi_{2k}\xi_2) (1 + \xi_{3k}\xi_3). \quad (10.44)$$

Для шестнадцати промежуточных узлов на остальных восьми ребрах следует в (10.44) соответствующим образом изменить индексы.

Мультиплексные КЭ можно построить не только в прямоугольных, но и в любых ортогональных системах координат, в том числе в полярных (или цилиндрических) и сферических координатах. Для трехмерных задач применимы лагранжевы КЭ в виде треугольных прямых призм (рис. 10.11), причем в плоскости, параллельной основанию призмы, аппроксимирующий многочлен является полным (линейным или более высокой

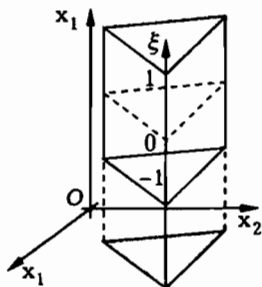


Рис. 10.11

степени) по двум координатам. После умножения этого многочлена на многочлен по координате ξ , отсчитываемой вдоль оси $O\xi$, перпендикулярной основанию, получают неполный многочлен по трем координатам. Однако КЭ, у которого хотя бы одна грань непараллельна координатным плоскостям, строго говоря, нельзя отнести ни к симплексным (или комплексным), ни к мультиплексным.

Например, для прямой треугольной призмы с узлами только в вершинах получим

$$\tilde{\varphi}_n(M) = \frac{1 + \xi_n \xi}{2} \varphi_n(M), \quad n = 1, 2, 3, \quad (10.45)$$

где ξ — координата точки M , перпендикулярная основанию, $\xi_n = \pm 1$ — координата основания, содержащего n -й узел, а $\varphi_n(M)$ — функция формы n -го узла двумерного симплексного КЭ, совпадающего с основанием призмы. При наличии дополнительных узлов в серединах всех ребер призмы имеем

$$\begin{cases} \tilde{\varphi}_n(M) = \frac{\xi^2 + (2\varphi_n(M) - 1)(1 + \xi_n \xi) - 1}{2} \varphi_n(M), & n = \overline{1, 3} \\ \tilde{\varphi}_m(M) = (1 - \xi^2) \varphi_n(M), & m = \overline{1, 3}, \\ \tilde{\varphi}_k(M) = 2(1 + \xi_k \xi) \varphi^*(M) \varphi_*(M), & \xi_k = \pm 1, \quad k = \overline{1, 3}, \end{cases} \quad (10.46)$$

для узлов в вершинах, серединах боковых ребер и сторон основания соответственно, причем $\varphi_n(M)$ — функция формы двумерного симплексного КЭ в сечении призмы, перпендикулярном боковому ребру в узле с номером n или m , а $\varphi^*(M)$ и $\varphi_*(M)$ — функции формы этого КЭ для узлов на концах стороны основания, содержащей узел с номером k .

Аналогично можно построить функции формы для КЭ в цилиндрических координатах r, φ, z , когда сечение области решаемой задачи, проходящее через ось Oz , разбито на двумерные симплексные КЭ, а ξ выполняет роль полярного угла φ .

Как и в случае комплексных КЭ, при аппроксимации действительной функции $u(M)$, $M \in \Omega$, в пределах мультиплексного КЭ Ω с общим числом узлов N справедливо приближение (10.32), причем элементами матриц-столбцов U_e и Φ_e размера $N \times 1$ являются узловые значения u_n этой функции и функции формы $\varphi_n(M)$, $n = \overline{1, N}$, мультиплексного КЭ соответственно.

Все рассмотренные типы КЭ обеспечивают непрерывную аппроксимацию искомой функции как внутри элемента, так и на границах между однотипными соседними элементами. Но может возникнуть необходимость обеспечить при переходе через границу между соседними КЭ непрерывную аппроксимацию не только функции, но и всех ее производных первого порядка. Такую **аппроксимацию** называют **гладкой**. Это требование можно выполнить, если при построении КЭ использовать в качестве *узловых параметров* не только значения функции, но и значения ее производных, т.е. перейти от лагранжевых к эрмитовым КЭ.

Эрмитовы конечные элементы. Простейшим одномерным эрмитовым КЭ является отрезок $[x_{n-1}, x_n]$ с узлами на концах и узловыми значениями u_{n-1} , u'_{n-1} , u_n , u'_n функции $u(x)$ и ее производной $u'(x)$. Эти значения единственным образом задают *кубический интерполяционный многочлен Эрмита*, аппроксимирующий в пределах этого КЭ дифференцируемую на концах отрезка функцию $u(x)$:

$$\begin{aligned} u(x) \approx & u_{n-1}(x_n - x)^2 \frac{2(x - x_{n-1}) + h}{h^3} + \\ & + u'_{n-1}(x_n - x)^2 \frac{x - x_{n-1}}{h^2} + u_n(x - x_{n-1})^2 \frac{2(x_n - x)^2 + h}{h^3} + \\ & + u'_n(x - x_{n-1})^2 \frac{x - x_{n-1}}{h^2}, \quad (10.47) \end{aligned}$$

где $h = x_n - x_{n-1}$. Для четырежды непрерывно дифференцируемой в интервале (x_{n-1}, x_n) функции $u(x)$ наибольшая возможная погрешность в (10.47) пропорциональна h^4 [II].

Если обозначить $\xi = \frac{x_n - x}{h}$, $\psi(\xi) = (1 - \xi)^2(1 + 2\xi)$ и $\omega(\xi) = \xi(1 - \xi)^2$, то вместо (10.47) аналогично (10.32) можно написать

$$u(M) \approx U_e^T \Phi_e(M), \quad U_e^T = (u_{n-1}, u'_{n-1}, u_n, u'_n), \quad (10.48)$$

где $\Phi_e(M) = (\psi(1-\xi) \quad h\omega(1-\xi) \quad \psi(\xi) \quad h\omega(\xi))^T$ — матрица-столбец размера 4×1 , элементами которой являются функции формы рассматриваемого одномерного эрмитового КЭ, а M — точка с координатой $x \in [x_{n-1}, x_n]$.

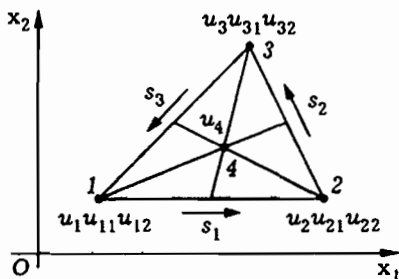


Рис. 10.12

Примером двумерного эрмитова КЭ с полным кубическим многочленом является треугольник с узлами в вершинах и четвертым узлом в точке пересечения медиан (рис. 10.12), в котором в качестве узловых параметра принимают значение u_4 аппроксимируемой функции $u(x_1, x_2)$, дифференцируемой в трех остальных узлах. Как и в случае комплексного КЭ функция формы $\tilde{\varphi}_4^{(3)} = 27\varphi_1\varphi_2\varphi_3$ для этого узла обращается в нуль на всех сторонах треугольника, поскольку на любой стороне одна из функций формы φ_1 , φ_2 или φ_3 двумерного симплексного КЭ равна нулю, а в „своем“ узле равна единице, поскольку $\varphi_1 = \varphi_2 = \varphi_3 = 1/3$. В вершинах треугольника узловыми параметрами являются значения u_n , $n = 1, 2, 3$, функции $u(x_1, x_2)$ и ее частных производных u_{n1} , u_{n2} первого порядка по координатам x_1 и x_2 соответственно. Итак, общее число узловых параметров

равно 10, что необходимо для однозначного построения полного кубического многочлена двух переменных. Функции формы (например, для узла с номером $n = 1$) можно записать в виде*

$$\begin{cases} \tilde{\varphi}_1^{(3)} = \varphi_1^2(\varphi_1 + 3\varphi_2 + \varphi_3) - 7\varphi_1\varphi_2\varphi_3, \\ \tilde{\varphi}_{11}^{(3)} = \varphi_1^2(b_{23}\varphi_2 - b_{22}\varphi_3) + (b_{22} - b_{23})\varphi_1\varphi_2\varphi_3, \\ \tilde{\varphi}_{12}^{(3)} = \varphi_1^2(b_{12}\varphi_3 - b_{13}\varphi_2) + (b_{13} - b_{12})\varphi_1\varphi_2\varphi_3, \end{cases} \quad (10.49)$$

где $b_{12} = (x_2)_3 - (x_2)_1$, $b_{13} = (x_2)_1 - (x_2)_2$, $b_{22} = (x_1)_1 - (x_1)_3$, $b_{23} = (x_1)_2 - (x_1)_1$, причем $(x_1)_i$ и $(x_2)_i$, $i = \overline{1, 3}$, — координаты узлов в вершинах треугольника. Если номера $n = 1, 2, 3$ узлов возрастают при движении против хода часовой стрелки, то для узлов с номерами 2 и 3 в (10.49) следует провести циклическую перестановку нижних индексов. Положим

$$\Phi_e(M) = (\tilde{\varphi}_1^{(3)} \quad \tilde{\varphi}_{11}^{(3)} \quad \tilde{\varphi}_{12}^{(3)} \quad \tilde{\varphi}_2^{(3)} \quad \tilde{\varphi}_{21}^{(3)} \quad \tilde{\varphi}_{22}^{(3)} \quad \tilde{\varphi}_3^{(3)} \quad \tilde{\varphi}_{31}^{(3)} \quad \tilde{\varphi}_{32}^{(3)} \quad \tilde{\varphi}_4^{(3)})^T,$$

$$U^{(e)} = (u_1 \quad u_{11} \quad u_{12} \quad u_2 \quad u_{21} \quad u_{22} \quad u_3 \quad u_{31} \quad u_{32} \quad u_4)^T.$$

Тогда для приближенного значения функции $u(M)$ в точке M , принадлежащей рассматриваемому КЭ, аналогично (10.48) получим $u(M) \approx U_e^T \Phi(M)_e$.

Изменение функции u вдоль каждой стороны такого КЭ аппроксимирует кубический интерполяционный многочлен Эрмита по координате s_n , направленной вдоль этой стороны и отсчитываемой от узла с номером $n = 1, 2, 3$ (см. рис. 10.12). Коэффициенты этого многочлена можно однозначно выразить через узловые значения функции u и ее производной $\frac{\partial u}{\partial s_n}$ в вершинах на концах каждой из сторон КЭ. Производная $\frac{\partial u}{\partial s_n}$ в каждой вершине с номером n является линейной комбинацией значений u_{n1} и u_{n2} , одинаковой для соседних КЭ с двумя

*См., например: Зарубин В.С., Селиванов В.В.

общими вершинами. Поэтому кубические многочлены, построенные в соседних КЭ, совпадают на стороне между этими вершинами. Это обеспечивает непрерывность аппроксимации функции u и ее производной $\frac{\partial u}{\partial s_n}$ при переходе через границу между КЭ. Однако непрерывность аппроксимации производной $\frac{\partial u}{\partial \nu}$ в направлении нормали ν к границе между соседними КЭ в общем случае отсутствует, поскольку изменение $\frac{\partial u}{\partial \nu}$ вдоль этой границы описывает многочлен, квадратичный по s_n , и его коэффициенты нельзя однозначно найти по двум значениям $\frac{\partial u}{\partial \nu}$ в вершинах на концах этой стороны. Таким образом, рассматриваемый КЭ не обеспечивает гладкой аппроксимации функции u .

Треугольный эрмитов КЭ с гладкой аппроксимацией функции $u(x_1, x_2)$ можно построить на основе полного многочлена пятой степени, содержащего 21 коэффициент. В качестве узловых параметров в вершинах с номерами $n = 1, 2, 3$ принимают по шесть значений $u_n, u_{n1}, u_{n2}, u_{n11} = \left(\frac{\partial^2 u}{\partial x_1^2}\right)_n, u_{n22} = \left(\frac{\partial^2 u}{\partial x_2^2}\right)_n$ и $u_{n12} = \left(\frac{\partial^2 u}{\partial x_1 \partial x_2}\right)_n$, а в серединах сторон — значения $u_{m\nu} = \left(\frac{\partial u}{\partial \nu}\right)_m, m = 4, 5, 6$, производных в направлении нормали к каждой из сторон (рис. 10.13), т.е. всего 21 параметр.

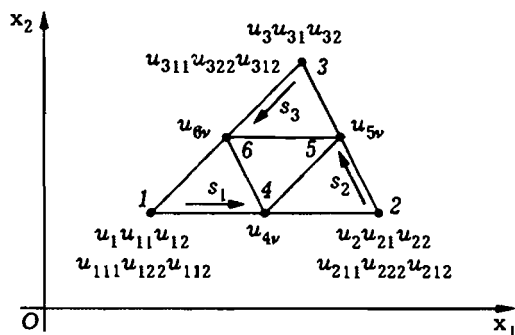


Рис. 10.13

Изменение $u(s_n)$ вдоль любой стороны такого КЭ описывает многочлен пятой степени по координате s_n . Его коэффициенты однозначно определены значениями u , $\frac{\partial u}{\partial s_n}$ и $\frac{\partial^2 u}{\partial s_n^2}$ в двух вершинах на концах этой стороны. Значения $\frac{\partial u}{\partial s_n}$ и $\frac{\partial^2 u}{\partial s_n^2}$ в каждой вершине являются линейными комбинациями узловых параметров (u_{n1} , u_{n2} и u_{n11} , u_{n22} , u_{n12} соответственно), одинаковыми для КЭ с общими вершинами. Поэтому многочлены пятой степени, построенные в соседних КЭ, совпадают на стороне между этими вершинами, что обеспечивает непрерывность аппроксимации u , $\frac{\partial u}{\partial s_n}$ и $\frac{\partial^2 u}{\partial s_n^2}$ при переходе через границу между соседними КЭ. Производную $\frac{\partial u}{\partial \nu}$ на любой стороне КЭ аппроксимирует многочлен четвертой степени по s_n , коэффициенты которого однозначно определены значениями $\frac{\partial u}{\partial \nu}$ и $\frac{\partial^2 u}{\partial \nu \partial s_n}$ в двух вершинах на концах этой стороны и значением $\frac{\partial u}{\partial \nu}$ в ее середине. Но значения $\frac{\partial u}{\partial \nu}$ и $\frac{\partial^2 u}{\partial \nu \partial s_n}$ в каждой вершине являются также линейными комбинациями узловых параметров в той же вершине, одинаковыми для соседних КЭ с общими вершинами. Следовательно, при переходе через границу между соседними КЭ будет обеспечена непрерывность аппроксимации и производной $\frac{\partial u}{\partial \nu}$, т.е. гладкая аппроксимация функции u .

Прямоугольный КЭ со сторонами, параллельными осям декартовой системы координат Ox_1x_2 , обеспечивает гладкую аппроксимацию функции $u(x_1, x_2)$, если в каждой вершине с номером m в качестве узловых параметров выбрать по четыре значения u_m , u_{m1} , u_{m2} и u_{m12} , $m = \overline{1, 4}$ (рис. 10.14). Вдоль каждой стороны изменение функции u и ее производной $\frac{\partial u}{\partial \nu}$ в направлении ν нормали к этой стороне описывают кубические многочлены по координате s_m , отсчитываемой от вершины с номером m . При этом коэффициенты многочлена, аппроксимирующего функцию u , однозначно связаны с ее значениями

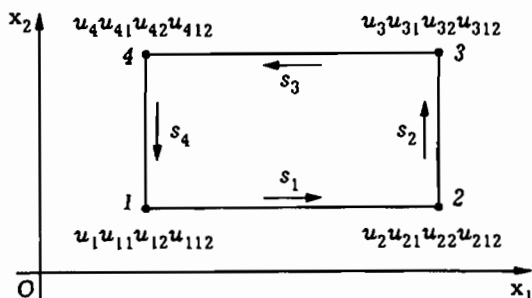


Рис. 10.14

и значениями производной $\frac{\partial u}{\partial s_m}$ в вершинах на концах рассматриваемой стороны, а коэффициенты многочлена для $\frac{\partial u}{\partial \nu}$ — со значениями $\frac{\partial u}{\partial \nu}$ и $\frac{\partial^2}{\partial \nu \partial s_m}$ в этих же вершинах. Таким образом, узловых параметров оказывается достаточно для обеспечения непрерывности аппроксимации функции и ее первых производных при переходе через границу между соседними КЭ. Внутри КЭ функцию u аппроксимирует многочлен, содержащий произведения координат x_1 и x_2 в степенях от 0 до 3. Такой многочлен является неполным и его обычно называют бикубическим. Для каждого из 16 узловых параметров можно построить функцию формы, используя произведения кубических многочленов от x_1 и x_2 вида (10.47).

Четырехугольные (в частности, прямоугольные) КЭ с гладкой аппроксимацией искомой функции можно построить путем объединения треугольных КЭ. Это может привести к уменьшению общего числа узловых параметров, но обычно связано с усложнением функций формы. Возможность уменьшения общего числа узловых параметров становится более важной для трехмерных КЭ. Параметры узлов, расположенных внутри таких КЭ, можно исключить приемом, получившим название конденсации*.

*См., например Норри Д., де Фриз Ж.

10.3. Матричная форма представления функций

Одним из преимуществ *метода конечных элементов* (МКЭ) при решении задач математической физики является простота и однотипность операций по подготовке задачи к решению. В значительной мере это преимущество связано с матричной формой представления основных соотношений, используемых в МКЭ.

Пусть для определенности речь идет о решении задачи в некоторой области $V \subset \mathbb{R}^d$. Сначала следует выбрать тип используемых *конечных элементов* (КЭ) и заполнить ими область V так, чтобы они не пересекались и не образовывали пустот, а также достаточно точно представляли границу области, если она имеет криволинейные участки. В результате получим *сетку* КЭ, занимающую область V_0 , в общем случае не совпадающую с V , но обычно такую, что $V_0 \subseteq V$.

Обозначим общее число КЭ через E . Необходимо установить взаимно однозначное соответствие между номерами $N = \overline{1, N_\Sigma}$ узлов образованной сетки и номерами узлов $n = \overline{1, N_e}$ каждого отдельного КЭ с номером $e = \overline{1, E}$. Если КЭ выбранного типа имеет внутренние узлы, то для таких узлов установление соответствия происходит независимо от соседних КЭ. В противоположность этому граничные узлы принадлежат одновременно нескольким соседним КЭ, что необходимо учитывать во избежание пересечения элементов или возникновения пустот между ними. Итогом установления указанного соответствия между глобальной (сквозной) нумерацией узлов сетки и локальной нумерацией узлов КЭ является построение для каждого КЭ с номером e матрицы Ω_e размера $N_e \times N_\Sigma$, элементы которой $\Omega_{nN}^{(e)} = 1$, если узел N сетки совпадает с узлом n этого КЭ, в противном случае $\Omega_{nN}^{(e)} = 0$.

Зависящие в общем случае от времени t узловые значения $u_N(t)$ действительной функции можно представить на сетке

КЭ матрицей-столбцом U размера $N_\Sigma \times 1$, а значения $u_n^{(e)}(t)$ этой функции в узлах КЭ — матрицами-столбцами U_e размера $N_e \times 1$. Если функция является векторной и имеет D координатных действительных функций, то размеры матриц U и U_e будут $N_\Sigma \times D$ и $N_e \times D$ соответственно. В случае эрмитовых КЭ элементами этих матриц будут не только значения функций, но и значения используемых при аппроксимации функций производных. Тогда под D следует понимать наибольшее число узловых параметров, задаваемых в узле выбранного варианта эрмитова КЭ. Связь между введенными матрицами и их элементами устанавливают соотношения

$$U_e = \Omega_e U, \quad u_n^{(e)}(t) = \sum_{N=1}^{N_\Sigma} \Omega_{nN}^{(e)} u_N(t). \quad (10.50)$$

Замечание 10.2. Рассмотрим случай векторной функции $u(t, M)$ времени t и координат точки $M \in V$, имеющей смысл перемещения (скорости или ускорения) среды. Тогда координатные функции $u_i(t, M)$ $i = \overline{1, d}$, являются проекциями $u(t, M)$ на оси Ox_i прямоугольной системы координат, называемой обычно глобальной. В каждом конечном элементе может быть определена своя так называемая локальная система координат, базис которой не обязательно совпадает с базисом глобальной системы координат. Если эти базисы совпадают, то в (10.50) размеры матриц U и U_e будут $N_\Sigma \times d$ и $N_e \times d$ соответственно, а их элементами — узловые значения $u_{N_i}(t)$ и $u_{n_i}^{(e)}(t)$ соответственно.

Если же базис в КЭ с номером e не совпадает с базисом глобальной системы, то вместо (10.50) в случае векторной функции будем иметь

$$U_e = \Omega_e U \alpha^{(e)}, \quad u_{n_i}^{(e)}(t) = \sum_{j=1}^d \sum_{N=1}^{N_\Sigma} \Omega_{nN}^{(e)} u_{N_j}(t) \alpha_{ij}^{(e)}, \quad (10.51)$$

где $\alpha^{(e)}$ — квадратная матрица перехода порядка d от базиса глобальной системы координат к базису локальной системы в элементе с номером e , причем элементы $\alpha_{ij}^{(e)}$, $i, j = \overline{1, d}$, этой матрицы равны косинусам углов между осями Ox_i и $Ox_j^{(e)}$ глобальной и локальной систем координат соответственно. В дальнейшем для упрощения будем считать базисы локальной системы во всех КЭ совпадающими с базисом глобальной системы координат. #

Пусть выбранный тип конечных элементов является лагранжевым. Для КЭ с номером e и числом узлов N_e , занимающего область $V_e \subset V_0 \subseteq V$, функции формы $\varphi_n^{(e)}(M)$, $n = \overline{1, N_e}$, зависят от координат $x_i^{(e)}(M)$, $i = \overline{1, d}$, точки $M \in V_e$ и образуют матрицу-столбец $\Phi_e(M)$ размера $N_e \times 1$ (см. 10.2). Тогда действительную функцию $u(M)$, $M \in V$, в пределах этого КЭ можно приближенно представить функцией

$$u^{(e)}(M) = U_e^T \Phi_e(M) = \sum_{n=1}^{N_e} u_n^{(e)} \varphi_n^{(e)}(M), \quad M \in V_e, \quad (10.52)$$

где U_e — матрица-столбец размера $N_e \times 1$ с элементами $u_n^{(e)}$, равными узловым значениям функции u .

В случае векторной функции $\underline{u}(M)$, $M \in V$, имеющей D координатных функций $u_i(M)$, $i = \overline{1, D}$, вместо (10.52) получим

$$U_e(M) = U_e^T \Phi_e(M), \quad M \in V_e. \quad (10.53)$$

Здесь U_e — матрица размера $N_e \times D$ с элементами $u_{ni}^{(e)}$, $n = \overline{1, N_e}$, $i = \overline{1, D}$, равными узловым значениям координатных функций, а $U_e(M)$ — матрица-столбец размера $D \times 1$, элементами которой являются координатные функции $u_i^{(e)}(M)$ векторной функции $\underline{u}^{(e)}(M)$, $M \in V_e$, приближенно описывающей в этом КЭ векторное поле, задаваемое функцией $\underline{u}(M)$. Таким

образом, как действительную, так и векторную функции в КЭ можно приближенно представить произведением двух матриц. Элементы одной из них являются узловыми значениями этой функции или ее координатных функций (в общем случае зависящими от времени), а элементы другой — функциями формы, зависящими только от координат точки $M \in V_e$.

Пример 10.1. Предположим, что векторная функция $u(M)$, $M \in V \subset \mathbb{R}^d$, задает поле перемещения среды. Тогда $U = d$ и элементами матрицы-столбца $U_e(M)$ размера $d \times 1$ в (10.53) будут функции $u_i^{(e)}(M)$, которые аппроксимируют в КЭ $V_e \subset V$ проекции $u_i(M)$, $M \in V$, перемещения на координатные оси $Ox_i^{(e)}$, $i = \overline{1, d}$.

Полагая локальные и глобальные координаты любой точки $M \in V_e \subset V$ совпадающими, т.е. $x_i^{(e)}(M) = x_i(M)$, $i = \overline{1, d}$, с учетом (10.50) и (10.53) получаем, что перемещение среды на сетке КЭ можно приближенно задать матрицей-столбцом

$$U(M) = U^T \sum_{e=1}^E \Omega_e^T \Phi_e(M) = U^T \Phi(M) \quad (10.54)$$

размера $d \times 1$ с элементами $u_i^*(M)$, аппроксимирующими проекции $u_i(M)$, $M \in V$. При этом элементами матрицы-столбца

$$\Phi(M) = \sum_{e=1}^E \Omega_e^T \Phi_e(M) \quad (10.55)$$

размера $N_\Sigma \times 1$ являются функции формы $\varphi_N(M)$, $M \in V_0$, сетки КЭ.

Тензор деформаций имеет в прямоугольных координатах компоненты (см. 3.3)

$$\varepsilon_{ij}(M) = \frac{1}{2} \left(\frac{\partial u_i(M)}{\partial x_j} + \frac{\partial u_j(M)}{\partial x_i} \right), \quad M \in V_0, \quad i, j = \overline{1, d}. \quad (10.56)$$

В КЭ с номером e эти компоненты можно приближенно представить в виде

$$\varepsilon_{ij}^{(e)}(M) = \frac{1}{2} \sum_{n=1}^{N_e} \left(u_{ni}^{(e)} \frac{\partial \varphi_n^{(e)}(M)}{\partial x_j} + u_{nj}^{(e)} \frac{\partial \varphi_n^{(e)}(M)}{\partial x_i} \right), \quad (10.57)$$

где $u_{ni}^{(e)}$ — проекции на оси Ox_i , $i = \overline{1, d}$, вектора u в n -м узле КЭ с номером e . В случае симплексных конечных элементов имеем (см. 10.2)

$$\varphi_n^{(e)}(M) = a_n^{(e)} + \sum_{i=1}^d b_{ni}^{(e)} x_i, \quad \frac{\partial \varphi_n^{(e)}(M)}{\partial x_i} = b_{ni}^{(e)} \quad (10.58)$$

и вместо (10.57) получим

$$\varepsilon_{ij}^{(e)}(M) = \frac{1}{2} \sum_{n=1}^{N_e} (u_{ni}^{(e)} b_{nj}^{(e)} + u_{nj}^{(e)} b_{ni}^{(e)}), \quad (10.59)$$

т.е. в каждом симплексном КЭ предполагается однородная (не зависящая от координат) деформация. Соотношения, аналогичные (10.56), (10.57) и (10.59), дают приближенное представление тензора скоростей деформаций. #

При использовании эрмитовых КЭ, когда в качестве узловых параметров выступают не только значения аппроксимируемой функции, но и ее производные по пространственным координатам, вид матричных соотношений остается прежним, если соответствующим образом изменить в (10.52) и (10.53) матрицы $I^{(e)}$ и $\Phi^{(e)}(M)$ (см. 10.2).

Вопросы и задачи

10.1. Вывести формулы (10.12) и (10.13)–(10.15).

10.2. Доказать справедливость (10.19), (10.28) и (10.31).

10.3. Убедиться, что сумма всех функций формы (10.35) и (10.36) для треугольного конечного элемента с шестью и десятью узлами соответственно равна единице.

10.4. Выразить коэффициенты билинейного многочлена через значения аппроксимируемой функции в вершинах прямоугольника.

10.5. Вывести формулы (10.39)–(10.46).

10.6. Получить выражения вида (10.49) для всех функций формы треугольного эрмитового конечного элемента и убедиться, что их сумма равна единице.

11. ПРИКЛАДНЫЕ ЗАДАЧИ

Схема применения *метода конечных элементов* (МКЭ) к решению задач математической физики кратко изложена в начале предыдущей главы. В этой главе мы более подробно остановимся на особенностях МКЭ и примерах его использования при решении прикладных задач.

11.1. Особенности применения метода конечных элементов

Использование *метода конечных элементов* (МКЭ) для решения любой задачи математической физики возможно при наличии *интегральной формулировки* этой задачи. Пути построения такой формулировки рассмотрены выше (см. 6). В частном случае эта формулировка может содержать *функционал* в виде интеграла по области V , где предстоит искать решение задачи, достигающий на искомом решении экстремального значения.

Выбор типа *конечного элемента* (КЭ) при решении задачи зависит от требований к классу функций, на котором допустимо рассматривать ее интегральную формулировку. Если в эту формулировку входят производные искомой функции до порядка p включительно, то допустимые функции должны принадлежать классу непрерывно дифференцируемых $p - 1$ раз, а производные порядка p в области V могут быть кусочно непрерывными.

Интегральная формулировка значительного числа задач математической физики содержит производные лишь первого порядка, т.е. $p = 1$. Поэтому достаточно, чтобы допустимые функции были непрерывны в области V , а их производные утрачивали непрерывность лишь на множестве точек, *мера Лебега*

которого в V равна нулю. Таким свойством обладают функции формы симплексных КЭ. Для некоторых задач интегральная формулировка содержит и производные второго порядка. Это приводит к необходимости использовать эрмитовы КЭ, функции формы которых обеспечивают непрерывность производных первого порядка при переходе через границу между соседними КЭ. Свойство функций формы КЭ обеспечивать выполнение требований к классу функций, диктуемых интегральной формулировкой задачи, называют согласованностью.

После выбора типа КЭ в случае границы области V , имеющей криволинейные участки, возникает проблема заполнения этой области конечными элементами так, чтобы была достаточно точно отражена геометрия границы. Образующей при этом сетке КЭ будет соответствовать область V_0 , в общем случае не совпадающая с V . Отличие V_0 от V является одним из источников погрешности, возникающей при применении МКЭ к решению задачи математической физики. Влияние этой погрешности можно ослабить использованием меньших по размерам элементов, т.е. измельчением сетки КЭ, или представлением границы области при помощи более сложного типа КЭ, одна или несколько сторон (или граней) которых являются криволинейными*. В дальнейшем в этой главе V_0 будем для упрощения отождествлять с V , т.е. считать, что сетка КЭ точно соответствует области решения задачи.

Вторым источником погрешности являются ошибки аппроксимации искомой функции функциями формы выбранного типа КЭ. Влияние этой причины не всегда удастся снизить путем уменьшения размеров КЭ. Более эффективным может оказаться переход от симплексных к комплексным КЭ, а в случае границы простой формы — к мультиплексным КЭ.

По существу МКЭ основан на том, что искомое решение удастся приближенно представить на сетке КЭ в виде разложения по функциям формы, причем коэффициентами этого

*См., например: Норри Д., де Фриз Ж.

разложения являются неизвестные *узловые параметры* — значения искомой функции (а в общем случае и ее производных) в узлах сетки, совпадающих с узлами отдельных КЭ. В случае линейной задачи математической физики подстановка такого разложения в интегральную формулировку задачи приводит к системе линейных алгебраических уравнений (СЛАУ) относительно узловых параметров. Если задача является нелинейной, то элементы матрицы СЛАУ будут зависеть от неизвестных узловых параметров. Тогда узловые параметры приходится находить последовательными приближениями, задаваясь ожидаемыми значениями и уточняя элементы матрицы после очередного решения СЛАУ.

Матрица этой СЛАУ благодаря свойствам функций формы КЭ содержит значительное число нулевых элементов, что упрощает практическую реализацию МКЭ на ЭВМ. Более того, простота и однотипность свойств КЭ позволяют поручить ЭВМ не только решение такой СЛАУ, но и автоматизировать ряд предшествующих этапов: разбиение области решения задачи на конечные элементы и построение их сетки с нумерацией элементов и узлов; построение системы функций формы в пределах каждого КЭ; вычисление вкладов отдельного КЭ в матричное уравнение; формирование глобальной матрицы для всей сетки КЭ. Существуют алгоритмы оптимальной нумерации узлов КЭ, обеспечивающие минимальную ширину ленты ненулевых элементов глобальной матрицы, что позволяет экономить вычислительные ресурсы при хранении и обработке этой матрицы.

Отмеченные особенности превращают МКЭ в один из наиболее гибких и универсальных современных методов численного решения широкого круга задач математической физики. Помимо его строгого обоснования на основе общей теории *проекторных методов* для ряда прикладных задач нетрудно дать физическую или механическую интерпретацию МКЭ с использованием приближенных дискретных моделей сплошной среды, имеющих конечное число степеней свободы. Первые разработ-

ки МКЭ были связаны именно с такими дискретными моделями в механике деформируемого твердого тела. В частности, в задачах статики упругих конструкций разбиение области сеткой КЭ, рассмотрение равновесия отдельных КЭ и установление связей между ними приводит к системе уравнений, совпадающей с той, которая следует из формальной процедуры МКЭ.

Наглядность, простота и возможность автоматизации процедуры МКЭ делает его весьма удобным для численного решения прикладных задач. На его основе разработаны достаточно универсальные программные комплексы, которые широко используют в инженерной практике. Однако для уверенного применения любого приближенного метода необходимо располагать возможностью оценки возникающей погрешности и иметь представление о скорости сходимости приближенного решения задачи к истинному. На практике эти вопросы обычно решают тестированием МКЭ на задачах, для которых известно точное решение. Но для определенного круга задач (в основном, для линейных задач математической физики) можно получить априорные оценки *погрешности аппроксимации* искомого решения на сетке КЭ и скорости его сходимости к истинному решению при измельчении этой сетки*.

11.2. Задачи теплопроводности в твердом теле

Последовательно рассмотрим применение *метода конечных элементов* (МКЭ) для решения линейной задачи стационарной теплопроводности в изотропном твердом теле, возможности учета при помощи МКЭ анизотропии свойств материала тела и нелинейных факторов, а также решение задачи нестационарной теплопроводности.

Стационарная теплопроводность. Пусть распределение температуры $T(M)$, зависящее от положения точки $M \in \bar{V}$ на

* См.: Норри Д., де Фриз Ж., а также Стренг Г., Фикс Дж. или Сьярле Ф.

замыкании $\bar{V} = V \cup S$ области V , ограниченной поверхностью S (рис. 11.1), удовлетворяет дифференциальному уравнению

$$\nabla(\lambda(M)\nabla T(M)) + I_V^{(q)}(M) = 0, \quad M \in V, \quad (11.1)$$

с граничными условиями (2.55) и (2.56)

$$T(P) = f_1(P), \quad P \in S_1, \quad (11.2)$$

$$\lambda(P)\nabla T(P)\mathbf{n}(P) + \beta(P)T(P) = f_2(P), \quad P \in S_2. \quad (11.3)$$

Здесь $\lambda > 0$ — коэффициент теплопроводности; $I_V^{(q)}$ — мощность объемного энергоснабжения; f_1 и f_2 — заданные функции положения точки P на участках $S_1 \subset S$ и $S_2 = S \setminus S_1 \neq \emptyset$ поверхности S соответственно; $\mathbf{n}$ — единичный вектор внешней нормали к поверхности S ; $\beta > 0$ — коэффициент теплообмена. Уравнение (11.1) является частным случаем (2.53) и следует из закона сохранения тепловой энергии в теле.

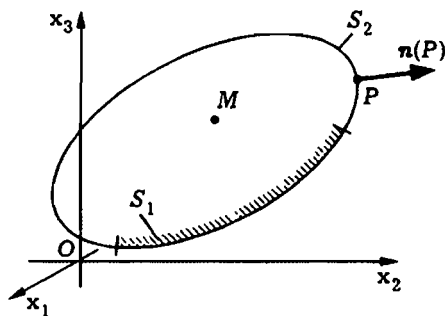


Рис. 11.1

Краевой задаче (11.1)–(11.3) соответствует ее *интегральная формулировка*, содержащая функционал [XV]

$$J[T] = \int_V \left(\frac{\lambda}{2} (\nabla T)^2 - I_V^{(q)} T \right) dV + \int_{S_2} \left(\frac{\beta}{2} T^2 - f_2 T \right) dS, \quad (11.4)$$

минимизируемый на распределениях температуры $T(M)$, удовлетворяющих (11.2), непрерывных на замыкании $\bar{V}$ и имеющих в V кусочно непрерывные производные по пространственным координатам x_i , $i = \overline{1, 3}$ (в случае трехмерной задачи). Функция $T^*(M)$, $M \in \bar{V}$, удовлетворяющая (11.1)–(11.3), является *точкой экстремума* (минимума) *функционала* (11.4). Этот функционал можно использовать для приближенного решения рассматриваемой задачи при помощи МКЭ. При этом *функции формы конечных элементов* (КЭ) обеспечивают непрерывную аппроксимацию искомой функции $T(M)$, $M \in \bar{V}$, на сетке КЭ, заполняющей область V , а в пределах каждого КЭ непрерывность производных функции $T(M)$.

Пусть сетка КЭ состоит из E лагранжеевых конечных элементов, каждый из которых имеет N_e узлов и занимает область $V_e \subset V$, $e = \overline{1, E}$. Тогда вместо (11.4) можно написать

$$J[T] = \sum_{e=1}^E J_e[T], \quad (11.5)$$

причем вклад КЭ с номером e в значение функционала $J[T]$ составляет

$$J_e[T] = \int_{V_e} \left(\frac{\lambda}{2} (\nabla T)^2 - I_V^{(q)} T \right) dV + \int_{S'_e} \left(\frac{\beta}{2} T^2 - f_2 T \right) dS. \quad (11.6)$$

Второй интеграл в правой части (11.6) отличен от нуля лишь при условии, что граница S'_e данного КЭ имеет общие участки с поверхностью S_2 , т.е. $S'_e \cap S_2 \neq \emptyset$.

В пределах каждого КЭ искомое распределение температуры аппроксимируем матричным выражением вида (10.52):

$$T(M) \approx T^{(e)}(M) = \Theta_e^T \Phi_e(M), \quad M \in V_e, \quad e = \overline{1, E}, \quad (11.7)$$

где Θ и $\Phi_e(M)$ — матрицы-столбцы размера $N_e \times 1$, элементами которых являются значения $T_n^{(e)}$ температуры в узлах КЭ с

номером e и функции формы $\varphi_n^{(e)}(M)$, $n = \overline{1, N_e}$, этого КЭ соответственно. Тогда в матричной записи будем иметь

$$\begin{aligned} T^2(P) &\approx (T^{(e)}(P))^2 = \Theta_e^T \Phi_e(P) (\Theta_e^T \Phi_e(P))^T = \\ &= \Theta_e^T \Phi_e(P) \Phi_e^T(P) \Theta_e, \quad P \in S'_e. \end{aligned} \quad (11.8)$$

Так как с учетом (11.7)

$$\frac{\partial T^{(e)}(M)}{\partial x_i} = \Theta_e^T B_i^{(e)}(M), \quad M \in V_e, \quad i = \overline{1, 3},$$

где $B_i^{(e)}(M)$ — матрица-столбец размера $N_e \times 1$ с элементами $\frac{\partial \varphi_n^{(e)}(M)}{\partial x_i}$, $n = \overline{1, N_e}$, $i = \overline{1, 3}$, то квадрат градиента температурного поля в КЭ можно представить в виде

$$\begin{aligned} (\nabla T(M))^2 &\approx (\nabla T^{(e)}(M))^2 = \Theta_e^T B_e(M) (\Theta_e^T B_e(M))^T = \\ &= \Theta_e^T B_e(M) B_e^T(M) \Theta_e, \quad M \in V_e, \quad e = \overline{1, E}. \end{aligned} \quad (11.9)$$

Здесь $B_e(M)$ — блочная матрица размера $N_e \times 3$, блоками которой являются матрицы $B_i^{(e)}$.

После подстановки (11.7)–(11.9) в (11.6) получим

$$\begin{aligned} J_e[T] &\approx \int_{V_e} \left(\frac{\lambda(M)}{2} (\nabla T^{(e)}(M))^2 - I_V^{(q)}(M) T^{(e)}(M) \right) dV + \\ &\quad + \int_{S'_e} \left(\frac{\beta(P)}{2} (T^{(e)}(P))^2 - f_2(P) T^{(e)}(P) \right) dS = \\ &= \int_{V_e} \left(\frac{\lambda(M)}{2} \Theta_e^T B_e(M) B_e^T(M) \Theta_e - I_V^{(q)}(M) \Theta_e^T \Phi_e(M) \right) dV + \\ &\quad + \int_{S'_e} \left(\frac{\beta(P)}{2} \Theta_e^T \Phi_e(P) \Phi_e^T(P) \Theta_e - f_2(P) \Theta_e^T \Phi_e(P) \right) dS. \end{aligned}$$

Элементы матрицы Θ_e узловых значений температуры не зависят от координат, и поэтому в силу линейности матричных операций и операции интегрирования можно записать

$$J_e[T] \approx \frac{1}{2} \Theta_e^T \Lambda_e \Theta_e - \Theta_e^T Q_e, \quad (11.10)$$

где Λ_e — квадратная матрица порядка N_e , называемая **матрицей теплопроводности КЭ**, с элементами

$$\Lambda_{ln}^{(e)} = \int_{V_e} \lambda(M) \sum_{i=1}^3 \frac{\partial \varphi_i^{(e)}(M)}{\partial x_i} \frac{\partial \varphi_n^{(e)}(M)}{\partial x_i} dV + \\ + \int_{S_e} \beta(P) \varphi_l^{(e)}(P) \varphi_n^{(e)}(P) dS, \quad l, n = \overline{1, N_e}, \quad (11.11)$$

а Q_e — матрица-столбец размера $N_e \times 1$ с элементами

$$Q_l^{(e)} = \int_{V_e} I_V^{(q)}(M) \varphi_l^{(e)}(M) dV + \int_{S_e} f_2(P) \varphi_l^{(e)}(P) dS. \quad (11.12)$$

Отметим, что матрица Λ_e является симметрической.

Перейдем в (11.10) от локальной нумерации узлов к глобальной (см. 10.3) при помощи матриц Ω_e размера $N_e \times N_\Sigma$, $e = \overline{1, E}$, где N_Σ — общее число узлов сетки КЭ. Тогда после подстановки (11.10) в (11.5) получим

$$J[T] \approx \frac{1}{2} \Theta^T \Lambda \Theta - \Theta^T Q, \quad (11.13)$$

где Θ — матрица-столбец размера $N_\Sigma \times 1$ значений температуры T_N в узлах сетки КЭ; Λ — симметрическая матрица порядка N_Σ , называемая **глобальной матрицей теплопроводности**, с элементами

$$\Lambda_{LN} = \sum_{e=1}^E \sum_{l=1}^{N_e} \sum_{n=1}^{N_e} \Omega_{lL}^{(e)} \Lambda_{ln}^{(e)} \Omega_{nN}^{(e)}, \quad (11.14)$$

причем $\Omega_{nN}^{(e)}$ — элементы матрицы Ω_e ; Q — матрица-столбец размера $N_\Sigma \times 1$ с элементами

$$Q_N = \sum_{e=1}^E \sum_{n=1}^{N_e} \Omega_{nN}^{(e)} Q_n^{(e)}. \quad (11.15)$$

Правую часть (11.13) можно рассматривать как конечно-мерный аналог $J[\tilde{T}_{N_\Sigma}]$ функционала (11.4), полученный при использовании метода Рунца и соответствующий элементу

$$\tilde{T}_{N_\Sigma}(M) = \Theta^T \Phi(M) = \sum_{N=1}^{N_\Sigma} T_N \varphi_N(M) \quad (11.16)$$

последовательности, минимизирующей этот функционал, причем $\Phi(M)$ — матрица-столбец размера $N_\Sigma \times 1$, элементами которой являются функции формы

$$\varphi_N(M) = \sum_{e=1}^E \sum_{n=1}^{N_e} \Omega_{nN}^{(e)} \varphi_n^{(e)}(M), \quad N = \overline{1, N_\Sigma}, \quad (11.17)$$

определенные на сетке КЭ. Правую часть (11.13) можно представить многочленом второй степени относительно N_Σ переменных T_N :

$$J[\tilde{T}_{N_\Sigma}] = \frac{1}{2} \sum_{L=1}^{N_\Sigma} \sum_{N=1}^{N_\Sigma} \Lambda_{LN} T_L T_N - \sum_{L=1}^{N_\Sigma} Q_L T_L. \quad (11.18)$$

Так как минимизируемый функционал (11.4) ограничен снизу $[XV]$, то функция (11.18) также ограничена снизу, поскольку в этом случае для минимального значения функционала, достигаемого на функции $T_*(M)$, $M \in \bar{V}$, справедливо неравенство

$$J[T_*] \leq J[\tilde{T}_{N_\Sigma}]. \quad (11.19)$$

Поэтому функция (11.18) также достигает минимума при некотором наборе значений T_N , $N = \overline{1, N_\Sigma}$.

В частном случае, когда в (11.1) $I_V^{(q)}(M) = 0$, $M \in V$, а в (11.3) $f_2(P) = 0$, $P \in S_2$, имеем $Q_L = 0$, $L = \overline{1, N_\Sigma}$, т.е. (11.18) будет квадратичной формой $\Theta^T \Lambda \Theta$. Если при этом в (11.2) $f_1(P) = 0$, $P \in S_1$, или же участки поверхности S_1 отсутствуют, то оператор $Au = -\nabla(\lambda \nabla u)$ в (11.1) при $\lambda(M) > 0$, $M \in V$, является положительно определенным (см. пример 5.9), а дифференциальное уравнение $\nabla(\lambda \nabla T) = 0$ имеет единственное решение $T_0(M) = 0$, $M \in \overline{V} [XV]$. На этом решении функционал (11.4) достигает минимального значения, причем $J[T_0] = \Theta_0^T \Lambda \Theta_0 = 0$, где Θ_0 — матрица-столбец (вектор) размера $N_\Sigma \times 1$ с нулевыми элементами.

Стационарная точка $T_0(M) = 0$, $M \in \overline{V}$, функционала (11.4), в которой он достигает минимума при $I_V^{(q)}(M) = 0$, $M \in V$, и $f_2(P) = 0$, $P \in S_2$, единственна. Поэтому для любой функции $\tilde{T}_{N_\Sigma}(M) \neq 0$, $M \in \overline{V}$, вида (11.16), удовлетворяющей (11.2) при $f_1(P) = 0$, $P \in S_1$, в соответствии с (11.19) имеем $J[\tilde{T}_{N_\Sigma}] > J[T_0] = 0$. Следовательно, для соответствующего этой функции ненулевого вектора Θ получаем $\Theta^T \Lambda \Theta > 0$.

Пусть $S_1 \neq \emptyset$ и сетка КЭ имеет N_1 узлов $P_N \in S_1$, номера которых упорядочены так, что $N = \overline{N^*+1, \overline{N}}$, где $N^* = N - N_1$. Предположим, что существует функция $\tilde{T}_{N_\Sigma}(M)$, $M \in \overline{V}$, принимающая в этих узлах значения $\tilde{T}_{N_\Sigma}(P_N) = f_{1N}$, $P_N \in S_1$, $N = \overline{N^*+1, \overline{N}}$, из которых хотя бы одно отлично от нуля, и такая, что $\Theta^T \Lambda \Theta \leq 0$ для соответствующего ей ненулевого вектора Θ . Тогда найдется непрерывная функция $T_*(M)$, $M \in \overline{V}$, имеющая в V кусочно непрерывные производные, удовлетворяющая условиям $T_*(P_N) = f_{1N}$, $P_N \in S_1$, $N = \overline{N^*+1, \overline{N}}$, и минимизирующая функционал (11.4) при $I_V^{(q)}(M) = 0$, $M \in V$, и $f_2(P) = 0$, $P \in S_2$. Для этой функции $J[T_*] > 0$. Действительно, по физическому смыслу задачи в этом случае на участках поверхности $S_2 \neq \emptyset$ происходит теплообмен тела, температура которого отлична от нуля, со средой, имеющей нулевую температуру. Так как функция $T_*(M)$, $M \in \overline{V}$, является стационарной точкой функционала (11.4), то выполнено условие стационарности (11.3), в

котором слагаемое $(\lambda(P)\nabla T(P))n(P)$ равно отличной от нуля плотности теплового потока, проходящего через поверхность S_2 . Следовательно, $T(P) \neq 0$, $P \in S_2$, а этого уже достаточно, чтобы $J[T_*] > 0$ при $\beta > 0$. Тогда с учетом (11.19) получаем, что $J[\tilde{T}_{N_\Sigma}] \geq J[T_*] > 0$, т.е. для ненулевого вектора Θ , соответствующего функции $\tilde{T}_{N_\Sigma}(M)$, $M \in \bar{V}$, имеем $\Theta^T \Lambda \Theta > 0$, а это противоречит сделанному предположению. Таким образом, для любого ненулевого вектора Θ квадратичная форма $\Theta^T \Lambda \Theta$ и ее матрица Λ являются положительно определенными.

Если $S_1 \neq \emptyset$ и сетка КЭ имеет N_1 узлов $P_N \in S_1$, $N = \overline{N^*+1, N}$, с заданными в соответствии с (11.2) температурами $T(P_N) = f_1(P_N) = f_{1N}$, то (11.16) должно принимать в этих узлах заданные значения. Тогда из необходимых условий $\frac{\partial J[\tilde{T}_{N_\Sigma}]}{\partial T_L} = 0$, $L = \overline{1, N^*}$, минимума функции (11.18) с учетом $\Lambda_{LN} = \Lambda_{NL}$ получим систему линейных алгебраических уравнений (СЛАУ)

$$\sum_{N=1}^{N^*} \Lambda_{LN} T_N = Q_L - \sum_{N=N^*+1}^{N_\Sigma} \Lambda_{LN} f_{1N}, \quad L = \overline{1, N^*}. \quad (11.20)$$

Симметрическая матрица Λ^* порядка N^* этой СЛАУ получена из положительно определенной матрицы Λ вычеркиванием последних N_1 строк и столбцов. Поэтому все ее угловые миноры останутся положительными, т.е., согласно критерию Сильвестра [IV], матрица Λ^* также положительно определенная, а СЛАУ (11.20) имеет единственное решение. Если же участки поверхности S_1 отсутствуют, то $N_* = N_\Sigma$ и сумма в правой части (11.20) исчезает.

Отметим, что в силу свойств функций формы КЭ в матрице Λ^* отличны от нуля лишь элементы на пересечении строк и столбцов с номерами, соответствующими номерам узлов сетки, принадлежащих одному и тому же КЭ. Поскольку обычно $N_e \ll N_\Sigma$, большинство элементов матрицы Λ^* равны нулю. Такие матрицы называют **слабозаполненными** (или **разреженными**). Для численного решения СЛАУ с такой матрицей

можно использовать способы, приводящие к существенной экономии вычислительных ресурсов ЭВМ*.

Учет анизотропии материала твердого тела. Если материал тела анизотропный, то λ в (11.1) и (11.3) следует заменить тензором теплопроводности $\tilde{\lambda}$, который можно представить симметрической матрицей (2.95) третьего порядка с элементами $\lambda_{ij}(M) = \lambda_{ji}(M)$, $i, j = \overline{1, 3}$. Тогда функционал (11.4) примет вид

$$\tilde{J}[T] = \int_V \left(\frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 \lambda_{ij} \frac{\partial T}{\partial x_i} \frac{\partial T}{\partial x_j} - I_V^{(q)} T \right) dV + \int_{S_2} \left(\frac{\beta}{2} T^2 - f_2 T \right) dS,$$

а вместо (11.11) для КЭ с номером e получим

$$\Lambda_{ln}^{(e)} = \int_{V_e} \sum_{i=1}^3 \sum_{j=1}^3 \lambda_{ij}(M) \frac{\partial \varphi_l^{(e)}(M)}{\partial x_i} \frac{\partial \varphi_n^{(e)}(M)}{\partial x_j} dV + \\ + \int_{S'_e} \beta(P) \varphi_l^{(e)}(P) \varphi_n^{(e)}(P) dS. \quad (11.21)$$

Зависимость компонент $\lambda_{ij}(M)$, $i, j = \overline{1, 3}$, тензора $\tilde{\lambda}$ от декартовых координат $x_i(M)$ точки $M \in V_e$ аппроксимируем в пределах КЭ с номером e соотношением вида (10.52)

$$\lambda_{ij}(M) \approx \sum_{m=1}^{N_e} \lambda_{ij}^{(e)}(M_m) \varphi_m^{(e)}(M),$$

где $\lambda_{ij}^{(e)}(M_m)$ — известное значение $\lambda_{ij}^{(e)}(M)$ в узле с номером m . Тогда первый интеграл в (11.21) примет вид

$$\bar{\Lambda}_{ln}^{(e)} = \sum_{m=1}^{N_e} \lambda_{ij}^{(e)}(M_m) \int_{V_e} \sum_{i=1}^3 \sum_{j=1}^3 \varphi_m^{(e)}(M) \frac{\partial \varphi_l^{(e)}(M)}{\partial x_i} \frac{\partial \varphi_n^{(e)}(M)}{\partial x_j} dV.$$

*См.: Джордж А., Лю Дж., а также: Писсанецки С.

Отсюда в случае симплексных КЭ, используя (10.58), получаем

$$\bar{\Lambda}_{ln}^{(e)} = \sum_{i=1}^3 \sum_{j=1}^3 b_{li}^{(e)} b_{nj}^{(e)} \sum_{m=1}^{N_e} \lambda_{ij}^{(e)}(M_m) \int_{V_e} \varphi_m^{(e)}(M) dV. \quad (11.22)$$

Интеграл в (11.22) не зависит от номера m узла. Поэтому для трехмерной задачи при $N_e = 4$ при помощи (10.31) находим

$$\bar{\Lambda}_{ln}^{(e)} = V_e \sum_{i=1}^3 \sum_{j=1}^3 b_{li}^{(e)} b_{nj}^{(e)} \bar{\lambda}_{ij}^{(e)}, \quad (11.23)$$

где через V_e обозначен объем области конечного элемента с номером e , а

$$\bar{\lambda}_{ij}^{(e)} = \frac{1}{N_e} \sum_{m=1}^{N_e} \lambda_{ij}^{(e)}(M_m)$$

является среднеарифметическим узловых значений компоненты λ_{ij} тензора $\tilde{\lambda}$ в симплексном КЭ. В случае двумерной задачи ($N_e = 3$) в (11.23) V_e следует заменить на площадь F_e КЭ, а для одномерной ($N_e = 2$) — на его длину l_e . Для изотропного материала вместо (11.23) получим

$$\bar{\Lambda}_{ln}^{(e)} = \bar{\lambda}^{(e)} V_e \sum_{i=1}^3 b_{li}^{(e)} b_{ni}^{(e)},$$

где $\bar{\lambda}^{(e)}$ — среднеарифметическое узловых значений λ в симплексном КЭ.

Пусть в трехмерном КЭ с номером e одна из граней имеет площадь S'_e и аппроксимирует участок поверхности S_2 , на котором задано граничное условие (11.3). Зависимость коэффициента теплообмена $\beta(P)$ от координат точки $P \in S'_e$ аппроксимируем соотношением

$$\beta(P) \approx \sum_{m=1}^{N'_e} \beta_m^{(e)} \varphi_m^{(e)}(P),$$

где N'_e — число узлов КЭ, принадлежащих рассматриваемой грани; $\beta_m^{(e)}$ — значение $\beta(P)$ в узле с номером m . Тогда второй интеграл в (11.21) примет вид

$$\bar{\Lambda}_{ln}^{(e)} = \sum_{m=1}^{N'_e} \beta_m^{(e)} \int_{S'_e} \varphi_m^{(e)}(P) \varphi_l^{(e)}(P) \varphi_n^{(e)}(P) dS.$$

Так как грань трехмерного симплексного КЭ является двумерным симплексным КЭ, для которого $N'_e = 3$, то, используя (10.28), находим

$$\bar{\Lambda}_{11}^{(e)} = \frac{3\beta_1^{(e)} + \beta_2^{(e)} + \beta_3^{(e)}}{30} S'_e, \quad \bar{\Lambda}_{12}^{(e)} = \frac{2\beta_1^{(e)} + 2\beta_2^{(e)} + \beta_3^{(e)}}{60} S'_e,$$

а остальные значения $\bar{\Lambda}_{ln}^{(e)}$ получаем соответствующей перестановкой нижних индексов. Для двумерной задачи $N'_e = 2$. Используя (10.19), вычисляем $\bar{\Lambda}_{11}^{(e)} = \frac{1}{12}(3\beta_1^{(e)} + \beta_2^{(e)})l'_e$, $\bar{\Lambda}_{12}^{(e)} = \frac{1}{12}(\beta_1^{(e)} + \beta_2^{(e)})l'_e$, где l'_e — длина стороны треугольного симплексного КЭ, на которой заданы граничные условия вида (11.3). Аналогично можно вычислить интегралы в (11.12).

Учет нелинейности. Зависимость λ и $I_V^{(q)}$ от координат точки $M \in V$ в (11.1) и λ , β и f_2 от координат точки $P \in S_2$ в (11.3) позволяет путем итераций учесть и зависимость этих функций от температуры, т.е. приближенно решить нелинейную задачу стационарной теплопроводности. Уточняя узловые значения этих функций по найденным в первом приближении из решения СЛАУ (11.20) значениям $T_N^{(1)}$, $N = \bar{1}, \bar{N}^*$, температуры в узлах сетки КЭ, можно вычислить уточненные значения элементов матриц Λ и Q , а затем, решая снова СЛАУ (11.20), получить узловые значения $T_N^{(2)}$ во втором приближении и т.д. Вместе с тем благодаря известному экстремальному свойству функционала (11.4) для решения нелинейной задачи применимы и различные модификации методов оптимизации [XIV]. Кратко остановимся на одном из наиболее простых вариантов с точки

зрения построения алгоритма решения нелинейных задач, называемом **методом локальных вариаций**.

Вариация δT_N температуры лишь в одном N -м узле сетки КЭ относительно принятого в качестве нулевого приближения значения $T_N^{(0)}$ вызывает изменение $\delta J[\tilde{T}_{N_\Sigma}]$ лишь той части функционала (11.18), которая непосредственно зависит от узлового значения T_N . Варьируя только это значение, можно добиться минимума $J[\tilde{T}_{N_\Sigma}]$, а затем перейти к варьированию температуры в соседнем узле и т.д. После обхода всех узлов подобную процедуру приходится повторять, так как каждое последующее варьирование температуры в соседнем узле может отклонить значение $J[\tilde{T}_{N_\Sigma}]$ от минимального. Поэтому сначала рационально вести варьирование во всех узлах с достаточно грубым шагом ΔT , а затем, получив совокупность значений $T_N^{(1)}$, $N = \overline{1, N^*}$, первого приближения, для которых переход к температурам $T_N = T_N^{(1)} \pm \Delta T$ уже не понижает значения $J[\tilde{T}_{N_\Sigma}]$, следует уменьшить шаг варьирования (например, вдвое) и повторять описанную процедуру до тех пор, пока шаг варьирования не станет менее заданной точности вычисления температуры.

Нестационарная теплопроводность. Перейдем к рассмотрению задачи нестационарной теплопроводности в твердом теле, описываемой дифференциальным уравнением (при $t > 0$, $M \in V$) вида (2.53)

$$c(M) \frac{\partial T(t, M)}{\partial t} = \nabla(\lambda(M) \nabla T(t, M)) + I_V^{(q)}(t, M), \quad (11.24)$$

где $c(M)$ — объемная теплоемкость материала тела, с граничными условиями вида (11.2), (11.3)

$$T(t, P) = f_1(t, P), \quad P \in S_1, \quad (11.25)$$

$$\lambda(P) \nabla T(t, P) \mathbf{n}(P) + \beta(t, P) T(t, P) = f_2(t, P), \quad P \in S_2, \quad (11.26)$$

и начальным условием $T(0, M) = f_0(M)$, $M \in V \cup S_2$, в момент времени $t = 0$, принимаемый за начало отсчета. Сформулиро-

ванной задаче не удастся поставить в соответствие функционал, для которого ее решение было бы его точкой экстремума*. Поэтому в качестве интегральной формулировки этой задачи примем условия вида (6.76)

$$\int_V \left(c(M) \frac{\partial T(t, M)}{\partial t} - \nabla(\lambda(M) \nabla T(t, M)) - I_V^{(q)}(t, M) \right) v_L(M) dV + \\ + \int_{S_2} (\lambda(P) \nabla T(t, P) \mathbf{n}(P) + \beta(t, P) T(t, P) - f_2(t, P)) v_L(P) dS = 0$$

ортогональности проекций невязки, возникающей при подстановке в (11.24) искомого приближенного решения, на элементы v_L , $L = \overline{1, N_\Sigma}$, базиса N_Σ -мерного подпространства $\mathcal{H}_{N_\Sigma}$ гильбертова пространства $\mathcal{H}$ непрерывных в $\bar{V}$ функций, имеющих в V кусочно непрерывные производные и обращающихся в нуль на S_1 .

Преобразовывая последнее равенство при помощи первой формулы Грина и учитывая, что $v_L(P) = 0$ при $P \in S_1$, получаем (обозначения аргументов функций опущены)

$$\int_V \left((\nabla(\lambda \nabla T)) \nabla v_L + \left(c \frac{\partial T}{\partial t} - I_V^{(q)} \right) v_L \right) dV + \\ + \int_{S_2} (\beta T - f_2) v_L dS = 0, \quad L = \overline{1, N_\Sigma}. \quad (11.27)$$

Приближенное решение рассматриваемой задачи будем искать в виде, аналогичном (11.16):

$$\tilde{T}_{N_\Sigma}(t, M) = \Phi^T(M) \Theta(t) = \sum_{N=1}^{N_\Sigma} T_N(t) \varphi_N(M), \quad (11.28)$$

где $T_N(t)$ — зависящие от времени значения температуры в узлах сетки КЭ, а в качестве базиса подпространства $\mathcal{H}_{N_\Sigma}$ примем систему $\{\varphi_N(M)\}_{N_\Sigma}$ функций (11.17), т.е. базисные и

*См. Зарубин В.С., Селиванов В.В

проекционные функции возьмем одинаковыми, что характерно для процедуры метода Бубнова — Галеркина. В примере 6.15 получена система обыкновенных дифференциальных уравнений (ОДУ) для нахождения зависящих от времени коэффициентов $a_n(t)$ в приближенном решении (6.173) рассматриваемой задачи. Однако отличие (11.28) от (6.173) состоит в том, что при использовании МКЭ нет необходимости в построении функции, которая удовлетворяла бы *главному граничному условию* (11.25), так как достаточно в (11.28) положить $T_N(t) = f_1(t, P_N) = f_{1N}(t)$ в N_1 узлах $P_N \in S_1$ конечноэлементной сетки, принадлежащих участкам S_1 поверхности S . Это позволяет из числа базисных функций исключить те функции φ_N , которые соответствуют указанным узлам, но сохранить их в представлении (11.28) приближенного решения.

Номера N оставшихся базисных функций $\varphi_N(M)$ упорядочим так, что $N = \overline{1, N^*}$, где $N^* = N_\Sigma - N_1$ (ясно, что в частном случае отсутствия участков S_1 , т.е. $S_2 = S$, имеем $N_1 = 0$ и $N^* = N_\Sigma$). В силу свойств функций формы КЭ имеем $\varphi_N(P) = 0$, $N = \overline{1, N^*}$, $P \in S_1$. Подставляя (11.28) в (11.27) вместо T , приходим к системе ОДУ (обозначения аргументов функций опущены)

$$\begin{aligned} & \sum_{N=1}^{N^*} \frac{dT_N}{dt} \int_V c \varphi_L \varphi_N dV + \\ & + \sum_{N=1}^{N^*} T_N \left(\int_V \lambda \sum_{i=1}^3 \frac{\partial \varphi_L}{\partial x_i} \frac{\partial \varphi_N}{\partial x_i} dV + \int_{S_2} \beta \varphi_L \varphi_N dS \right) = \\ & = \int_V I_V^{(q)} \varphi_L dV + \int_{S_2} f_2 \varphi_L dS - \sum_{N=N^*+1}^{N_\Sigma} \frac{df_{1N}}{dt} \int_V c \varphi_L \varphi_N dV - \\ & - \sum_{N=N^*+1}^{N_\Sigma} f_{1N} \left(\int_V \lambda \sum_{i=1}^3 \frac{\partial \varphi_L}{\partial x_i} \frac{\partial \varphi_N}{\partial x_i} dV + \int_{S_2} \beta \varphi_L \varphi_N dS \right), \quad L = \overline{1, N^*} \end{aligned}$$

или, используя обозначения элементов матриц,

$$\sum_{N=1}^{N^*} \left(C_{LN} \frac{dT_N(t)}{dt} + \Lambda_{LN} T_N(t) \right) = \\ = Q_L - \sum_{N=N^*+1}^{N_\Sigma} \left(C_{LN} \frac{df_{1N}(t)}{dt} + \Lambda_{LN} f_{1N}(t) \right), \quad L = \overline{1, N^*}. \quad (11.29)$$

Начальные условия для этой системы составляют значения $T_N(0) = f_0(M_N)$, где M_N — узлы сетки КЭ с номерами $N = \overline{1, N^*}$.

Отметим, что элементы Λ_{LN} , $L, N = \overline{1, N_\Sigma}$, глобальной матрицы теплопроводности Λ могут зависеть в (11.29) от времени в силу возможной зависимости от времени коэффициента теплообмена β в граничном условии (11.26). Аналогичное замечание относится и к элементам Q_L матрицы-столбца Q . В (11.29) входят также значения

$$C_{LN} = \int_V c(M) \varphi_L(M) \varphi_N(M) dV, \quad L, N = \overline{1, N_\Sigma}, \quad (11.30)$$

которые являются элементами симметрической матрицы порядка N_Σ , называемой **глобальной матрицей теплоемкости**. Ее (как и глобальную матрицу теплопроводности) удобнее формировать из вкладов отдельных КЭ. Учитывая (11.17) и (11.30), эти значения можно представить в виде

$$C_{LN} = \sum_{e=1}^E \int_{V_e} c(M) \left(\sum_{l=1}^{N_e} \Omega_{lL}^{(e)} \varphi_l^{(e)}(M) \right) \left(\sum_{n=1}^{N_e} \Omega_{nN}^{(e)} \varphi_n^{(e)}(M) \right) dV = \\ = \sum_{e=1}^E \sum_{l=1}^{N_e} \sum_{n=1}^{N_e} C_{ln}^{(e)} \Omega_{lL}^{(e)} \Omega_{nN}^{(e)},$$

где

$$C_{ln}^{(e)} = \int_{V_e} c(M) \varphi_l^{(e)}(M) \varphi_n^{(e)}(M) dV. \quad (11.31)$$

Элементы $C_{ln}^{(e)}$, $l, n = \overline{1, N_e}$, образуют симметрическую **матрицу теплоемкости КЭ** с номером e , имеющую порядок N_e . Элементы этой матрицы распределяют полную теплоемкость КЭ

$$C^{(e)} = \int_{V_e} c(M) dV$$

по его узлам в виде сосредоточенных теплоемких масс таким образом, что изменение внутренней энергии дискретной системы совпадает с ее изменением в КЭ при непрерывных распределениях $c(M)$ и $T(t, M)$, $M \in V_e$. Если зависимость $c(M)$ от координат точки $M \in V_e$ аппроксимировать в пределах этого КЭ соотношением

$$c(M) \approx \sum_{m=1}^{N_e} c_m^{(e)} \varphi_m^{(e)}(M),$$

где $c_m^{(e)}$ — значение $c(M)$ в узле с номером m , то вместо равенства (11.31) получим

$$C_{ln}^{(e)} = \sum_{m=1}^{N_e} c_m^{(e)} \int_{V_e} \varphi_m^{(e)}(M) \varphi_l^{(e)}(M) \varphi_n^{(e)}(M) dV. \quad (11.32)$$

Отсюда в случае симплексных КЭ для трехмерной задачи при $N_e = 4$, используя (10.31), находим

$$C_{11}^{(e)} = \frac{3c_1^{(e)} + c_2^{(e)} + c_3^{(e)} + c_4^{(e)}}{60} V_e, \quad C_{12}^{(e)} = \frac{2c_1^{(e)} + 2c_2^{(e)} + c_3^{(e)} + c_4^{(e)}}{120} V_e,$$

а остальные значения $C_{ln}^{(e)}$ получаем соответствующей перестановкой нижних индексов. Для двумерной ($N_e = 3$) и одномерной задач значения $C_{ln}^{(e)}$ нетрудно вычислить, используя (10.28) и (10.19) соответственно.

Систему ОДУ (11.29) запишем в матричном виде

$$C^* \frac{d\Theta^*}{dt} + \Lambda^* \Theta = Q^*, \quad (11.33)$$

где C^* и Λ^* — симметрические матрицы порядка N^* с элементами C_{LN} и Λ_{LN} , $L, N = \overline{1, N^*}$, соответственно; Θ^* и Q^* матрицы-столбцы размера $N^* \times 1$ с элементами $T_N(t)$ и

$$Q_L^* = Q_L - \sum_{N=N^*+1}^{N_\Sigma} \left(C_{LN} \frac{df_{1N}}{dt} + \Lambda_{LN} f_{1N} \right), \quad L = \overline{1, N^*},$$

соответственно. Обращением матрицы C^* систему (11.33) можно привести к *нормальной системе ОДУ* и для ее решения использовать один из вариантов *метода Рунге — Кутты*. Однако обращение матрицы C^* далеко не всегда оправдано с точки зрения экономии вычислительных ресурсов ЭВМ при полном решении задачи, особенно в тех случаях, когда существенно зависимость теплоемкости тела от температуры и эту зависимость необходимо учесть путем решения (11.33) итерациями, последовательно уточняя элементы матрицы C^* .

Рассмотрим некоторые способы приближенного решения системы (11.33), имея в виду для нелинейных задач возможность учета зависимости от температуры не только теплоемкости, но и коэффициента теплопроводности тела. Используем конечно-разностную аппроксимацию производной в (11.33) в пределах интервала времени $\Delta t_k = t_k - t_{k-1}$, приняв

$$\frac{d\Theta^*}{dt} \approx \frac{\Theta_k^* - \Theta_{k-1}^*}{\Delta t_k},$$

где индексы $k-1$ и k соответствуют моментам времени в начале и конце k -го интервала. Применим *двухслойную разностную схему с весами* и вместо (11.33) получим

$$\begin{aligned} ((1-\eta)C_{k-1}^* + \eta C_k^*) \frac{\Theta_k^* - \Theta_{k-1}^*}{\Delta t_k} = \\ = (1-\eta)(Q_{k-1}^* - \Lambda_{k-1}^* \Theta_{k-1}^*) + \eta(Q_k^* - \Lambda_k^* \Theta_k^*), \end{aligned} \quad (11.34)$$

где $\eta \in [0, 1]$. Индексы $k-1$ и k в обозначениях матриц C^* и Λ^* указывают на то, что элементы этих матриц в общем случае

не остаются постоянными. В частности, они могут изменяться при изменении температуры в силу возможной зависимости от нее теплоемкости и коэффициента теплопроводности тела.

В предельных случаях $\eta = 0$ и $\eta = 1$ (11.34) соответствует *явной* и *неявной разностным схемам* аппроксимации (11.33). Но в отличие от метода конечных разностей (МКР) выбор $\eta = 0$ не приводит к распаду (11.34) на независимые уравнения относительно неизвестных узловых значений температуры. Сохраняется лишь преимущество явной схемы, связанное с возможностью решения нелинейных задач без итераций на каждом интервале времени, поскольку после обращения матрицы C_{k-1}^* можно разрешить (11.34) явно относительно искомой матрицы-столбца

$$\Theta_k^* = \Theta_{k-1}^* + (C_{k-1}^*)^{-1} (Q_{k-1}^* - \Lambda_{k-1} \Theta_{k-1}^*) \Delta t_k.$$

В правой части этого равенства все параметры известны или могут быть вычислены по найденным на предыдущем интервале времени (при $k = 1$ — по начальным) узловым значениям температуры, составляющим матрицу-столбец Θ_{k-1}^* . Следует иметь в виду, что использование явной схемы накладывает ограничение на выбор Δt_k из условия устойчивости (см. 8.3).

При $\eta = 1$ (11.34) принимает вид

$$\Theta_k^* = \left(\frac{C_k^*}{\Delta t_k} + \Lambda_k \right)^{-1} \left(C_{k-1}^* \frac{\Theta_{k-1}^*}{\Delta t_k} + Q_k^* \right). \quad (11.35)$$

Выбирая Δt_k , можно руководствоваться лишь соображениями точности расчета, поскольку неявная схема устойчива при любых значениях Δt_k . В случае существенной зависимости теплоемкости и коэффициента теплопроводности тела от температуры вычисления по (11.35) на каждом интервале времени приходится проводить несколько раз, последовательно уточняя значения элементов матриц C_k^* и Λ_k^* . При слабой зависимости этих элементов от температуры обычно достаточно ограни-

читься лишь первым приближением, приняв

$$\Theta_k^* = \left(\frac{C_{k-1}^*}{\Delta t_k} + \Lambda_{k-1} \right)^{-1} \left(C_{k-1}^* \frac{\Theta_{k-1}^*}{\Delta t_k} + Q_k^* \right). \quad (11.36)$$

Отметим, что в случае нелинейных граничных условий и зависимости объемной мощности $I_V^{(q)}$ энерговыделения от температуры элементы матрицы-столбца Q_k^* также заранее неизвестны при расчете на k -м интервале и возникает необходимость их последовательного уточнения в (11.35) и (11.36).

Если в (11.24) приближенно принять

$$\frac{\partial T(t, M)}{\partial t} \approx \frac{T_k(M) - T_{k-1}(M)}{\Delta t_k}, \quad M \in V,$$

рассматривая правую часть (11.24) в момент времени t_k , то в соответствии с *методом прямых* получим последовательность краевых задач для дифференциально-разностного уравнения

$$\nabla(\lambda(M) \nabla T_k(M)) - \frac{c(M)}{\Delta t_k} T_k(M) + q_k(M) = 0, \quad M \in V, \quad k \in \mathbb{N},$$

с граничными условиями $T_k(P) = f_1(t_k, P)$, $P \in S_1$, и

$$\lambda(P) \nabla T_k(P) \mathbf{n}(P) + \beta_k(P) T_k(P) = f_k(P), \quad P \in S_2 = S \setminus S_1,$$

где $q_k(M) = I_V^{(q)}(t_k, M) + \frac{c(M) T_{k-1}(M)}{\Delta t_k}$, $\beta_k(P) = \beta(t_k, P)$ и $f_k(P) = f_2(t_k, P)$. Каждой из этих задач соответствует функционал [XV]

$$J[T_k] = \int_V \left(\frac{\lambda}{2} (\nabla T_k)^2 + \frac{c}{2 \Delta t_k} T_k^2 - q_k T_k \right) dV + \int_{S_2} \left(\frac{\beta_k}{2} T_k^2 - f_k T_k \right) dS,$$

минимизируемый на распределениях температуры $T_k(M)$, удовлетворяющих граничному условию $T_k(P) = f_1(t_k, P)$, $P \in S_1$,

непрерывных на замыкании $\bar{V} = V \cup S$ и имеющих кусочно непрерывные производные в области V , ограниченной поверхностью S . Используя процедуру МКЭ, этот функционал можно привести к виду (11.18) и затем из условий его минимума получить СЛАУ вида (11.20), решение которой будет эквивалентно (11.35). Но в случае нелинейной задачи, минимизируя такой функционал, например, методом локальных вариаций, можно непосредственно найти приближенное распределение $T_k(M)$ ($M \in V$) температуры в момент времени t_k .

Выбор $\eta = 1/2$ в (11.34) приводит к *двухслойной симметричной разностной схеме* и более высокому порядку погрешности по Δt_k по сравнению с остальными значениями η . Того же эффекта можно добиться, если для решения задачи нестационарной теплопроводности использовать КЭ в пространственно-временной области*, однако и в том, и в другом случае необходимо уточнять элементы матриц C^* , Λ^* и Q^* при решении нелинейных задач.

Избежать последовательного уточнения элементов этих матриц можно с помощью *явной трехслойной разностной схемы*

$$C_{k-1}^* \frac{\Theta_k^* - \Theta_{k-2}^*}{2\Delta t} + \Lambda_{k-1}^* \frac{\Theta_k^* + \Theta_{k-1}^* + \Theta_{k-2}^*}{3} = Q_{k-1}^*, \quad (11.37)$$

устойчивой при любых значениях Δt . Здесь зависящие от температуры элементы матриц C_{k-1}^* , Λ_{k-1}^* и Q_{k-1}^* вычисляют по известным на k -м интервале времени элементам матрицы-столбца Θ_{k-1}^* , соответствующим моменту времени t_{k-1} в середине удвоенного интервала $2\Delta t = t_k - t_{k-2}$. Поэтому (11.37) можно явно разрешить относительно искомой матрицы-столбца Θ_k^* . Отметим, что при $k = 1$ из-за отсутствия информации об элементах матрицы-столбца Θ_{k-2}^* для нахождения Θ_k^* придется использовать одну из двухслойных разностных схем.

*См., например: Зарубин В.С., Селиванов В.В.

11.3. Двумерное течение вязкой жидкости

Метод конечных элементов (МКЭ) можно применить к решению задач, математическая формулировка которых включает несколько искомых функций и в нее входит более чем одно дифференциальное уравнение. В качестве примера, связанного с решением системы уравнений, рассмотрим достаточно медленное установившееся двумерное течение однородной несжимаемой вязкой жидкости параллельно координатной плоскости $x_1 O x_2$ (см. 3.2) в области F , ограниченной кусочно гладким контуром Γ . Такое движение описывает система трех уравнений вида (3.22)

$$b_i - \frac{\partial p}{\partial x_i} + \eta \sum_{j=1}^2 \frac{\partial^2 v_i}{\partial x_j^2} = 0, \quad i = 1, 2; \quad \sum_{i=1}^2 \frac{\partial v_i}{\partial x_i} = 0, \quad (11.38)$$

где b_i и v_i — проекции на оси Ox , векторов плотности и скорости объемных сил соответственно; p — давление; η — коэффициент сдвиговой вязкости жидкости.

Разобьем область F на E лагранжевых конечных элементов (КЭ), образующих сетку КЭ с общим числом N_Σ узлов. На этой сетке введем систему $\{\varphi_N\}_{N_\Sigma}$ функций вида (11.17)

$$\varphi_N(M) = \sum_{e=1}^E \sum_{n=1}^{N_e} \Omega_{nN}^{(e)} \varphi_n^{(e)}(M), \quad N = \overline{1, N_\Sigma}, \quad (11.39)$$

где $\varphi_n^{(e)}(M)$, $n = \overline{1, N_e}$, — функции формы КЭ с номером $e = \overline{1, E}$ и числом узлов N_e ; $\Omega_{nN}^{(e)}$ — элементы матрицы Ω_e размера $N_e \times N_\Sigma$, устанавливающей соответствие между номерами узлов в отдельном КЭ и в сетке КЭ (см. 10.3).

Приняв (11.39) в качестве проекционных функций, в соответствии с условиями (6.76) ортогональности невязки N_Σ -мерному подпространству $\mathcal{H}_{N_\Sigma}$ гильбертова пространства $\mathcal{H}$ непрерывных на $\overline{F}$ функций, имеющих в F кусочно непрерывные

производные, умножим (11.38) на функцию $\varphi_L(M)$, $L = \overline{1, N_\Sigma}$, и проинтегрируем по области F :

$$\int_F \varphi_L(M) \left(b_i(M) - \frac{\partial p(M)}{\partial x_i} + \eta \nabla^2 v_i(M) \right) dF = 0, \quad i = 1, 2, \quad (11.40)$$

$$\int_F \varphi_L(M) \sum_{i=1}^2 \frac{\partial v_i(M)}{\partial x_i} dF = 0, \quad L = \overline{1, N_\Sigma}. \quad (11.41)$$

Здесь ∇^2 — двумерный оператор Лапласа в плоскости $x_1 O x_2$. Функции $\varphi_L(M)$ имеют в области F кусочно непрерывные производные по координатам x_i , $i = 1, 2$. Поэтому, используя первую формулу Грина, запишем

$$\begin{aligned} \int_F \varphi_L(M) \nabla^2 v_i(M) dF &= \int_\Gamma \varphi_L(P) (\nabla v_i(P)) \mathbf{n}(P) d\Gamma - \\ &- \int_F (\nabla \varphi_L(M)) \nabla v_i(M) dF, \quad i = 1, 2, \end{aligned}$$

где ∇ и $\mathbf{n}(P)$ — двумерный оператор Гамильтона и единичный вектор внешней нормали в точках $P \in \Gamma$ контура Γ соответственно. Тогда вместо (11.40) получим

$$\begin{aligned} \int_F \left(\varphi_L(M) \left(b_i(M) - \frac{\partial p(M)}{\partial x_i} \right) - \eta (\nabla \varphi_L(M)) \nabla v_i(M) \right) dF + \\ + \eta \int_\Gamma \varphi_L(P) (\nabla v_i(P)) \mathbf{n}(P) d\Gamma = 0, \quad i = 1, 2, \quad L = \overline{1, N_\Sigma}. \quad (11.42) \end{aligned}$$

Распределения давления и проекций скорости в области F аппроксимируем при помощи их узловых значений p_N и v_{Ni} соотношениями

$$p(M) = \sum_{N=1}^{N_\Sigma} p_N \varphi_N(M), \quad v_i(M) = \sum_{N=1}^{N_\Sigma} v_{Ni} \varphi_N(M), \quad i = 1, 2,$$

которые подставим в (11.41)

$$\int_F \varphi_L(M) \sum_{i=1}^2 v_{N_i} \sum_{N=1}^{N_\Sigma} \frac{\partial \varphi_N(M)}{\partial x_i} dF = 0, \quad i = 1, 2, \quad L = \overline{1, N_\Sigma},$$

и (11.42)

$$\begin{aligned} & \int_F \varphi_L(M) \left(b_i(M) - p_N \sum_{N=1}^{N_\Sigma} \frac{\partial \varphi_N(M)}{\partial x_i} \right) dF - \\ & - \eta \int_F (\nabla \varphi_L(M)) v_{N_i} \sum_{N=1}^{N_\Sigma} \nabla \varphi_N(M) dF + \\ & + \eta \int_\Gamma \varphi_L(P) \sum_{N=1}^{N_\Sigma} (\nabla v_i(P)) n(P) d\Gamma = 0, \quad i = 1, 2, \quad L = \overline{1, N_\Sigma}. \end{aligned}$$

Отсюда следует система линейных алгебраических уравнений (СЛАУ) в матричной записи

$$DU = Q, \quad (11.43)$$

где U — матрица-столбец размера $3N_\Sigma \times 1$, элементами которой являются

$$U_{3(N-1)+i} = v_{N_i}, \quad U_{3N} = p_N, \quad i = 1, 2, \quad N = \overline{1, N_\Sigma}.$$

Матрица-столбец Q имеет тот же размер и элементы

$$Q_{3(L-1)+i} = \int_F b_i(M) \varphi_L(M) dF + \eta \int_\Gamma \varphi_L(P) (\nabla v_i(P)) n(P) d\Gamma$$

и $Q_{3L} = 0$, $i = 1, 2$, $L = \overline{1, N_\Sigma}$. Элементами симметрической матрицы D порядка $3N_\Sigma$ являются

$$D_{3(L-1)+i, 3(N-1)+j} = \eta \delta_{ij} \int_F \sum_{k=1}^2 \frac{\partial \varphi_L(M)}{\partial x_k} \frac{\partial \varphi_N(M)}{\partial x_k} dF,$$

$$D_{3(L-1)+i, 3N} = D_{3L, 3(N-1)+i} = \int_F \varphi_L(M) \frac{\partial \varphi_N(M)}{\partial x_i} dF$$

и $D_{3L, 3N} = 0$, $i, j = 1, 2$, $L, N = \overline{1, N_\Sigma}$, где $\delta_{ij} = 1$ при $i = j$ и $\delta_{ij} = 0$ при $i \neq j$.

Отметим, что второй интеграл в выражении для $Q_{3(L-1)+i}$ отличен от нуля, если узел с номером L принадлежит участку границы Γ , на котором заданы значения $(\nabla u_i) \mathbf{n}$. Так как в выражениях для элементов матрицы D подынтегральные функции отличны от нуля лишь при условии, что узлы с номерами $L, N = \overline{1, N_\Sigma}$ принадлежат одному КЭ, то эту матрицу, а заодно и матрицу Q удобнее формировать (аналогично матрицам теплопроводности и теплоемкости) из вкладов отдельных КЭ.

Пусть в простейшем случае на контуре Γ скорость жидкости равна нулю, т.е. $v_{N'_i} = 0$, $i = 1, 2$, во всех N_1 узлах с номерами N' , принадлежащих границе, а в узле с номером N_* задано значение давления p_* (последнее необходимо для однозначного нахождения давления в остальных узлах). Тогда для $N^* = 3N_\Sigma - 2N_1 - 1$ неизвестных узловых значений получим СЛАУ $D^* U^* = Q^*$, в которой симметрическая матрица D^* порядка N^* получена из матрицы D вычеркиванием строк и столбцов с номерами $3(N' - 1) + i$, $i = 1, 2$, и $3N_*$, а матрица-столбец U^* размера $N^* \times 1$ — вычеркиванием строк с указанными номерами из матрицы U . Для формирования матрицы-столбца Q^* размера $N^* \times 1$ предварительно следует составить матрицу-столбец $\tilde{Q}$ размера $N^* \times 1$ с элементами

$$\tilde{Q}_{3(L-1)+i} = \int_F \varphi_L(M) \left(b_i(M) - p_* \frac{\partial \varphi_{N_*}(M)}{\partial x_i} \right) dF$$

и $\tilde{Q}_{3L} = 0$, $i = 1, 2$, $L = \overline{1, N_\Sigma}$, а затем вычеркнуть из нее строки с указанными выше номерами.

В более общем случае могут возникнуть трудности при задании корректных граничных условий для давления (см. 3.2). Тогда для решения задачи целесообразно использовать равносильное системе (11.38) уравнение вида (3.27)

$$\nabla^2(\nabla^2\psi) = \frac{1}{\eta} \left(\frac{\partial b_2}{\partial x_1} - \frac{\partial b_1}{\partial x_2} \right) \quad (11.44)$$

относительно функции тока $\psi(M)$, $M \in F$. Проследим путь применения МКЭ к решению краевой задачи, описываемой (11.44) с характерными для функции тока граничными условиями вида (3.29):

$$\psi(P) = f_0(P), \quad \frac{\partial \psi(P)}{\partial n(P)} = f_1(P), \quad P \in \Gamma_1 \subset \Gamma; \quad (11.45)$$

$$\psi(P) = \psi^0 = \text{const}, \quad \frac{\partial^2 \psi(P)}{\partial n^2(P)} = 0, \quad P \in \Gamma_2 = \Gamma \setminus \Gamma_1, \quad (11.46)$$

где $n(P)$ — направление внешней нормали к границе Γ в точке $P \in \Gamma$; $f_0(P)$ и $f_1(P)$ — заданные функции положения точки на этой границе, а число ψ^0 положим равным нулю.

Краевой задаче (11.44)–(11.46) соответствует функционал*

$$J[\psi] = \frac{1}{2} \int_F (\nabla^2 \psi)^2 dF - \int_F f \psi dF, \quad (11.47)$$

где функция $f(M)$, $M \in F$, равна правой части (11.44), минимизируемый на функциях, непрерывно дифференцируемых на замыкании $\overline{F} = F \cup \Gamma$, имеющих кусочно непрерывные вторые производные в области F и удовлетворяющих (11.45). Убедимся, что стационарная точка этого функционала является

*См., например: Ректорис К.

решением рассматриваемой краевой задачи. Для этого вычислим вариацию функционала (11.47), вызванную вариацией $\delta\psi$:

$$\delta J[\psi, \delta\psi] = \int_F (\nabla^2 \psi) \nabla^2 \delta\psi dF - \int_F f \delta\psi dF. \quad (11.48)$$

Преобразуем первый интеграл в (11.48) при помощи *второй формулы Грина*, положив в ней $u = \nabla^2 \psi$ и $v = \delta\psi$:

$$\begin{aligned} \int_F (\nabla^2 \psi) \nabla^2 \delta\psi dF &= \int_F \nabla^2 (\nabla^2 \psi) \delta\psi dF + \\ &+ \int_{\Gamma} (\nabla^2 \psi) (\nabla \delta\psi) \mathbf{n} d\Gamma - \int_{\Gamma} \nabla (\nabla^2 \psi) \mathbf{n} \delta\psi d\Gamma. \end{aligned} \quad (11.49)$$

Последний интеграл в правой части (11.49) исчезает, так как на контуре значения ψ в соответствии с (11.45) и (11.46) заданы и поэтому $\delta\psi \equiv 0$.

Второй интеграл в правой части (11.49) представим в виде

$$\int_{\Gamma} (\nabla^2 \psi) (\nabla \delta\psi) \mathbf{n} d\Gamma = \int_{\Gamma_1} (\nabla^2 \psi) \delta(\nabla \psi) \mathbf{n} d\Gamma + \int_{\Gamma_2} (\nabla^2 \psi) \delta(\nabla \psi) \mathbf{n} d\Gamma.$$

Здесь интеграл по участкам Γ_1 контура будет равен нулю, поскольку на этих участках в соответствии с (11.45) заданы значения $(\nabla \psi) \mathbf{n}$ и поэтому $\delta(\nabla \psi) \mathbf{n} \equiv 0$. В итоге с учетом (11.48) и (11.49) условие стационарности функционала (11.47) примет вид

$$\delta J[\psi, \delta\psi] = \int_F (\nabla^2 (\nabla^2 \psi) - f) \delta\psi dF + \int_{\Gamma_2} (\nabla^2 \psi) \delta(\nabla \psi) \mathbf{n} d\Gamma = 0.$$

При произвольных вариациях $\delta\psi$ в области F и на участках Γ_2 ее границы это условие будет выполнено, если ψ удовлетворяет (11.44) и (11.46), т.е. является решением рассматриваемой краевой задачи.

Функционал (11.47) будет строго выпуклым, если для любых функций $\psi_1 \neq \psi_2$, на которых его допустимо рассматривать, при $\sigma \in (0, 1)$ выполнено строгое неравенство [XV]

$$R(\psi_1, \psi_2) = \sigma J[\psi_1] + (1-\sigma)J[\psi_2] - J[\sigma\psi_1 + (1-\sigma)\psi_2] > 0. \quad (11.50)$$

При исследовании функционала на выпуклость в его представлении можно опустить линейное относительно функции ψ слагаемое [XV]. Тогда с учетом (11.47) и (11.50) получим

$$\begin{aligned} R(\psi_1, \psi_2) &= \frac{\sigma}{2} \int_F (\nabla^2 \psi_1)^2 dF + \frac{1-\sigma}{2} \int_F (\nabla^2 \psi_2)^2 dF - \\ &\quad - \int_F (\sigma \nabla^2 \psi_1 + (1-\sigma) \nabla^2 \psi_2)^2 dF = \\ &= \sigma \frac{1-\sigma}{2} \int_F (\nabla^2 \psi_1 - \nabla^2 \psi_2)^2 dF = \sigma \frac{1-\sigma}{2} \int_F (\nabla^2 (\psi_1 - \psi_2))^2 dF > 0, \end{aligned}$$

поскольку неравенство значений функций ψ_1 и ψ_2 хотя бы в одной точке $M \in F$ в силу непрерывности этих функций означает, что $\psi_1 - \psi_2 \neq 0$ в некоторой окрестности такой точки.

Так как вариация (11.48) строго выпуклого функционала (11.47) линейна относительно $\delta\psi$ всюду в его области определения, то всюду в этой области существует его дифференциал Гато. Кроме того, функционал (11.47) имеет единственную стационарную точку [XV], в которой он достигает наименьшего значения.

Отсюда следует, что краевая задача (11.44)–(11.46) имеет единственное решение, а условие минимума функционала (11.47) можно использовать в качестве интегральной формулировки этой задачи. Так как функционал (11.47) допустимо рассматривать на функциях, непрерывно дифференцируемых на замыкании $\bar{F}$, то при разбиении области F необходимо использовать сетку КЭ, обеспечивающую гладкую аппроксимацию функций не только в пределах отдельных элементов, но

и при переходе через их границы. Это обстоятельство требует применения эрмитовых КЭ, узловыми параметрами которых наряду со значениями искомой функции являются и значения ее производных. В частности, можно использовать треугольные КЭ с шестью узлами и функциями формы в виде многочленов пятой степени (см. 10.2), что позволит при решении задачи получить не только значения функции тока во всех узлах, но и значения проекции скорости в вершинах треугольников, выражаемые через производные функции тока по координатам x_i , $i = 1, 2$ (см. 3.2).

Отметим, что процедура применения МКЭ при использовании эрмитовых КЭ остается прежней. После разбиения области F на E эрмитовых КЭ представим функционал (11.47) в виде

$$J[\psi] = \sum_{e=1}^E J_e[\psi] = \frac{1}{2} \sum_{e=1}^E \int_{F_e} (\nabla^2 \psi)^2 dF - \sum_{e=1}^E \int_{F_e} f \psi dF, \quad (11.51)$$

где $J_e[\psi]$ — вклад в функционал КЭ с номером $e = \overline{1, E}$ и площадью F_e . Распределение функции тока в пределах отдельного КЭ с номером e аппроксимируем соотношением

$$\psi(M) = \Psi_e^T \Phi^{(e)}(M), \quad (11.52)$$

где Ψ_e и $\Phi^{(e)}(M)$ — матрицы-столбцы размера $N'_e \times 1$, элементами которых являются узловые параметры и функции формы $\varphi_n^{(e)}(M)$, $n = \overline{1, N'_e}$, КЭ соответственно, N'_e — общее число узловых параметров (а не узлов!) КЭ. С учетом (11.52) имеем

$$\frac{\partial^2 \psi}{\partial x_i^2} = \Psi_e^T \tilde{B}_i^{(e)}, \quad (\nabla^2 \psi)^2 = \Psi_e^T \tilde{B}_e (\Psi_e^T \tilde{B}_e)^T, \quad i = 1, 2, \quad (11.53)$$

где $\tilde{B}_i^{(e)}$ — матрица-столбец размера $N'_e \times 1$ с элементами $\frac{\partial^2 \varphi_n^{(e)}(M)}{\partial x_i^2}$, а $\tilde{B}_e$ — блочная матрица размера $N'_e \times 2$, блоками которой являются матрицы $\tilde{B}_i^{(e)}$.

Используя (11.52) и второе соотношение (11.53), для вклада КЭ с номером $e = \overline{1, E}$ в функционал $J[\psi]$ получаем

$$\begin{aligned} J_e[\psi] &\approx \frac{1}{2} \sum_{e=1}^E \int_{F_e} \Psi_e^T \tilde{B}_e (\Psi_e^T \tilde{B}_e)^T dF - \int_{F_e} f \Psi_e^T \Phi_e dF = \\ &= \frac{1}{2} \Psi_e^T \tilde{A}^{(e)} \Psi_e - \Psi_e^T \hat{Q}^{(e)}, \quad (11.54) \end{aligned}$$

где $A^{(e)}$ — симметрическая матрица порядка N'_e с элементами

$$A_{ln}^{(e)} = \int_{F_e} \sum_{i=1}^2 \frac{\partial^2 \varphi_l^{(e)}(M)}{\partial x_i^2} \frac{\partial^2 \varphi_n^{(e)}(M)}{\partial x_i^2} dF, \quad l, n = \overline{1, N'_e},$$

а $\hat{Q}^{(e)}$ — матрица-столбец размера $N'_e \times 1$ с элементами

$$\hat{Q}_l^{(e)} = \int_{F_e} f(M) \varphi_l^{(e)}(M) dF.$$

В (11.54) перейдем от локальной нумерации узлов к глобальной (см. 10.3) при помощи матриц Ω_e размера $N_e \times N'_e$, $e = \overline{1, E}$, где N'_e — общее число узловых параметров (а не узлов!) сетки КЭ. Тогда после подстановки (11.54) в (11.51) получим

$$J[\psi] \approx \frac{1}{2} \Psi^T \hat{A} \Psi - \Psi^T \hat{Q}, \quad (11.55)$$

где Ψ — матрица-столбец размера $N'_\Sigma \times 1$ значений функции тока и ее производных в узлах сетки КЭ; $\hat{A}$ — симметрическая матрица порядка N'_Σ с элементами

$$\hat{A}_{LN} = \sum_{e=1}^E \sum_{l=1}^{N_e} \sum_{n=1}^{N_e} \Omega_{lL}^{(e)} A_{ln}^{(e)} \Omega_{nN}^{(e)}, \quad (11.56)$$

причем $\Omega_{nN}^{(e)}$ — элементы матрицы Ω_e ; $\hat{Q}$ — матрица-столбец размера $N'_\Sigma \times 1$ с элементами

$$\hat{Q}_N = \sum_{e=1}^E \sum_{n=1}^{N_e} \Omega_{nN}^{(e)} \hat{Q}_n^{(e)}. \quad (11.57)$$

Правая часть (11.55) является конечномерным аналогом $J[\tilde{\psi}_{N'_\Sigma}]$ функционала (11.47), полученным при использовании метода Рунца и соответствующим элементу

$$\tilde{\psi}_{N'_\Sigma}(M) = \Psi^T \Phi(M) = \sum_{N=1}^{N_\Sigma} \psi_N \varphi_N(M) \quad (11.58)$$

последовательности, минимизирующей этот функционал, причем $\Phi(M)$ — матрица-столбец размера $N'_\Sigma \times 1$, элементами которой являются функции

$$\varphi_N(M) = \sum_{e=1}^E \sum_{n=1}^{N_e} \Omega_{nN}^{(e)} \varphi_n^{(e)}(M), \quad N = \overline{1, N'_\Sigma}. \quad (11.59)$$

Эти функции определены на сетке КЭ и образуют базис конечномерного функционального пространства.

Ясно, что правую часть (11.55) можно представить многочленом второй степени относительно N'_Σ переменных ψ_N :

$$J[\tilde{\psi}_{N'_\Sigma}] = \frac{1}{2} \sum_{L=1}^{N_\Sigma} \sum_{N=1}^{N'_\Sigma} \hat{A}_{LN} \psi_L \psi_N - \sum_{L=1}^{N'_\Sigma} \hat{Q}_L \psi_L. \quad (11.60)$$

Так как минимизируемый функционал (11.47) ограничен снизу [XV], то функция (11.60) также ограничена снизу, поскольку в этом случае для минимального значения функционала, достигаемого на функции $\psi_*(M)$, $M \in \overline{F}$, справедливо неравенство $J[\psi_*] \leq J[\tilde{\psi}_{N'_\Sigma}]$. Поэтому функция (11.60) также достигает минимума при некотором наборе значений ψ_N , $N = \overline{1, N'_\Sigma}$.

В общем случае часть узловых значений функции тока и ее первых производных в (11.60) может быть задана, если существуют узлы, принадлежащие участкам Γ_1 границы, поскольку (11.58) должно удовлетворять граничным условиям (11.45) на Γ_1 . Пусть число заданных узловых значений в таких узлах равно N'_1 , а число искоемых узловых значений составляет $N^* = N'_\Sigma - N'_1$ и их номера укорячены так, что $N = \overline{1, N^*}$. Тогда из необходимых условий $\frac{\partial J[\tilde{\psi}_{N_\Sigma}]}{\partial \psi_L} = 0$, $L = \overline{1, N^*}$, минимума функции (11.60) с учетом $\hat{A}_{LN} = \hat{A}_{NL}$ получим СЛАУ

$$\sum_{N=1}^{N^*} \hat{A}_{LN} \psi_N = \hat{Q}_L - \sum_{N'=N^*+1}^{N'_\Sigma} \hat{A}_{LN'} \psi_{N'}, \quad L = \overline{1, N^*}. \quad (11.61)$$

Если же участки Γ_1 границы отсутствуют, то $N_* = N'_\Sigma$ и сумма в правой части (11.61) исчезает.

Отметим, что матрица СЛАУ (11.61) и симметрическая матрица D СЛАУ (11.43) являются *слабозаполненными*. Поэтому для численного решения СЛАУ с такими матрицами целесообразно использовать способы, позволяющие сокращать число арифметических операций и экономить память ЭВМ*.

11.4. Задачи теории упругости

Переменное во времени t векторное поле малых перемещений $u(t, M)$, $M \in V$, в однородном линейно упругом теле, занимающем область V , можно описать *уравнениями Ламе* в векторной форме (3.35), если их левую часть заменить на $\rho^0 \frac{\partial^2 u}{\partial t^2}$ (см. 3.3), где $\rho^0 = \text{const}$ — плотность материала тела. Тогда в проекциях на оси Ox_i , $i = \overline{1, 3}$, декартовой прямоугольной системы координат эти уравнения примут вид

$$\rho^0 \frac{\partial^2 u_i}{\partial t^2} = b_i + (\tilde{\lambda} + \tilde{\mu}) \frac{\partial \Theta}{\partial x_i} + \tilde{\mu} \nabla^2 u_i, \quad i = \overline{1, 3}. \quad (11.62)$$

*См.: Джордж А., Лю Дж., а также: Писсанецки С'

Здесь u_i и b_i — проекции на координатные оси векторов перемещений и плотности объемных сил соответственно; $\tilde{\lambda}$ и $\tilde{\mu}$ — константы Ламе; $\Theta = \frac{\partial u_i}{\partial x_i}$ — объемная деформация (здесь и далее происходит суммирование слагаемых по повторяющемуся нижнему индексу).

Статическая задача. Рассмотрим сначала статическую задачу теории упругости, положив в (11.62) $\frac{\partial^2 u_i}{\partial t^2} \equiv 0$, что соответствует *уравнениям равновесия* (2.87)

$$\frac{\partial \sigma_{ji}}{\partial x_j} + b_i = 0, \quad i, j = \overline{1, 3}, \quad (11.63)$$

где σ_{ij} — компоненты тензора напряжений, причем $\sigma_{ij} = \sigma_{ji}$. На участках S_1 и $S_2 = S \setminus S_1$ поверхности S , ограничивающей область V (рис. 11.2), примем граничные условия

$$\begin{cases} u_i(P) = u_i^0(P), & P \in S_1; \\ \sigma_{ij}(P) n_j(P) = p_i^0(P), & P \in S_2, \end{cases} \quad (11.64)$$

где $u_i^0(P)$ и $p_i^0(P)$ — заданные функции координат точки P ; $n_j = \mathbf{n} \mathbf{e}_j$ — направляющие косинусы единичного вектора $\mathbf{n}$ внешней нормали к поверхности S_2 , причем $\mathbf{e}_j$, $j = \overline{1, 3}$, — *орты* декартовой прямоугольной системы координат.

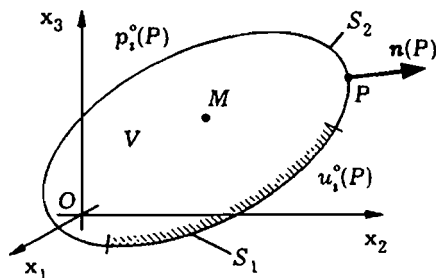


Рис. 11.2

Покажем, что стационарная точка функционала

$$J[u] = \int_V \left(\frac{\tilde{\lambda}}{2} \Theta^2 + \tilde{\mu} \Upsilon \right) dV - \int_V b_i u_i dV - \int_{S_2} p_i^\circ u_i dS, \quad (11.65)$$

где с учетом соотношений Коши (3.30)

$$\Theta = \frac{\partial u_i}{\partial x_i} = \varepsilon_{ii} = \varepsilon_{11} + \varepsilon_{22} + \varepsilon_{33}, \quad (11.66)$$

$$\Upsilon = \frac{1}{4} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) = \varepsilon_{ij} \varepsilon_{ij}, \quad i, j = \overline{1, 3}, \quad (11.67)$$

($\varepsilon_{ij} = \varepsilon_{ji}$, $i, j = \overline{1, 3}$, — компоненты симметрического тензора деформаций), на множестве функций u_i , дважды непрерывно дифференцируемых на замыкании $\bar{V} = V \cup S$ области V и удовлетворяющих первому условию (11.64), является решением рассматриваемой краевой задачи. Действительно, вычисляя вариацию этого функционала, вызванную вариациями δu_i , $i = \overline{1, 3}$, и учитывая равенства $\frac{\partial u_i}{\partial x_i} = \delta_{ij} \frac{\partial u_i}{\partial x_j}$, где δ_{ij} — компоненты единичного тензора второго ранга (т.е. $\delta_{ij} = 1$ при $i = j$ и $\delta_{ij} = 0$ при $i \neq j$), получаем

$$\delta J[u, \delta u_i] = \int_V \left(\tilde{\lambda} \Theta \delta_{ij} + 2\tilde{\mu} \varepsilon_{ij} \right) \frac{\partial \delta u_i}{\partial x_j} dV - \int_V b_i \delta u_i dV - \int_{S_2} p_i^\circ \delta u_i dS.$$

Подынтегральную функцию в первом интеграле, используя соотношения обобщенного закона Гука (3.31), преобразуем к виду

$$\begin{aligned} \left(\tilde{\lambda} \Theta \delta_{ij} + 2\tilde{\mu} \varepsilon_{ij} \right) \frac{\partial \delta u_i}{\partial x_j} &= \sigma_{ij} \frac{\partial \delta u_i}{\partial x_j} = \\ &= \frac{\partial \sigma_{ji} \delta u_i}{\partial x_j} - \frac{\partial \sigma_{ji}}{\partial x_j} \delta u_i = \nabla(\sigma_{ji} e_j \delta u_i) - \frac{\partial \sigma_{ji}}{\partial x_j} \delta u_i, \end{aligned}$$

поскольку $\sigma_{ji} \delta u_i$ можно рассматривать как проекции вектора $\sigma_{ji} e_j \delta u_i$ на оси координат Ox_j , $j = \overline{1, 3}$. В итоге, используя

формулу Гаусса — Остроградского (1.18) и учитывая, что $(\sigma_j, e_j, \delta u_i) n = \sigma_j n_j \delta u_i$, находим

$$\delta J[u, \delta u_i] = \int_{S_1} \sigma_j n_j \delta u_i dS - \int_V \left(\frac{\partial \sigma_j}{\partial x_j} + b_i \right) \delta u_i + \int_{S_2} (\sigma_j n_j - p_i^0) \delta u_i dS.$$

В этом равенстве первый интеграл обращается в нуль, поскольку функции u_i должны удовлетворять первому условию (11.64), так что $\delta u_i(P) = 0$, $i = \overline{1, 3}$, при $P \in S_1$. В стационарной точке u^* функционала (11.65) $J[u^*, \delta u_i] = 0$, $i = \overline{1, 3}$. Поэтому при произвольных вариациях δu_i на участках поверхности S_2 и в области V из равенства нулю второго и третьего интегралов следует, что проекции u_i^* должны удовлетворять (11.63) и второму условию (11.64) соответственно, т.е. функции u_i^* будут решением задачи (11.63), (11.64).

Входящие в функционал (11.65) скалярные величины Θ и Υ являются инвариантами первого и второго порядка [IV] тензора деформаций $\tilde{\epsilon}$, т.е. свертками этого тензора, и не зависят от выбора координатных осей. Первая из них является следом симметрической матрицы третьего порядка, соответствующей $\tilde{\epsilon}$. Эту матрицу ортогональным преобразованием можно привести к диагональному виду с элементами ϵ_1 , ϵ_2 и ϵ_3 на главной диагонали, которые имеют смысл линейных деформаций (удлинений) вдоль осей нового базиса. В этом базисе в соответствии с (11.66), (11.67) будем иметь

$$\Upsilon = \epsilon_{ij} \epsilon_{ij} = \epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2, \quad i, j = \overline{1, 3}. \quad (11.68)$$

Функционал (11.65) будет строго выпуклым, если для любых функций $u_1(M) \neq u_2(M)$ с проекциями $u_i^{(1)}$ и $u_i^{(2)}$, $i = \overline{1, 3}$, удовлетворяющими первому условию (11.64), при $\sigma \in (0, 1)$

$$R(u_1, u_2) = \sigma J[u_1] + (1 - \sigma) J[u_2] - J[\sigma u_1 + (1 - \sigma) u_2] > 0. \quad (11.69)$$

В соответствии с соотношениями Коши (3.30) обозначим

$$\Delta \epsilon_{ij} = \frac{1}{2} \left(\frac{\partial \Delta u_i}{\partial x_j} + \frac{\partial \Delta u_j}{\partial x_i} \right), \quad \Delta u_i = u_i^{(1)} - u_i^{(2)}, \quad i, j = \overline{1, 3}.$$

Тогда с учетом (11.66), (11.67) и (11.69) получим

$$\begin{aligned}
 R(u_1, u_2) &= \frac{\tilde{\lambda}}{2} \int_V \left(\sigma \Theta_1^2 + (1 - \sigma) \Theta_2^2 - (\sigma \Theta_1 + (1 - \sigma) \Theta_2)^2 \right) dV + \\
 &\quad + \tilde{\mu} \int_V \left(\sigma \Upsilon_1 + (1 - \sigma) \Upsilon_2 - \right. \\
 &\quad \left. - (\sigma \varepsilon_{ij}^{(1)} + (1 - \sigma) \varepsilon_{ij}^{(2)}) (\sigma \varepsilon_{ij}^{(1)} + (1 - \sigma) \varepsilon_{ij}^{(2)}) \right) dV = \\
 &= \tilde{\lambda} \sigma \frac{1 - \sigma}{2} \int_V (\Theta_1 - \Theta_2)^2 dV + \tilde{\mu} \sigma (1 - \sigma) \int_V \Delta \varepsilon_{ij} \Delta \varepsilon_{ij} dV > 0,
 \end{aligned}$$

поскольку первый интеграл в правой части неотрицателен, а второй положителен, так как неравенство $u_1(M) \neq u_2(M)$, $M \in V \cup S_2$, означает, что $\Delta u_i(M) = u_i^{(1)}(M) - u_i^{(2)}(M) \neq 0$ хотя бы для одного из значений $i = \overline{1, 3}$, причем $\Delta u_i(P) = 0$, $P \in S_1$, т.е. $\Delta u_i(M)$ не может быть константой. Поэтому хотя бы один из компонентов $\Delta \varepsilon_{ij}$ тензора $\Delta \tilde{\varepsilon}$ в некоторой точке $M \in V$ (а в силу непрерывной дифференцируемости функций u_i и в ее окрестности) будет отличен от нуля. Тогда при ортогональном преобразовании к диагональному виду симметрической матрицы, соответствующей этому тензору, хотя бы один из элементов на главной диагонали будет отличен от нуля, поэтому с учетом (11.68) функционал (11.65) является строго выпуклым на рассматриваемом множестве функций.

Так как вариация $\delta J[u, \delta u_i]$ функционала (11.65) линейна относительно δu_i всюду в области его определения, то у него всюду в этой области существует дифференциал Гато [XV]. Поэтому строго выпуклый функционал (11.65) имеет единственную стационарную точку, в которой он достигает наименьшего значения [XV]. Таким образом, задача (11.63), (11.64) имеет единственное решение, а (11.65) входит в интегральную формулировку этой задачи.

Используем функционал (11.65) для решения задачи (11.63), (11.64) *методом конечных элементов* (МКЭ). Если считать границы между соседними конечными элементами (КЭ) *поверхностями слабого разрыва* функции перемещений, то этот функционал можно рассматривать на множестве непрерывных функций u_i , непрерывно дифференцируемых лишь в пределах отдельных КЭ (см. 6.10). Поэтому возможно использование *лагранжевых КЭ*, в том числе *симплексных*.

Разобьем область V на E лагранжевых КЭ, каждый из которых имеет N_e узлов и занимает область V_e , $e = \overline{1, E}$. Общее число узлов образованной таким образом *сетки КЭ* обозначим N_Σ . Тогда вместо (11.65) с учетом свойства аддитивности интеграла по области V получим

$$J[u] = \sum_{e=1}^E J_e[u], \quad (11.70)$$

причем вклад КЭ с номером e в значение функционала $J[u]$

$$J_e[u] = \int_{V_e} \left(\frac{\tilde{\lambda}}{2} (\Theta)^2 + \tilde{\mu} \Upsilon - b_i u_i \right) dV - \int_{S'_e} p_i^\circ u_i dS. \quad (11.71)$$

Второй интеграл в правой части (11.71) отличен от нуля лишь при условии, что граница данного элемента имеет общие участки с поверхностью S_2 , т.е. $S'_e = S_e \cap S_2 \neq \emptyset$.

В пределах каждого КЭ с номером e векторное поле перемещений $u(M)$, $M \in V_e$, приближенно представим матричным выражением вида (10.53):

$$U_e(M) = (U^{(e)})^T \Phi_e(M), \quad M \in V_e, \quad (11.72)$$

где $U^{(e)}$ — матрица размера $N_e \times 3$ (в случае трехмерной задачи) с элементами $u_{n_i}^{(e)}$, $n = \overline{1, N_e}$, $i = \overline{1, 3}$, равными искомым узловым значениям проекций $u_i(M)$ вектора перемещений $u(M)$;

$U_e(M)$ — матрица-столбец размера 3×1 , элементами которой являются координатные функции $u_i^{(e)}(M)$ векторной функции $u^{(e)}(M)$, $M \in V_e$, аппроксимирующей в этом КЭ функцию $u(M)$; $\Phi_e(M)$ — матрица-столбец размера $N_e \times 1$, элементами которой являются функции формы $\varphi_n^{(e)}(M)$ КЭ, зависящие от координат $x_i^{(e)}(M)$ точки $M \in V_e$.

Выразим вклад $J_e[u]$ КЭ с номером e в функционал (11.65) через искомые узловые значения $u_{ni}^{(e)}$. В соответствии с (10.52) при $i = \overline{1, 3}$ и $n = \overline{1, N_e}$

$$u_i^{(e)}(M) = (U_i^{(e)})^T \Phi_e(M) = u_{ni}^{(e)} \varphi_n^{(e)}(M), \quad M \in V_e, \quad (11.73)$$

где $U_i^{(e)}$ — матрица-столбец размера $N_e \times 1$, являющаяся одним из трех блоков матрицы $U^{(e)}$. Аналогично в пределах КЭ

$$b_i(M) = b_{ni}^{(e)} \varphi_n^{(e)}(M), \quad p_i(M) = p_{ni}^{(e)} \varphi_n^{(e)}(M), \quad (11.74)$$

где $b_{ni}^{(e)}$ и $p_{ni}^{(e)}$, $n = \overline{1, N_e}$, $i = \overline{1, 3}$, — узловые значения проекций векторов плотности объемных и поверхностных сил соответственно.

Используя (10.52), (10.57), (11.66), (11.67) и (11.73), для общего случая зависимости $\tilde{\lambda}(M)$ и $\tilde{\mu}(M)$ от положения точки $M \in V_e$ находим (опуская обозначение этой точки и применяя правило суммирования по повторяющимся нижним индексам $i, j = \overline{1, 3}$, $m, n, q = \overline{1, N_e}$)

$$\frac{\tilde{\lambda}}{2} \Theta^2 = \frac{\tilde{\lambda}_q^{(e)} \varphi_q^{(e)}}{2} u_{mi}^{(e)} \frac{\partial \varphi_m^{(e)}}{\partial x_i} u_{nj}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j},$$

$$\tilde{\mu} \Upsilon = \frac{\tilde{\mu}_q^{(e)} \varphi_q^{(e)}}{4} \left(u_{mi}^{(e)} \frac{\partial \varphi_m^{(e)}}{\partial x_j} + u_{mj}^{(e)} \frac{\partial \varphi_m^{(e)}}{\partial x_i} \right) \left(u_{ni}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} + u_{nj}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_i} \right).$$

Подставляя эти соотношения, а также соотношения (11.73), (11.74) в равенство (11.71), для вклада КЭ с номером e в

функционал (11.65) получаем

$$\begin{aligned}
 J_e[\mathbf{u}] \approx & \frac{1}{2} \tilde{\lambda}_q^{(e)} u_{m_i}^{(e)} u_{n_j}^{(e)} \int_{V_e} \frac{\partial \varphi_m^{(e)}}{\partial x_i} \varphi_q^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} dV + \\
 & + \frac{1}{4} \tilde{\mu}_q^{(e)} \int_{V_e} \varphi_q^{(e)} \left(u_{m_i}^{(e)} \frac{\partial \varphi_m^{(e)}}{\partial x_j} + u_{m_j}^{(e)} \frac{\partial \varphi_m^{(e)}}{\partial x_i} \right) \left(u_{n_i}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} + u_{n_j}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_i} \right) dV - \\
 & - b_{n_i}^{(e)} u_{m_i}^{(e)} \int_{V_e} \varphi_m^{(e)} \varphi_n^{(e)} dV - p_{n_i}^{(e)} u_{m_i}^{(e)} \int_{S'_e} \varphi_m^{(e)} \varphi_n^{(e)} dS. \quad (11.75)
 \end{aligned}$$

Из (11.75) следует, что $J_e[\mathbf{u}]$ аппроксимируется квадратичной функцией $3N_e$ переменных $u_{n_i}^{(e)}$. Это позволяет вариацию функционала (11.65), используя (11.70), представить в виде

$$\delta J[\mathbf{u}, \delta \mathbf{u}] = \sum_{e=1}^E \delta J_e[\mathbf{u}, \delta \mathbf{u}] = \sum_{e=1}^E \frac{\partial J_e[\mathbf{u}]}{\partial u_{lk}^{(e)}} \delta u_{lk}^{(e)}, \quad k = \overline{1, 3}. \quad (11.76)$$

Учитывая, что $\frac{\partial u_{m_j}^{(e)}}{\partial u_{lk}^{(e)}} = \delta'_{lm} \delta_{jk}$, где $\delta'_{lm} = 1$ при $l = m$ и $\delta'_{lm} = 0$ при $l \neq m$, $l, m = \overline{1, N_e}$, из (11.75) находим

$$\begin{aligned}
 \frac{\partial J_e[\mathbf{u}]}{\partial u_{lk}^{(e)}} \approx & \frac{\tilde{\lambda}_q^{(e)}}{2} (\delta'_{lm} \delta_{ik} u_{n_j}^{(e)} + u_{m_i}^{(e)} \delta'_{ln} \delta_{jk}) \int_{V_e} \frac{\partial \varphi_m^{(e)}}{\partial x_i} \varphi_q^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} dV + \\
 & + \frac{\tilde{\mu}_q^{(e)}}{4} \int_{V_e} \varphi_q^{(e)} \left(\delta'_{lm} \delta_{ik} \frac{\partial \varphi_m^{(e)}}{\partial x_j} + \delta'_{lm} \delta_{jk} \frac{\partial \varphi_m^{(e)}}{\partial x_i} \right) \left(u_{n_i}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} + u_{n_j}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_i} \right) dV + \\
 & + \frac{\tilde{\mu}_q^{(e)}}{4} \int_{V_e} \varphi_q^{(e)} \left(u_{m_i}^{(e)} \frac{\partial \varphi_m^{(e)}}{\partial x_j} + u_{m_j}^{(e)} \frac{\partial \varphi_m^{(e)}}{\partial x_i} \right) \left(\delta'_{ln} \delta_{ik} \frac{\partial \varphi_n^{(e)}}{\partial x_j} + \delta'_{ln} \delta_{jk} \frac{\partial \varphi_n^{(e)}}{\partial x_i} \right) dV - \\
 & - b_{n_i}^{(e)} \delta'_{lm} \delta_{ik} \int_{V_e} \varphi_m^{(e)} \varphi_n^{(e)} dV - p_{n_i}^{(e)} \delta'_{lm} \delta_{ik} \int_{S'_e} \varphi_m^{(e)} \varphi_n^{(e)} dS. \quad (11.77)
 \end{aligned}$$

Так как $\varphi_m^{(e)} \delta_{lm} = \varphi_n^{(e)} \delta_{ln} = \varphi_l^{(e)}$, а результат суммирования по повторяющемуся индексу не зависит от обозначения этого индекса, то второй и третий интегралы в правой части (11.77) равны. Кроме того, $\delta_{ik} \frac{\partial \varphi_n^{(e)}}{\partial x_i} = \frac{\partial \varphi_n^{(e)}}{\partial x_k}$, а также

$$\begin{aligned} & (\delta'_{lm} \delta_{ik} u_{nj}^{(e)} + u_{mi}^{(e)} \delta'_{ln} \delta_{jk}) \frac{\partial \varphi_m^{(e)}}{\partial x_i} \frac{\partial \varphi_n^{(e)}}{\partial x_j} = \\ & = u_{nj}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} \frac{\partial \varphi_l^{(e)}}{\partial x_k} + u_{mi}^{(e)} \frac{\partial \varphi_m^{(e)}}{\partial x_i} \frac{\partial \varphi_l^{(e)}}{\partial x_k} = 2 u_{nj}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} \frac{\partial \varphi_l^{(e)}}{\partial x_k} \end{aligned}$$

и

$$\begin{aligned} & \left(\delta'_{lm} \delta_{ik} \frac{\partial \varphi_m^{(e)}}{\partial x_j} + \delta'_{lm} \delta_{jk} \frac{\partial \varphi_m^{(e)}}{\partial x_i} \right) \left(u_{ni}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} + u_{nj}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_i} \right) = \\ & = \frac{\partial \varphi_l^{(e)}}{\partial x_j} \left(u_{nk}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} + u_{nj}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_k} \right) + \frac{\partial \varphi_l^{(e)}}{\partial x_i} \left(u_{ni}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_k} + u_{nk}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_i} \right) = \\ & = 2 \frac{\partial \varphi_l^{(e)}}{\partial x_i} u_{nk}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_i} + 2 \frac{\partial \varphi_l^{(e)}}{\partial x_j} u_{nj}^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_k} = \\ & = 2 u_{nj}^{(e)} \left(\frac{\partial \varphi_l^{(e)}}{\partial x_j} \frac{\partial \varphi_n^{(e)}}{\partial x_k} + \delta_{jk} \frac{\partial \varphi_l^{(e)}}{\partial x_i} \frac{\partial \varphi_n^{(e)}}{\partial x_i} \right). \end{aligned}$$

В итоге (11.77) примет вид

$$\frac{\partial J_e[\mathbf{u}]}{\partial u_{lk}^{(e)}} \approx k_{lk,nj}^{(e)} u_{nj}^{(e)} - \tilde{b}_{lk}, \quad j, k = \overline{1, 3}, \quad l, n = \overline{1, N_e}, \quad (11.78)$$

где

$$\begin{aligned} k_{lk,nj}^{(e)} &= \tilde{\lambda}_q^{(e)} \int_{V_e} \frac{\partial \varphi_l^{(e)}}{\partial x_k} \varphi_q^{(e)} \frac{\partial \varphi_n^{(e)}}{\partial x_j} dV + \\ &+ \tilde{\mu}_q^{(e)} \int_{V_e} \varphi_q^{(e)} \left(\frac{\partial \varphi_l^{(e)}}{\partial x_j} \frac{\partial \varphi_n^{(e)}}{\partial x_k} + \delta_{jk} \frac{\partial \varphi_l^{(e)}}{\partial x_i} \frac{\partial \varphi_n^{(e)}}{\partial x_i} \right) dV \quad (11.79) \end{aligned}$$

и

$$\tilde{b}_{lk} = b_{nk}^{(e)} \int_{V_e} \varphi_l^{(e)} \varphi_n^{(e)} dV - p_{nk}^{(e)} \int_{S'_e} \varphi_l^{(e)} \varphi_n^{(e)} dS. \quad (11.80)$$

Элементы $k_{lk,nj}$ составляют симметрическую матрицу $K^{(e)}$ порядка $3N_e$, называемую **матрицей жесткости КЭ**, а элементы $\tilde{b}_{lk}$ — матрицу-столбец $\tilde{B}^{(e)}$ размера $3N_e \times 1$. Отметим, что при использовании симплексных КЭ в виде тетраэдров (для трехмерной задачи) $\frac{\partial \varphi_l^{(e)}}{\partial x_k} = \varphi_{lk}^{(e)} = \text{const}$, так что с учетом (10.31)

$$k_{lk,nj}^{(e)} = (\tilde{\lambda}^{(e)} \varphi_{lk}^{(e)} \varphi_{nj}^{(e)} + \tilde{\mu}^{(e)} (\varphi_{lj}^{(e)} \varphi_{nk}^{(e)} + \delta_{jk} \varphi_{li}^{(e)} \varphi_{ni}^{(e)})) V_e,$$

где $\tilde{\lambda}^{(e)}$ и $\tilde{\mu}^{(e)}$ — средние арифметические узловых значений констант Ламе КЭ с номером $e = \overline{1, E}$.

Подставив (11.78) в (11.76), запишем необходимое условие $J[u, \delta u] = 0$ экстремума функционала (11.65):

$$\sum_{e=1}^E (k_{lk,nj}^{(e)} u_{nj}^{(e)} - \tilde{b}_{lk}) \delta u_{lk}^{(e)} = 0, \quad j, k = \overline{1, 3}, \quad l, n = \overline{1, N_e}.$$

Чтобы использовать это условие для нахождения узловых значений перемещений, перейдем в нем при помощи матриц Ω_e , $e = \overline{1, E}$, устанавливающих соответствие между номерами узлов каждого КЭ с номерами N узлов сетки КЭ (см. 10.3), от $u_{nj}^{(e)}$ и $\delta u_{lk}^{(e)}$ к u_{Nj} и δu_{Lk} , $L, N = \overline{1, N_\Sigma}$, соответственно:

$$\sum_{e=1}^E (k_{lk,nj}^{(e)} \omega_{nN}^{(e)} u_{Nj} - \tilde{b}_{lk}) \omega_{lL}^{(e)} \delta u_{Lk} = 0, \quad j, k = \overline{1, 3}, \quad (11.81)$$

или в матричной записи $\delta U^T (KU - \tilde{B}) = 0$, где K — симметрическая **глобальная матрица жесткости** порядка $3N_\Sigma$ с элементами

$$k_{Lk,Nj} = \sum_{e=1}^E \omega_{lL}^{(e)} k_{lk,nj}^{(e)} \omega_{nN}^{(e)}, \quad L, N = \overline{1, N_\Sigma}, \quad j, k = \overline{1, 3},$$

U и $\tilde{B}$ — матрицы-столбцы размера $3N_\Sigma \times 1$ с элементами u_N , и $\tilde{b}_{Lk} = \sum_{e=1}^E \omega_{iL}^{(e)} \tilde{b}_{lk}^{(e)}$ соответственно.

Если игнорировать граничные условия (11.64), заданные на участках S_1 поверхности S , то вариации δu_{Lk} в (11.81) можно считать произвольными, что приводит к системе линейных алгебраических уравнений (СЛАУ)

$$k_{Lk, N_j} u_{N_j} = \tilde{b}_{Lk}, \quad L, N = \overline{1, N_\Sigma}, \quad j, k = \overline{1, 3}, \quad (11.82)$$

или в матричной записи $KU = \tilde{B}$. Однако в узлах сетки с номерами N' , соответствующих точкам $P_{N'} \in S_1$, значения $u_{N'_j} = u_k^\circ(P_{N'})$ заданы, т.е. в (11.81) $\delta u_{Lk} = 0$ при $L = N'$. Упорядочим номера N' так, чтобы $N' = \overline{N^*+1, N_\Sigma}$, где $N^* = N_\Sigma - N_1$ и N_1 — общее число узлов на участках S_1 . Тогда в остальных узлах с номерами $L = N = \overline{1, N^*}$ вариации δu_{Lk} будут произвольны.

Таким образом, условие (11.81) будет выполнено, если искомые узловые значения u_{N_j} , $N = \overline{1, N^*}$, удовлетворяют СЛАУ $K^*U^* = \tilde{B}^*$, где K^* — симметрическая матрица порядка $3N_\Sigma^*$, полученная из глобальной матрицы жесткости K вычеркиванием строк и столбцов, соответствующих номерам $N' = \overline{N^*+1, N_\Sigma}$, а U^* и $\tilde{B}^*$ — матрицы-столбцы размера $3N^* \times 1$ с элементами u_{N_j} , $N = \overline{1, N^*}$, и

$$\tilde{b}_{Lk}^* = \tilde{b}_{Lk} - k_{Lk, N'_j} u_j^\circ(P_{N'}), \quad L = \overline{1, N^*},$$

соответственно. Отметим, что матрица K^* является *слабозаполненной*, поскольку любой ее элемент отличен от нуля лишь в том случае, если номера строки и столбца соответствуют номерам узлов сетки, принадлежащих одному и тому же КЭ. Это позволяет использовать для численного решения СЛАУ с такой матрицей способы, приводящие к экономии вычислительных ресурсов ЭВМ*.

*См.: Джордж А., Лю Дж., а также: Лиссанецки С.

Динамическая задача. Перейдем к рассмотрению динамической задачи теории упругости, описываемой уравнениями Ламе (11.62), которым соответствуют *уравнения движения* (2.85)

$$\rho^{\circ} \frac{\partial^2 u_i}{\partial t^2} = \frac{\partial \sigma_{ji}}{\partial x_j} + b_i = 0, \quad i, j = \overline{1, 3}. \quad (11.83)$$

Аналогично статической задаче теории упругости на участках S_1 и $S_2 = S \setminus S_1$ поверхности S , ограничивающей область V (см. рис. 11.2), зададим граничные условия вида (11.64)

$$u_i(t, P) = u_i^{\circ}(t, P), \quad P \in S_1; \quad \sigma_{ij}(t, P) n_j(P) = p_i^{\circ}(t, P), \quad P \in S_2,$$

где $u_i^{\circ}(i, P)$ и $p_i^{\circ}(t, P)$ — заданные функции времени и координат точки P . В качестве начальных условий примем зависимости проекций $u_i(0, M) = \bar{u}_i(M)$ перемещений и $\frac{\partial u_i(0, M)}{\partial t} = \bar{v}_i(M)$ скоростей от координат точек $M \in V$ в момент времени $t = 0$, принимаемый за начало отсчета.

Сформулированной задаче не удастся поставить в соответствие функционал, для которого ее решение было бы его *точкой экстремума*. Поэтому в качестве интегральной формулировки этой задачи примем условия (6.76) ортогональности невязок, которые возникают при подстановке в (11.83) и в граничные условия искомого приближенного решения, N_{Σ} -мерному (по числу узлов сетки лагранжевых КЭ в области V) подпространству $\mathcal{H}_{N_{\Sigma}}$ гильбертова пространства $\mathcal{H}$ непрерывных на $\bar{V} = V \cup S$ функций, имеющих в V кусочно непрерывные производные и принимающие на S_1 нулевые значения. Это решение будем искать в форме

$$u_i(t, M) = \sum_{N=1}^{N_{\Sigma}} u_{N_i}(t) \varphi_N(M), \quad M \in V, \quad i = \overline{1, 3}, \quad (11.84)$$

где $u_{N_i}(t)$ — зависящие от времени проекции перемещений в узлах сетки КЭ, а линейно независимые функции формы узлов

этой сетки

$$\varphi_N(M) = \sum_{e=1}^E \Omega_{nN} \varphi_n^{(e)}(M), \quad N = \overline{1, N_\Sigma}, \quad (11.85)$$

образуют базис N_Σ -мерного подпространства.

Можно показать*, что указанные условия ортогональности приведут к системе $3N_\Sigma$ обыкновенных дифференциальных уравнений (ОДУ)

$$\tilde{m}_{LN} \delta_{kj} \frac{d^2 u_{Nj}(t)}{dt^2} + k_{Lk, Nj} u_{Nj}(t) = \tilde{b}_{Lk}(t), \quad (11.86)$$

$L, N = \overline{1, N_\Sigma}$, $k, j = \overline{1, 3}$, с начальными условиями $u_{Nj}(0) = \bar{u}_{j}(M_N)$, где $M_N \in \bar{V}$ — точки, соответствующие узлам с номерами N . Отметим, что элементы $\tilde{b}_{Lk}$ в общем случае могут зависеть от времени, поскольку в них входят заданные в (11.83) и граничных условиях функции b_k и p_k° , $k = \overline{1, 3}$, вообще говоря, зависящие от t .

Входящие в (11.86) значения

$$\tilde{m}_{LN} = \int_V \rho^\circ(M) \varphi_L(M) \varphi_N(M) dV$$

являются элементами симметрической **глобальной матрицы масс** $\tilde{M}$ порядка N_Σ . Ее (как и глобальную матрицу жесткости) удобнее формировать из вкладов отдельных КЭ. Учитывая (11.85), эти значения можно представить в виде

$$\begin{aligned} \tilde{m}_{LN} &= \sum_{e=1}^E \int_{V_e} \rho^\circ(M) \left(\Omega_{lL}^{(e)} \varphi_l^{(e)}(M) \right) \left(\Omega_{nN}^{(e)} \varphi_n^{(e)}(M) \right) dV = \\ &= \sum_{e=1}^E \tilde{m}_{ln}^{(e)} \Omega_{lL}^{(e)} \Omega_{nN}^{(e)}, \end{aligned}$$

*См. Зарубин В.С., Селиванов В.В.

где

$$\tilde{m}_{ln}^{(e)} = \int_{V_e} \rho^o(M) \varphi_l^{(e)}(M) \varphi_n^{(e)}(M) dV.$$

Элементы $\tilde{m}_{ln}^{(e)}$, $l, n = \overline{1, N_e}$, образуют симметрическую **матрицу масс КЭ** с номером e , имеющую порядок N_e . Эти элементы распределяют массу КЭ

$$m^{(e)} = \int_{V_e} \rho^o(M) dV$$

по его узлам в виде сосредоточенных масс так, что изменение кинетической энергии дискретной системы совпадает с ее изменением в КЭ при непрерывных распределениях плотности $\rho^o(M)$ и проекций $\frac{\partial u_i(t, M)}{\partial t}$, $M \in V_e$, скорости*. Матрица масс аналогична **матрице теплоемкости**. Поэтому для вычисления ее элементов в случае симплексных КЭ можно использовать соотношения, полученные для матрицы теплоемкости (см. 11.2).

Чтобы представить (11.86) в матричной записи, перейдем к глобальной матрице масс $\overline{M}$ порядка $3N_\Sigma$ с элементами $\overline{m}_{Lk, N_j} = \overline{m}_{3(L-1)+k, 3(N-1)+j} = \tilde{m}_{LN} \delta_{kj}$ и запишем

$$\overline{M} \frac{d^2 U}{dt^2} + K U = \tilde{B}. \quad (11.87)$$

Снова примем, что в N_1 узлах сетки с номерами N' , соответствующих точкам $P_{N'} \in S_1$, в силу граничных условий заданы зависимости $u_{N'_i}(t) = u_i^o(t, P'_{N'})$. Поэтому вместо (11.87) будем иметь

$$\overline{M}^* \frac{d^2 U^*}{dt^2} + K^* U = \tilde{B}^*, \quad (11.88)$$

где $\overline{M}^*$ — симметрическая матрица порядка $3N^* = 3(N_\Sigma - N_1)$, полученная из матрицы $\overline{M}$ вычеркиванием строк и столбцов, соответствующих номерам $N' = \overline{N^*+1, N_\Sigma}$, а U^* и $\tilde{B}^*$ — матри-

*См., например: Зарубин В.С., Селиванов В.В.

цы-столбцы размера $3N^* \times 1$ с элементами $u_{N_j}(t)$, $N = \overline{1, N^*}$, и

$$\begin{aligned} \tilde{b}_{Lk}^*(t) = \tilde{b}_{Lk}(t) - k_{Lk, N_j'} u_j^0(t, P_{N'}) - \\ - \overline{m}_{Lk, N_j'} \frac{d^2 u_j^0(t, P_{N'})}{dt^2}, \quad L = \overline{1, N_\Sigma^*}, \end{aligned}$$

соответственно. От (11.88) обращением матрицы $\overline{M}^*$ можно перейти к *нормальной системе ОДУ* и найти решение этой системы с учетом заданных начальных условий численно, например, *методом Рунге — Кутты*.

11.5. Электромагнитное поле в цилиндрическом волноводе

При проектировании устройств для передачи энергии электромагнитными волнами возникает необходимость анализировать электромагнитное поле в волноводе — цилиндрическом канале (не обязательно с круговым поперечным сечением), образованном электропроводными (обычно металлическими) стенками*. Примем, что стенки волновода являются идеальными проводниками, что соответствует поверхности, отражающей без потерь падающую на нее волну (аналогично прохождению света через трубу с зеркальной внутренней поверхностью).

Пусть среда, заполняющая волновод, имеет значения диэлектрической и магнитной проницаемостей $\epsilon = 1$, $\mu = 1$ и в ней отсутствуют электрические токи и заряды. Тогда при распространении в волноводе электромагнитной волны с частотой ω форма волны для любой проекции векторов напряженности электрического $\mathbf{E}$ и магнитного $\mathbf{H}$ полей, описываемая функцией u пространственных координат, удовлетворяет *уравнению Гельмгольца* (3.60)

$$\nabla^2 u + k^2 u = 0, \quad k = \frac{\omega}{c}, \quad (11.89)$$

где $c \approx 2,9979 \cdot 10^8$ м/с — скорость света в вакууме.

*См.: Сильвестер Л., Феррари Р.

Ось Ox_3 декартовой прямоугольной системы координат направим вдоль образующей цилиндрического волновода с произвольным замкнутым контуром Γ его поперечного сечения, ограничивающим односвязную область F (рис. 11.3). Вдоль этой оси в достаточно длинном волноводе могут распространяться бегущие волны [XII]. Опишем изменение формы одной из таких волн в направлении оси Ox_3 функцией $\cos \beta x_3$ ($\beta > 0$) и примем, что в этой волне $H_3 \equiv 0$, т.е. отсутствует проекция вектора $\mathbf{H}$ на ось Ox_3 . В этом случае проекции H_1 и H_2 вектора $\mathbf{H}$, а также проекции E_1 и E_2 вектора $\mathbf{E}$ можно выразить через проекцию E_3 [XII]. Так как в идеальном проводнике вектор напряженности электрического поля является нулевым, а на внутренней поверхности стенок волновода как на *поверхности разрыва* тангенциальная проекция вектора $\mathbf{E}$ в соответствии с (3.89) непрерывна, то $E_3 = 0$ в точках $P \in \Gamma$.

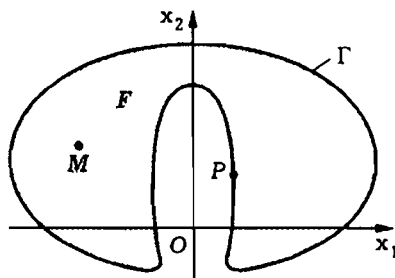


Рис. 11.3

Подставляя $u = E_3 = w(x_1, x_2) \cos \beta x_3$ в (11.89), получаем уравнение Гельмгольца относительно функции $w(M) = w(x_1, x_2)$ двух переменных:

$$\nabla_2^2 w + \lambda w = 0, \quad (11.90)$$

где ∇_2^2 — двумерный оператор Лапласа в координатной плоскости $x_1 O x_2$, а $\lambda = k^2 - \beta^2$. Эта функция должна удовлетворять однородному граничному условию

$$w(P) = 0, \quad P \in \Gamma. \quad (11.91)$$

Отсюда следует, что в волноводе возможно распространение волн рассматриваемого типа лишь с такими значениями β , для которых при заданных значениях k существует ненулевое решение краевой задачи (11.90), (11.91), т.е. задачи на собственные значения для оператора Лапласа ∇_2^2 . Этот оператор при условии (11.91) является *положительно определенным* (см. Д.5.3) и все его собственные значения λ_n , $n \in \mathbb{N}$, положительны (см. 5.4).

Таким образом, каждому собственному значению $\lambda_n > 0$ соответствуют значение $\beta_n = \sqrt{k^2 - \lambda_n}$ и собственная функция $w(M)$, $M \in F$, характеризующая форму волны в плоскости поперечного сечения волновода. Наименьшему собственному значению λ_1 отвечает также наименьшая (называемая критической) частота $\omega_* = c\sqrt{\lambda_1}$, ограничивающая снизу частоты электромагнитных волн, которые могут распространяться в рассматриваемом волноводе. Следовательно, наименьшее собственное значение λ_1 задачи (11.90), (11.91) позволяет вычислить одну из важнейших характеристик волновода. Это значение для волновода с произвольным поперечным сечением можно найти при помощи *метода конечных элементов* (МКЭ).

Для применения МКЭ от краевой задачи (11.90), (11.91) перейдем к ее вариационной формулировке, содержащей *функционал*

$$J[w] = \frac{1}{2} \int_F ((\nabla w)^2 - \lambda w^2) dF, \quad (11.92)$$

минимизируемый на множестве непрерывных функций $w(M)$, принимающих в соответствии с (11.91) нулевое значение на контуре Γ и имеющих в области F кусочно непрерывные производные. Используя аналогичную описанной выше (см. 11.2) процедуру разбиения области F на *конечных элементов* (КЭ) с общим числом N_Σ узлов и необходимые условия минимума функционала (11.92), получаем однородную систему линейных алгебраических уравнений (СЛАУ)

$$AW = \lambda BW, \quad (11.93)$$

где W — матрица-столбец размера $N_\Sigma \times 1$, элементами которой являются *узловые значения* функции w ; A и B — симметрические матрицы порядка N_Σ с элементами

$$A_{LN} = \sum_{e=1}^E \sum_{l=1}^{N_e} \sum_{n=1}^{N_e} \Omega_{lL}^{(e)} A_{ln}^{(e)} \Omega_{nN}^{(e)}, \quad B_{LN} = \sum_{e=1}^E \sum_{l=1}^{N_e} \sum_{n=1}^{N_e} \Omega_{lL}^{(e)} B_{ln}^{(e)} \Omega_{nN}^{(e)}.$$

В этих равенствах

$$A_{ln}^{(e)} = \int_{F_e} \sum_{i=1}^2 \frac{\partial \varphi_l^{(e)}}{\partial x_i} \frac{\partial \varphi_n^{(e)}}{\partial x_i} dF, \quad B_{ln}^{(e)} = \int_{F_e} \varphi_l^{(e)} \varphi_n^{(e)} dF,$$

$\Omega_{nN}^{(e)}$ — элементы матрицы Ω_e , устанавливающей соответствие между нумерацией узлов в каждом КЭ с номером $e = \overline{1, E}$ и глобальной нумерацией, а $\varphi_n^{(e)}$, $n = \overline{1, N_e}$, — *функции формы* этого КЭ, имеющего N_e узлов и занимающего область F_e .

Пусть N_1 узлов, принадлежащих контуру Γ , имеют номера $N = \overline{N_*+1, N_\Sigma}$, где $N_* = N_\Sigma - N_1$. Поскольку в этих узлах заданы нулевые значения функции w , вместо (11.93) имеем однородную СЛАУ

$$\sum_{N=1}^{N_*} (A_{LN} - \lambda B_{LN}) W_N = 0, \quad L = \overline{1, N_*}. \quad (11.94)$$

Элементы A_{LN} и B_{LN} в (11.94) составляют симметрические матрицы A_* и B_* порядка N_* соответственно, полученные вычеркиванием из матриц A и B последних N_1 строк и столбцов. С помощью этих матриц (11.94) можно представить в матричной записи

$$A_* W_* = \lambda B_* W_*, \quad (11.95)$$

где W_* — матрица-столбец размера $N_* \times 1$, элементами которой являются значения функции w во внутренних узлах *сетки* КЭ.

Матрица B является положительно определенной. Это свойство сохраняет и матрица B_* . Поэтому существует матрица B_*^{-1} , обратная к матрице B_* , так что искомое значение λ_1 будет наименьшим корнем характеристического уравнения $\det(A_* B_*^{-1} - \lambda I_*) = 0$ матрицы $A_* B_*^{-1}$, где I_* — единичная матрица порядка N_* . Существуют различные способы решения такого уравнения [II], [IV].

Матрица AB_*^{-1} в общем случае не является симметрической. Поэтому более удобный способ нахождения собственных значений и собственных векторов СЛАУ (11.95) связан с использованием разложения Холецкого $B_* = SS^T$, где S — нижняя треугольная матрица порядка N_* . Тогда после умножения (11.95) слева на S^{-1} получим

$$S^{-1}A_*((S^T)^{-1}S^T)W_* = \lambda S^{-1}(SS^T)W_*,$$

или

$$\hat{A}y = \lambda y, \quad (11.96)$$

где $\hat{A} = S^{-1}A_*(S^{-1})^T$ — симметрическая матрица, а $y = S^T W_*$ — N_* -мерный вектор, координаты которого являются линейными комбинациями значений функции w во внутренних узлах сетки КЭ и характеризуют форму волны.

Задачу нахождения собственных значений и собственных векторов СЛАУ вида (11.96) обычно называют алгебраической проблемой собственных значений. Алгоритмы численного решения этой проблемы хорошо разработаны и входят в математическое обеспечение современных ЭВМ*.

Отметим, что аналогичная рассмотренной задача возникает при исследовании акустических колебаний в цилиндрическом канале, по которому распространяется бегущая звуковая волна. В замкнутом объеме, называемом резонатором, могут возникнуть стоячие акустические или электромагнитные волны. Их частота и форма также удовлетворяют уравнению Гельмгольца

*См., например: Воеводин В.В., а также: Уилкинсон, Райнш.

с однородными граничными условиями. Краевая задача (11.90), (11.91) описывает установившиеся колебания гибкой мембраны, закрепленной по контуру Γ . К этой же задаче приводит анализ возможности теплового взрыва в твердом теле с внутренними источниками энерговыделения, объемная мощность которых возрастает с увеличением температуры*.

Вопросы и задачи

11.1. Вывести (11.11) и (11.12).

11.2. Получить (11.20).

11.3. Вычислить значения $C_{ln}^{(e)}$ в (11.32) в случае одномерных и двумерных симплексных конечных элементов.

11.4. Получить выражения для элементов матрицы $\tilde{A}(e)$, входящей в (11.54).

11.5. Вывести формулы (11.79) и (11.80).

11.6. Проверить справедливость выражений для элементов глобальной матрицы масс, входящих в (11.86).

11.7. Вывести выражения для элементов матриц A и B в (11.93) в случае одномерных и двумерных симплексных конечных элементов.

*См.: Зарубин В.С., 1983.

12. ВВЕДЕНИЕ В МЕТОД ГРАНИЧНЫХ ЭЛЕМЕНТОВ

Одним из эффективных путей приближенного решения задач математической физики является использование их интегральной формулировки в сочетании с применением *метода граничных элементов* (МГЭ). Этот метод можно рассматривать как модификацию *метода конечных элементов* (МКЭ) для аппроксимации искомых функций, но не в области V решения задачи, а на ее границе S . Это позволяет понизить размерность решаемой задачи: вместо трехмерной задачи в пространственной области решать двумерную задачу на ограничивающей поверхности, а вместо двумерной — одномерную на ограничивающем плоскую область контуре.

В теоретической основе МГЭ лежит переход от задачи, описываемой дифференциальными уравнениями с частными производными, к ее *интегральной формулировке*. Но в отличие от МКЭ эта формулировка включает интегралы от искомых функций и их производных, вычисляемые лишь по границе области. **Интегральные уравнения**, содержащие такие интегралы, принято называть **граничными**.

12.1. Граничные интегральные уравнения

Рассмотрим последовательность перехода от математической формулировки задачи, описываемой дифференциальными уравнениями с частными производными, к *граничному интегральному уравнению* (ГИУ) сначала на примере двумерной краевой задачи для уравнения Пуассона.

Обозначим через $r = r(M, M_0)$ расстояние между точками M_0 и M двумерной области F , ограниченной кусочно гладкой

кривой Γ (рис. 12.1). Будем считать, что Γ может содержать лишь конечное множество угловых точек (точка P_0 на рис. 12.1), причем в этих точках существуют односторонние касательные и конечные односторонние пределы кривизны гладких участков кривой Γ .

Напомним, что функция

$$v(M, M_0) = -\ln \frac{r(M, M_0)}{R_0}, \quad (12.1)$$

где $R_0 > 0$ — некоторое произвольно выбранное расстояние (например, характерный размер рассматриваемой области F), является *фундаментальным решением уравнения Лапласа* $\nabla^2 u = 0$ в $\mathbb{R}^2$ [XII]. Эту функцию иногда называют *сингулярным решением уравнения Лапласа на плоскости*, так как $\nabla^2 v(M, M_0) = 0$ во всех точках $M \in \mathbb{R}^2$, за исключением точки $M_0 \in \mathbb{R}^2$, в которой $v \rightarrow \infty$. Действительно, поместив в точку M_0 начало прямоугольной декартовой системы координат Ox_1x_2 , запишем $r(M, M_0) = \sqrt{x_1^2(M) + x_2^2(M)}$ и вычислим

$$\frac{\partial v}{\partial x_1} = -\frac{x_1}{r^2}, \quad \frac{\partial^2 v}{\partial x_1^2} = -\frac{1}{r^2} + \frac{2x_1^2}{r^4},$$

$$\nabla^2 v = \frac{\partial^2 v}{\partial x_1^2} + \frac{\partial^2 v}{\partial x_2^2} = -\frac{2}{r^2} + 2\frac{x_1^2 + x_2^2}{r^4} = -\frac{2}{r^2} + 2\frac{r^2}{r^4} = 0$$

при $r(M, M_0) \neq 0$.

По физическому смыслу $v(M, M_0)$ описывает в координатной плоскости x_1Ox_2 осесимметричное скалярное поле, создаваемое источниками (например, температурное поле, создаваемое тепловыми источниками), равномерно распределенными на прямой, перпендикулярной этой плоскости и проходящей через точку M_0 . Если на такой прямой расположить с постоянной линейной плотностью положительные электрические заряды, то

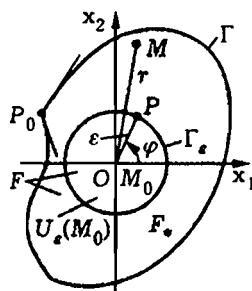


Рис. 12.1

функция $v(M, M_0)$ будет пропорциональна потенциалу возникающего при этом плоского осесимметричного электростатического поля.

Пусть на замыкании $\bar{F} = F \cup \Gamma$ области F определена дважды непрерывно дифференцируемая функция $u(M)$. Введем круговую окрестность $U_\epsilon(M_0)$ с центром в фиксированной точке $M_0 \in F$, ограниченную окружностью Γ_ϵ радиуса ϵ , не имеющей общих точек с контуром Γ , и рассмотрим область $F_* = F \setminus (U_\epsilon(M_0) \cup \Gamma_\epsilon)$ (см. рис. 12.1). На замыкании $\bar{F}_*$ этой области обе функции $u(M)$ и $v(M, M_0)$ дважды непрерывно дифференцируемы, причем $\nabla^2 v \equiv 0$, и к ним применима *вторая формула Грина*:

$$\begin{aligned} \int_{F_*} v \nabla^2 u \, dF &= \int_{F_*} (v \nabla^2 u - u \nabla^2 v) \, dF = \\ &= \int_{\Gamma} (v \nabla u - u \nabla v) \mathbf{n} \, d\Gamma + \int_{\Gamma_\epsilon} (v \nabla u - u \nabla v) \mathbf{n} \, d\Gamma, \quad (12.2) \end{aligned}$$

где $\mathbf{n}$ — единичный вектор внешней нормали к границе области F_* .

Будем стягивать при $\epsilon \rightarrow 0$ окрестность $U_\epsilon(M_0)$ в точку M_0 . В точках окружности Γ_ϵ внешняя нормаль направлена к точке M_0 , так что

$$(\nabla u) \mathbf{n} = - \frac{\partial u}{\partial r} \Big|_{r=\epsilon}, \quad (\nabla v) \mathbf{n} = - \frac{\partial v}{\partial r} \Big|_{r=\epsilon} = \frac{1}{r} \Big|_{r=\epsilon} = \frac{1}{\epsilon}.$$

Кроме того, на этой окружности $v = -\ln \frac{\epsilon}{R_0}$ и $d\Gamma = \epsilon d\varphi$, где $\varphi \in [0, 2\pi)$ — угол, отсчитываемый от оси Ox_1 (см. рис. 12.1). Второй интеграл в правой части (12.2) можно представить в виде

$$I_2 = \epsilon \ln \frac{\epsilon}{R_0} \int_0^{2\pi} \frac{\partial u(P)}{\partial r} d\varphi - \int_0^{2\pi} u(P) d\varphi, \quad P \in \Gamma_\epsilon. \quad (12.3)$$

Первые производные функции u непрерывны на $\bar{F}$ и потому ограничены. Следовательно, ограничен и первый интеграл в правой части (12.3). Так как $\varepsilon \ln \frac{\varepsilon}{R_0} \rightarrow 0$ при $\varepsilon \rightarrow 0$, то первое слагаемое в правой части (12.3) стремится к нулю при $\varepsilon \rightarrow 0$. Второй интеграл в (12.3) равен $2\pi \bar{u}_\varepsilon$, где $\bar{u}_\varepsilon$ — среднее значение функции u на окружности Γ_ε . Так как $\bar{u}_\varepsilon \rightarrow u(M_0)$ при $\varepsilon \rightarrow 0$, то в итоге $I_2 \rightarrow -2\pi u(M_0)$ при $\varepsilon \rightarrow 0$.

Первый интеграл в правой части (12.2) не зависит от ε и поэтому при $\varepsilon \rightarrow 0$ сохраняет свое значение. Левая часть (12.2) при $\varepsilon \rightarrow 0$ является несобственным интегралом

$$I_0 = \int_F v \nabla^2 u dF = - \int_F \left(\ln \frac{r(M, M_0)}{R_0} \right) \nabla^2 u(M) dF(M) \quad (12.4)$$

по области F от неограниченной в окрестности точки M_0 функции. Этот интеграл сходится, так как существует предел правой части (12.2). Итак, учитывая предельное значение интеграла I_2 , получаем интегральное представление дважды непрерывно дифференцируемой функции $u(M)$ в произвольной точке $M_0 \in F$ в виде

$$\begin{aligned} 2\pi u(M_0) = & \int_{\Gamma} u(P) \left(\nabla \ln \frac{r(P, M_0)}{R_0} \right) \mathbf{n}(P) d\Gamma(P) - \\ & - \int_{\Gamma} (\nabla u(P)) \mathbf{n}(P) \ln \frac{r(P, M_0)}{R_0} d\Gamma(P) + \\ & + \int_F (\nabla^2 u(M)) \ln \frac{r(M, M_0)}{R_0} dF(M), \quad M, M_0 \in F, \quad P \in \Gamma. \end{aligned} \quad (12.5)$$

Пусть теперь точка M_0 находится на кусочно гладкой границе Γ области F . В частности, M_0 может быть угловой точкой этой границы (рис. 12.2). Дугой Γ_ε окружности радиуса ε с центром в точке M_0 выделим подобласть $F_\varepsilon \subset F$. Тогда на замыкании $\bar{F}_* = F_* \cup \Gamma_* \cup \Gamma_\varepsilon$ области $F_* = F \setminus (F_\varepsilon \cup \Gamma_\varepsilon)$ с участком

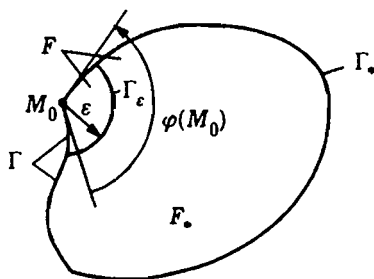


Рис. 12.2

$\Gamma_* \subset \Gamma$ внешней границы имеем $\nabla^2 v \equiv 0$ и к функциям u, v снова можно применить вторую формулу Грина:

$$\begin{aligned} \int_{F_*} v \nabla^2 u \, dF &= \int_{F_*} (v \nabla^2 u - u \nabla^2 v) \, dF = \\ &= \int_{\Gamma_*} (v \nabla u - u \nabla v) \mathbf{n} \, d\Gamma + \int_{\Gamma_\epsilon} (v \nabla u - u \nabla v) \mathbf{n} \, d\Gamma. \quad (12.6) \end{aligned}$$

Аналогично предыдущему можно показать, что средняя часть (12.6) при $\epsilon \rightarrow 0$ имеет пределом несобственный интеграл (12.4), который сходится на пересечении области F и окрестности $U_\delta(M_0)$ точки $M_0 \in \Gamma$. Второй интеграл в правой части (12.6) по аналогии с (12.3) представим в виде

$$I_2^* = \epsilon \ln \frac{\epsilon}{R_0} \int_0^{\Delta\varphi(\epsilon)} \frac{\partial u(P)}{\partial r} \, d\varphi - \int_0^{\Delta\varphi(\epsilon)} u(P) \, d\varphi, \quad P \in \Gamma_\epsilon, \quad (12.7)$$

где $\Delta\varphi(\epsilon) = l_\epsilon/\epsilon$, l_ϵ — длина дуги Γ_ϵ . В силу ограниченности первых производных функции u на $\overline{F}$ первое слагаемое в правой части (12.7) при $\epsilon \rightarrow 0$ исчезает. Второе слагаемое можно представить в виде $-\Delta\varphi(\epsilon) \bar{u}_\epsilon$, где $\bar{u}_\epsilon$ — среднее значение функции u на дуге Γ_ϵ . Так как $\bar{u}_\epsilon \rightarrow u(M_0)$ при $\epsilon \rightarrow 0$, а $\Delta\varphi(\epsilon)$ при этом стремится к значению $\varphi(M_0)$ угла, образованного касательными к контуру Γ в угловой точке $M_0 \in \Gamma$ и обращенного

в сторону области F (см. рис. 12.2), то $I_2^* \rightarrow -\varphi(M_0)u(M_0)$ при $\varepsilon \rightarrow 0$. Ясно, что если точка $M_0 \in \Gamma$ лежит на гладком участке контура, то $\varphi(M_0) = \pi$.

Но теперь при $\varepsilon \rightarrow 0$ несобственным станет и первый интеграл в правой части (12.6). Представим его в виде

$$I_1 = \int_{\Gamma} u(P) \left(\nabla \ln \frac{r(P, M_0)}{R_0} \right) \mathbf{n}(P) d\Gamma(P) - \\ - \int_{\Gamma} (\nabla u(P)) \mathbf{n}(P) \ln \frac{r(P, M_0)}{R_0} d\Gamma(P), \quad M_0 \in \Gamma, \quad (12.8)$$

и убедимся, что он сходится на участке $\Gamma_{\delta} = \Gamma \cap U_{\delta}(M_0)$ границы Γ , попавшем в некоторую δ -окрестность $U_{\delta}(M_0)$ точки $M_0 \in \Gamma$. Если Γ_{δ} состоит из прямолинейных участков, то на них $(\nabla \ln r(P, M_0)) \mathbf{n}(P) \equiv 0$ и первый интеграл в правой части (12.8) является собственным. Если же Γ_{δ} состоит из криволинейных участков Γ'_{δ} и Γ''_{δ} (рис. 12.3), то получим неопределенность вида $0/0$

$$(\nabla \ln r(P, M_0)) \mathbf{n}(P) = \frac{\partial}{\partial n} \ln r = \frac{1}{r} \frac{\partial r}{\partial n} = \frac{\cos \gamma}{r},$$

где γ — угол между направлением $\mathbf{n}$ вектора внешней нормали $\mathbf{n}(P)$ и вектора $\mathbf{r}$, проведенного из точки M_0 в точку $P \in \Gamma_{\delta}$, а

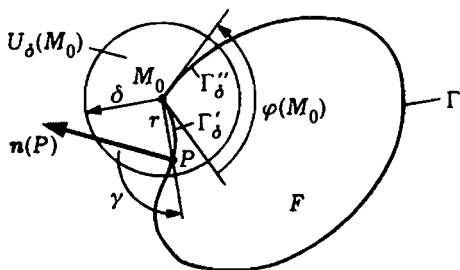


Рис. 12.3

$r = |\mathbf{r}|$. Раскроем эту неопределенность по правилу Бернулли — Лопиталя [II]:

$$\lim_{r \rightarrow 0} \frac{\cos \gamma}{r} = \lim_{r \rightarrow 0} \frac{-\sin \gamma}{1} \frac{\partial \gamma}{\partial r} = - \lim_{r \rightarrow 0} \frac{\partial \gamma}{\partial r},$$

поскольку $\gamma \rightarrow \pi/2$ при $r \rightarrow 0$. Величина $\frac{\partial \gamma}{\partial r}$ характеризует кривизну участка Γ_δ в окрестности точки $M_0 \in \Gamma_\delta$ и для кусочно гладкой границы Γ имеет в любой точке M_0 конечные односторонние пределы. Таким образом, и в этом случае первый интеграл в правой части (12.8) является собственным. Ясно, что этот вывод справедлив и в случае, если Γ_δ состоит из прямолинейного и криволинейного участков, имеющих общую точку $M_0 \in \Gamma_\delta$.

Второй интеграл в правой части (12.8) будет сходящимся, если для любого $\alpha > 0$ существует такое $\delta(\alpha) > 0$, что

$$\left| \int_{\Gamma_\delta} (\nabla u(P)) \mathbf{n}(P) \ln \frac{r(P, M_0)}{R_0} d\Gamma(P) \right| < \alpha, \quad M_0 \in \Gamma_\delta. \quad (12.9)$$

Функция $u(M)$ непрерывно дифференцируема на участке Γ_δ и поэтому ее производные ограничены, т.е. существует такое число $C_1 > 0$, что $|(\nabla u(P)) \mathbf{n}(P)| \leq C_1$ при $P \in \Gamma_\delta$. Но тогда справедлива оценка

$$\begin{aligned} & \left| \int_{\Gamma_\delta} (\nabla u(P)) \mathbf{n}(P) \ln \frac{r(P, M_0)}{R_0} d\Gamma(P) \right| \leq \\ & \leq \int_{\Gamma_\delta} \left| (\nabla u(P)) \mathbf{n}(P) \right| \left| \ln \frac{r(P, M_0)}{R_0} \right| d\Gamma(P) \leq \\ & \leq C_1 \int_{\Gamma_\delta} \left| \ln \frac{r(P, M_0)}{R_0} \right| d\Gamma(P), \quad M_0 \in \Gamma_\delta. \end{aligned} \quad (12.10)$$

Если Γ_δ состоит из прямолинейных участков, то на каждом из них в предположении, что $\delta < 1$, имеем

$$\begin{aligned} \int_{\Gamma_\delta} \left| \ln \frac{r(P, M_0)}{R_0} \right| d\Gamma(P) &= \int_0^{R_0\delta} \left| \ln \frac{r}{R_0} \right| dr = \\ &= R_0 (\delta |\ln \delta| - \delta - \lim_{\xi \rightarrow 0} \xi |\ln \xi|) = R_0 (\delta |\ln \delta| - \delta). \end{aligned}$$

Тогда при выборе числа $\delta > 0$, такого, что $\delta |\ln \delta| - \delta < \frac{\alpha}{C_1 R_0}$, условие (12.9) будет выполнено.

Если же Γ_δ состоит из криволинейных участков Γ'_δ и Γ''_δ (рис. 12.4), то на любом из них (например, на Γ'_δ) интеграл в правой части (12.10) можно представить в виде

$$\begin{aligned} I_3 &= \int_{\Gamma'_\delta} |\ln r(P, M_0)| d\Gamma(P) = \\ &= \int_0^\delta \left| \ln \frac{\sqrt{\xi^2 + \eta^2}}{R_0} \right| \sqrt{1 + \left(\frac{d\eta}{d\xi} \right)^2} d\xi, \quad (12.11) \end{aligned}$$

где $\xi \geq 0$ — координата, отсчитываемая от точки M_0 вдоль касательной к кривой Γ'_δ , а $\eta = \eta(\xi)$ — уравнение этой кривой, причем $\eta(0) = \frac{d\eta}{d\xi} \Big|_{\xi=0} = 0$. Убедимся, что этот интеграл схо-

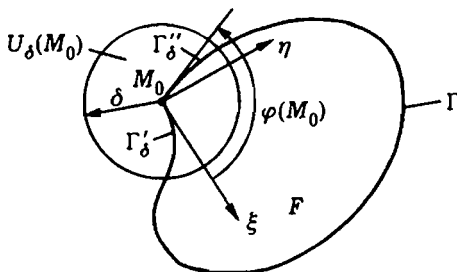


Рис. 12.4

дящийся и его значение путем уменьшения числа $\delta > 0$ можно сделать меньше любого наперед заданного числа α/C_1 . Для подынтегральной функции $f(x)$ в (12.11) выберем в качестве функции сравнения $g(\xi) = \frac{1}{\xi^\lambda}$ и вычислим при $\lambda \in (0, 1)$

$$\lim_{\xi \rightarrow 0} \frac{f(x)}{g(\xi)} = \lim_{\xi \rightarrow 0} \xi^\lambda \left| \ln \frac{\xi}{R_0} \right| \ln \sqrt{1 + \frac{\eta^2}{x^2}} \sqrt{1 + \left(\frac{d\eta}{d\xi} \right)^2} = 0.$$

В силу признака сравнения несобственный интеграл I_3 от неограниченной функции $f(x)$ сходится, поскольку сходится интеграл от функции $g(x)$. Так как в некотором промежутке $[0, \delta)$ справедливо неравенство $f(\xi) \leq K g(\xi)$, где $K = \text{const} > 0$, то

$$I_3 \leq K \int_0^\delta g(\xi) d\xi = K \int_0^\delta \frac{d\xi}{\xi^\lambda} = \frac{K}{1-\lambda} \delta^{1-\lambda}.$$

Тогда при выборе δ , удовлетворяющем условию

$$\delta < \left((1-\lambda) \frac{\alpha}{KC_1} \right)^{\frac{1}{1-\lambda}},$$

условие (12.9) будет выполнено.

Итак, интеграл I_1 в (12.8) является сходящимся, а из (12.6) следует, что для случая $M_0 \in \Gamma$ будет справедливо интегральное представление (12.5), если в нем в соответствии с предельным значением интеграла (12.7) заменить 2π на $\varphi(M_0)$. Тогда для общего случая $M_0 \in \bar{F}$, учитывая, что $\varphi(M_0) = 2\pi$ при $M_0 \in F$, получаем

$$\begin{aligned} \varphi(M_0) u(M_0) = & \int_{\Gamma} (q(P) v(P, M_0) - u(P) v^*(P, M_0)) d\Gamma(P) - \\ & - \int_F v(M, M_0) \nabla^2 u(M) dF(M), \quad M \in F, \quad M_0 \in \bar{F}, \quad P \in \Gamma, \quad (12.12) \end{aligned}$$

где $q(P) = (\nabla u(P)) \mathbf{n}(P)$ и

$$v^*(P, M_0) = (\nabla v(P, M_0)) \mathbf{n}(P) = - \left(\nabla \ln \frac{r(P, M_0)}{R_0} \right) \mathbf{n}(P).$$

Если функция $u(M)$ в области F удовлетворяет уравнению Пуассона

$$\nabla^2 u(M) + f(M) = 0, \quad M \in F, \quad (12.13)$$

где $f(M)$ — заданная функция координат точки M , то из (12.12) следует ГИУ

$$\begin{aligned} \varphi(M_0) u(M_0) = & \int_{\Gamma} (q(P) v(P, M_0) - u(P) v^*(P, M_0)) d\Gamma(P) + \\ & + \int_F f(M) v(M, M_0) dF(M), \quad M \in F, \quad M_0 \in \overline{F}, \quad P \in \Gamma. \end{aligned} \quad (12.14)$$

Отметим, что подынтегральная функция в последнем интеграле в правой части (12.14) известна и его можно вычислить при любом положении точки $M_0 \in \overline{F}$. Для решения этого ГИУ необходимо на контуре Γ задать граничные условия.

Если в точках $P \in \Gamma$ заданы значения $u(P) = f_1(P)$, то, полагая в (12.14) $M_0 \in \Gamma$ и подставляя заданную функцию $u(M_0) = f_1(M_0)$, получим ГИУ

$$\begin{aligned} \varphi(M_0) f_1(M_0) = & \int_{\Gamma} (q(P) v(P, M_0) - f_1(P) v^*(P, M_0)) d\Gamma(P) + \\ & + \int_F f(M) v(M, M_0) dF(M), \quad M \in F, \quad M_0, P \in \Gamma, \end{aligned}$$

относительно производной $q(P)$, $P \in \Gamma$, искомой функции $u(M)$, $M \in F$, по направлению внешней нормали в точках $P \in \Gamma$.

После нахождения функции $q(P)$ можно вычислить при помощи (12.14) значение $u(M_0)$ искомой функции в любой точке $M_0 \in F$.

При задании в точках $P \in \Gamma$ значений $q(P) = f_2(P)$, полагая в (12.14) $M_0 \in \Gamma$ и подставляя вместо $q(P)$ заданную функцию $f_2(P)$, $P \in \Gamma$, приходим к ГИУ

$$\begin{aligned} \varphi(M_0) u(M_0) = & \int_{\Gamma} (f_2(P) v(P, M_0) - u(P) v^*(P, M_0)) d\Gamma(P) + \\ & + \int_F f(M) v(M, M_0) dF(M), \quad M \in F, \quad M_0, P \in \Gamma, \end{aligned}$$

относительно значений искомой функции $u(M_0)$, $M_0 \in \Gamma$, на границе, которые входят и в левую часть, и в подынтегральную функцию первого интеграла в правой части этого ГИУ. После нахождения этих значений можно также вычислить при помощи (12.14) значение $u(M_0)$ искомой функции в любой точке $M_0 \in F$.

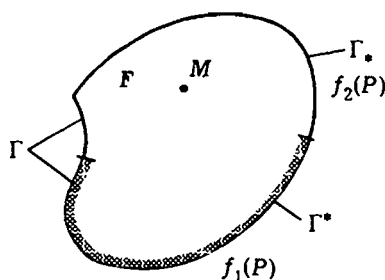


Рис. 12.5

В более общем случае на контуре Γ (рис. 12.5) могут быть заданы граничные условия

$$\begin{cases} u(P) = f_1(P), & P \in \Gamma^* \subset \Gamma; \\ q(P) + \beta(P) u(P) = f_2(P), & P \in \Gamma_* = \Gamma \setminus \Gamma^*, \end{cases} \quad (12.15)$$

где f_1, f_2, β — известные функции точки на контуре. Тогда вместо (12.14) получим

$$\begin{aligned} \varphi(M_0)u(M_0) = & \int_{\Gamma^*} (q(P)v(P, M_0) - f_1(P)v^*(P, M_0)) d\Gamma(P) - \\ & - \int_{\Gamma} ((f_2(P) - \beta(P)u(P))v(P, M_0) - u(P)v^*(P, M_0)) d\Gamma(P) + \\ & + \int_F f(M)v(M, M_0) dF(M), \quad M \in F, \quad M_0 \in \overline{F}, \quad P \in \Gamma. \quad (12.16) \end{aligned}$$

В (12.16) неизвестными являются значения $u(P)$ в точках $P \in \Gamma$ и значения $q(P)$ производной функции u по направлению вектора внешней нормали в точках $P \in \Gamma^*$. Получить аналитическое решение ГИУ (12.16) удастся лишь для области F простейшего вида (например, для внутренности и внешности окружности и для полуплоскости).

Перейдем к рассмотрению трехмерных задач.

Фундаментальное решение уравнения Лапласа $w(M, M_0) = \frac{1}{r(M, M_0)}$ в $\mathbb{R}^3$ позволяет найти интегральное представление функции $u(M)$ в области V , ограниченной поверхностью S [XII]. Для случая, когда поверхность S кусочно гладкая, а функция $u(M)$ дважды непрерывно дифференцируема на замыкании $\overline{V} = V \cup S$, это представление можно записать в виде

$$\begin{aligned} 4\pi u(M_0) = & \int_S (q(P)w(P, M_0) - u(P)w^*(P, M_0)) dS(P) - \\ & - \int_V w(M, M_0) \nabla^2 u(M) dV(M), \quad M, M_0 \in V, \quad P \in S, \quad (12.17) \end{aligned}$$

где $w^*(P, M_0) = (\nabla w)\mathbf{n}(P)$ — производная функции w по направлению единичного вектора $\mathbf{n}(P)$ внешней нормали к поверхности S в точке $P \in S$. Для функции $u(M)$, удовлетворяющей

уравнению Пуассона

$$\nabla^2 u(M) + f(M) = 0, \quad M \in V, \quad (12.18)$$

где $f(M)$ — заданная функция, путем, аналогичным рассмотренному для плоской области, можно получить, используя (12.17), ГИУ

$$\begin{aligned} \omega(M_0) u(M_0) = & \int_S (q(P) w(P, M_0) - u(P) w^*(P, M_0)) dS(P) + \\ & + \int_V f(M) w(M, M_0) dV(M), \quad M \in V, \quad P \in S, \quad M_0 \in \bar{V}, \end{aligned} \quad (12.19)$$

где $\omega(M_0)$ — телесный угол, под которым из точки $M_0 \in \bar{V}$ видна внутренность области V (в частности, $\omega(M_0) = 4\pi$ при $M_0 \in V$ и $\omega(M_0) = 2\pi$, если точка M_0 лежит на гладком участке поверхности S).

Если граничные условия имеют вид

$$\begin{cases} u(P) = f_1(P), & P \in S^* \subset S; \\ (\nabla u(P)) \mathbf{n}(P) + \beta(P) u(P) = f_2(P), & P \in S_* = S \setminus S^*, \end{cases} \quad (12.20)$$

где f_1, f_2, β — заданные функции, то из (12.19) получим

$$\begin{aligned} \omega(M_0) u(M_0) = & \int_{S^*} (q(P) w(P, M_0) - f_1(P) w^*(P, M_0)) dS(P) + \\ & + \int_{S_*} ((f_2(P) - \beta(P) u(P)) w(P, M_0) - u(P) w^*(P, M_0)) dS(P) + \\ & + \int_V f(M) w(M, M_0) dV(M), \quad M \in V, \quad M_0 \in \bar{V}, \quad P \in S. \end{aligned} \quad (12.21)$$

В этом ГИУ неизвестными являются функция $u(P)$, определенная на S_* , и функция $q(P)$, определенная на S^* . После их

нахождения можно вычислить из (12.21) значение $u(M_0)$ в любой точке $M_0 \in V$.

Как и в двумерном случае, получить аналитическое решение ГИУ (12.21) удастся лишь для простейших областей (например, для шара, внешности сферы и полупространства [XII]). Но ГИУ (12.16) и (12.21) можно решить приближенно численными методами.

12.2. Способы аппроксимации функций на границе

Первым этапом численного решения *граничных интегральных уравнений* (ГИУ) (12.16) и (12.21) является аппроксимация искомых функций и функций, заданных в граничных условиях на границе области решаемой задачи. Для этой цели удобно использовать разбиение границы на участки, которые по аналогии с *конечными элементами* также принято называть *элементами*, но *граничными* (ГЭ), что и дало название соответствующему численному методу решения ГИУ — *методу граничных элементов* (МГЭ).

Способы аппроксимации функций рассмотрим сначала для случая двумерной области F , ограниченной кусочно гладким контуром Γ (рис. 12.6). Границу Γ разобьем на N_Γ граничных элементов Γ_n , $n = \overline{1, N_\Gamma}$. Если на границе есть угловые точки,

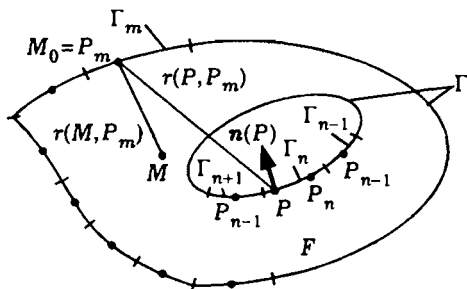


Рис. 12.6

то ГЭ расположим так, чтобы эти точки оказались на стыке элементов. В простейшем варианте МГЭ в каждом ГЭ можно взять один узел (точку $P_n \in \Gamma_n$) и поместить его в середине элемента. Это означает, что в пределах каждого ГЭ функции аппроксимированы их *узловыми значениями*. Так как каждый ГЭ является гладким участком контура Γ , то в любой точке $P \in \Gamma_n$, $n = \overline{1, N_\Gamma}$, определен единичный вектор $n(P)$ внешней нормали к Γ .

Пусть для определенности в области F требуется найти решение уравнения Пуассона (12.13) с заданными граничными условиями, т.е. решить ГИУ (12.14). Совместим точку $M_0 \in \Gamma$ с m -м узлом, т.е. положим $M_0 = P_m \in \Gamma_m$. Для этого узла, лежащего на гладком участке границы, в (12.16) имеем $\varphi(M_0) = \pi$. Поэтому вместо (12.16) с учетом аддитивности интеграла по Γ [VII] запишем

$$\pi u_m = \sum_{n=1}^{N_\Gamma} \int_{\Gamma_n} (q_n v_m(P) - u_n v_m^*(P)) d\Gamma(P) + \int_F f(M) v_m(M) dF(M), \quad (12.22)$$

где в соответствии с (12.1) $v_m(P) = -\ln \frac{r(P, P_m)}{R_0}$;

$$v_m^*(P) = (\nabla v(P, P_m)) n(P) = -\left(\nabla \ln \frac{r(P, P_m)}{R_0} \right) n(P);$$

$R_0 > 0$ — некоторый характерный размер области F ; $r(P, P_m)$ — расстояние от точки $P_m \in \Gamma_m$ до точки P (см. рис. 12.6).

Используя (12.22) для каждого узла, получаем в матричной записи

$$HU = GQ + B, \quad (12.23)$$

где U и Q — матрицы-столбцы размера $N_\Gamma \times 1$, элементами которых являются узловые значения u_n , q_n , $n = \overline{1, N_\Gamma}$; B —

матрица-столбец размера $N_\Gamma \times 1$ с элементами

$$b_m = \int_F f(M) v_m(M) dF(M), \quad m = \overline{1, N_\Gamma}; \quad (12.24)$$

H и G — матрицы порядка N_Γ с элементами

$$h_{mn} = \pi \delta_{mn} + \int_{\Gamma_n} v_m^*(P) d\Gamma(P), \quad g_{mn} = \int_{\Gamma_n} v_m(P) d\Gamma(P),$$

$m, n = \overline{1, N_\Gamma}$. Здесь $\delta_{mn} = 1$ при $m = n$ и $\delta_{mn} = 0$ при $m \neq n$.

Интегралы по ГЭ Γ_n , $n = \overline{1, N_\Gamma}$, входящие в выражения для h_{mn} и g_{mn} , в случае $n \neq m$ можно вычислить при помощи обычных *квадратурных формул*. Но в выражениях для диагональных элементов h_{mm} и g_{mm} матриц H и G интегралы являются несобственными, поскольку при сближении точки $P \in \Gamma_m$ с точкой $P_m \in \Gamma_m$ (m -м узлом) расстояние $r(P, P_m) \rightarrow 0$, так что $v_m(P) \rightarrow \infty$. Для прямолинейного ГЭ Γ_m длиной l_m , полагая, что нуль отсчета координаты ξ расположен в средней точке $M_0 = P_m \in \Gamma_m$ этого элемента (рис. 12.7), имеем

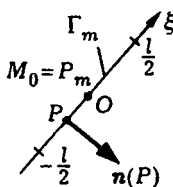


Рис. 12.7

$$\begin{aligned} g_{mm} &= \int_{\Gamma_m} v_m(P) d\Gamma(P) = \\ &= -2 \int_0^{l_m/2} \ln \frac{\xi}{R_0} d\xi = l_m \left(1 - \ln \frac{l_m}{2R_0} \right). \end{aligned} \quad (12.25)$$

Если же ГЭ криволинейный, то в окрестности точки $P_m \in \Gamma_m$ его можно приближенно представить прямолинейным участком Γ'_m длиной l'_m с точкой M_0 посередине, что дает

$$g_{mm} = g'_{mm} + \int_{\Gamma_m \setminus \Gamma'_m} v_m(P) d\Gamma(P).$$

Теперь функция $v_m(P)$ при $P \in \Gamma_m \setminus \Gamma'_m$ не имеет особенностей и можно применить обычные квадратурные формулы, а значение $g'_{m,m}$ найти из (12.25), заменив l_m на l'_m .

В случае прямолинейного ГЭ интеграл в выражении для $h_{m,m}$ равен нулю, так как тогда $v_m^*(P) = 0$, $P \in \Gamma_m \setminus \{P_m\}$, поскольку расстояние $r(P, P_m) = \xi$ между точками $P_m, P \in \Gamma_m$ изменяется в направлении, перпендикулярном единичному вектору $n(P)$ нормали в точке P (см. рис. 12.7). Криволинейный ГЭ Γ_m в малой окрестности точки $P_m \in \Gamma_m$ снова можно представить прямолинейным участком Γ'_m длиной l'_m и этот участок исключить при вычислении интеграла, положив

$$\int_{\Gamma_m} v_m^*(P) d\Gamma(P) \approx \int_{\Gamma_m \setminus \Gamma'_m} v_m^*(P) d\Gamma(P).$$

Но проще воспользоваться тем обстоятельством, что (12.23) справедливо и для частного случая $f(M) = 0$, $M \in F$. Тогда при задании на контуре Γ значений $u(P) = u_0 = \text{const}$ решением уравнения Лапласа $\nabla^2 u(M) = 0$ будет $u(M) = u_0$, причем $q(P) = 0$, $P \in \Gamma$, а вместо (12.23) получим $HU = O_{N_\Gamma}$, где O_{N_Γ} — нулевая матрица-столбец размера $N_\Gamma \times 1$. Теперь значения $u_n = u_0$, $n = \overline{1, N_\Gamma}$, будут удовлетворять матричному уравнению $HU = O_{N_\Gamma}$ лишь при условии равенства нулю суммы элементов в каждой строке матрицы H , т.е. при

$$h_{mm} = - \sum_{\substack{n=\overline{1, N_\Gamma} \\ n \neq m}} h_{mn}. \quad (12.26)$$

Отсюда можно вычислить значение h_{mm} , не прибегая к интегрированию по ГЭ Γ_m .

Интеграл в (12.24) также является несобственным и сходится в окрестности точки $P_m \in \Gamma_m$, если функция $f(M)$ ограничена в этой окрестности (см. 12.1). Его можно вычислить, например, при помощи *квадратурной формулы Гаусса*, не ис-

пользующей значения подынтегральной функции в граничных точках. Иной путь состоит в том, чтобы в малой δ -окрестности $U_\delta(P_m)$ точки $P_m \in \Gamma_m$ радиуса $R_0\delta$ принять $f(M) \approx f_m$, где f_m — предел функции $f(M)$ при стремлении точки $M \in F$ к точке P_m . Тогда для пересечения $F_\delta = F \cap U_\delta(P_m)$, считая, что $\delta < 1$, получим

$$\begin{aligned} I_\delta &= \int_{F_\delta} f(M) v_m(M) dF(M) \approx -f_m \int_0^{\varphi(P_m) R_0 \delta} \int_0^r \ln \frac{r}{R_0} d\varphi dr = \\ &= -\varphi(P_m) f_m R_0^2 \left(\frac{\delta^2}{2} \ln \delta + \frac{\delta^2}{4} - \lim_{\xi \rightarrow 0} \frac{\xi^2}{2} \ln \xi \right) = \\ &= \varphi(P_m) f_m R_0^2 \left(\frac{\delta^2}{2} \ln \frac{1}{\delta} - \frac{\delta^2}{4} \right) \end{aligned}$$

и в результате вместо (12.24) запишем

$$b_m \approx I_\delta + \int_{F \setminus F_\delta} f(M) v_m(M) dF(M), \quad m = \overline{1, N_\Gamma},$$

где последний интеграл можно вычислить по обычным квадратурным формулам.

Таким образом, для выбранного разбиения на $\Gamma \Delta$ границы Γ области F и заданной функции $f(M)$, $M \in F$, в (12.23) определены все элементы матриц H , G и B . Если на Γ заданы значения $u(P) = f_1(P)$ искомой функции $u(M)$, то в (12.23) известны все элементы u_n , $n = \overline{1, N_\Gamma}$, матрицы U . Тогда (12.23) переходит в систему линейных алгебраических уравнений (СЛАУ) $GQ = B - HU$ относительно элементов q_n , $n = \overline{1, N_\Gamma}$, матрицы Q , которые являются узловыми значениями производных функции u по направлению единичного вектора $n(P_n)$ внешней нормали к контуру Γ в узловых точках $P_n \in \Gamma_n$. После решения этой СЛАУ, используя (12.14) при $\varphi(M_0) = 2\pi$, в любой точке $M_0 \in F$

можно вычислить значение искомой функции

$$u(M_0) = \frac{1}{2\pi} \sum_{n=1}^{N_\Gamma} \left(q_n \int_{\Gamma_n} v(P, M_0) d\Gamma(P) - u_n \int_{\Gamma_n} v^*(P, M_0) d\Gamma(P) \right) + \\ + \frac{1}{2\pi} \int_F f(M) v(M, M_0) dF(M). \quad (12.27)$$

Погрешность вычислений по формуле (12.27) зависит от точности представления функций $u(P)$ и $q(P)$ их постоянными значениями в пределах каждого ГЭ и от погрешности при вычислении входящих в (12.27) интегралов. Если в пределах каждого ГЭ Γ_n длиной l_n принять $v(P, M_0) \approx v(P_n, M_0)$ и $v^*(P, M_0) \approx v^*(P_n, M_0)$, $P_n \in \Gamma_n$, $n = \overline{1, N_\Gamma}$, то (12.27) переходит в менее точную, но более простую формулу

$$u(M_0) \approx \frac{1}{2\pi} \sum_{n=1}^{N_\Gamma} (q_n v(P_n, M_0) - u_n v^*(P_n, M_0)) l_n + \\ + \frac{1}{2\pi} \int_V f(M) w(M, M_0) dV(M). \quad (12.28)$$

Предположим, что функция $f(M)$ непрерывна в окрестности любой точки M_0 области F или V . Тогда несобственные интегралы по F в (12.27) и по V в (12.28) сходятся равномерно в окрестности этой точки (см. 12.1) и их можно дифференцировать по координатам точки M_0 как по параметрам [VII]. Это позволяет вычислить не только значения искомой функции в точке M_0 , но и значения ее производных в этой точке, что важно при решении некоторых прикладных задач. Например, при решении задачи теплопроводности наряду с вычислением температуры часто необходимо вычислить плотность теплового потока, выражаемого через градиент температурного поля.

В более общем случае, когда заданы граничные условия (12.15), будем иметь N'_Γ узлов на участках $\Gamma^* \subset \Gamma$ контура Γ и $N''_\Gamma = N_\Gamma - N'_\Gamma$ узлов на участках $\Gamma_* = \Gamma \setminus \Gamma^*$. Подставив в

(12.23) $u_n = f_1(P_n)$ в узлах $P_n \in \Gamma^*$ и $q_n = f_2(P_n) - u_n$ в узлах $P_n \in \Gamma_*$, придем к СЛАУ

$$AX = B^0, \quad (12.29)$$

где X — матрица-столбец размера $N_\Gamma \times 1$, элементами которой являются неизвестные значения q_n в узлах $P_n \in \Gamma^*$ и u_n в узлах $P_n \in \Gamma_*$, упорядоченные в соответствии с нумерацией узлов. При этом для элементов матрицы A порядка N_Γ и матрицы-столбца B^0 размера $N_\Gamma \times 1$ получим выражения

$$\begin{cases} a_{mk} = h_{mk} + g_{mk}\beta(P_k), & a_{mn} = -g_{mn}, \\ b_m^0 = b_m - \sum_{n=1}^{N'_\Gamma} h_{mn}f_1(P_n) + \sum_{k=1}^{N''_\Gamma} g_{mk}f_2(P_k), \end{cases} \quad (12.30)$$

если $P_n \in \Gamma^*$, $P_k \in \Gamma_*$, и

$$\begin{cases} a_{mk} = -g_{mk}, & a_{mn} = h_{mn} + g_{mn}\beta(P_n), \\ b_m^0 = b_m - \sum_{k=1}^{N'_\Gamma} h_{mk}f_1(P_k) + \sum_{n=1}^{N''_\Gamma} g_{mn}f_2(P_n), \end{cases} \quad (12.31)$$

если $P_n \in \Gamma_*$, $P_k \in \Gamma^*$. Отметим, что для некоторых (или для всех) $\Gamma \ni \Gamma_n \subset \Gamma_*$ возможно $\beta(P_n) = 0$, т.е. $q_n = f_2(P_n)$ и $a_{mn} = h_{mn}$. После решения СЛАУ (12.29) все значения u_n и q_n , $n = \overline{1, N_\Gamma}$, будут известны, так что для вычисления значения $u(M_0)$ искомой функции u в любой точке $M_0 \in F$ можно использовать (12.27).

Распределение функций $u(P)$ и $q(P)$ на границе Γ области F можно представить более точно, если в пределах каждого $\Gamma \ni \Gamma_n$ аппроксимировать эти функции линейными зависимостями. Тогда узлы целесообразно расположить на стыке соседних $\Gamma \ni \Gamma_n$, причем так, чтобы узлы не находились на границе участков Γ^* и Γ_* с различными типами заданных граничных условий (12.15). При этом в каждом n -м узле, $n = \overline{1, N_\Gamma}$, будет известно либо значение $u_n = f_1(P_n)$, $P_n \in \Gamma^*$, либо значение $f_2(P_n) = q_n + \beta(P_n)u_n$, $P_n \in \Gamma_*$.

Пусть узлы P_n и P_{n+1} с номерами n и $n+1$ принадлежат $\Gamma\mathcal{E}$ Γ_l с номером $l = \overline{1, N_{\Gamma}}$, а $\psi_n^{(l)}(P)$ и $\psi_{n+1}^{(l)}(P)$ — линейные функции расстояния точки $P \in \Gamma_l$ от этих узлов, отсчитываемого вдоль $\Gamma\mathcal{E}$ Γ_l , причем $\psi_n^{(l)} = 1$, $\psi_{n+1}^{(l)} = 0$, когда P совпадает с P_n , и $\psi_n^{(l)} = 0$, $\psi_{n+1}^{(l)} = 1$, когда P совпадает с P_{n+1} . Тогда по аналогии с одномерными симплексными конечными элементами в пределах $\Gamma\mathcal{E}$ Γ_m функции $u(P)$ и $q(P)$ аппроксимируем выражениями

$$\begin{cases} u^{(l)}(P) = u_n \psi_n^{(l)}(P) + u_{n+1} \psi_{n+1}^{(l)}(P), \\ q^{(l)}(P) = q_n \psi_n^{(l)}(P) + q_{n+1} \psi_{n+1}^{(l)}(P), \end{cases} \quad (12.32)$$

$P \in \Gamma_l$. После их подстановки в (12.14) снова придем к (12.23), но теперь, учитывая, что узел P_n является общим для $\Gamma\mathcal{E}$ Γ_m и Γ_{m-1} , получим, что

$$h_{mn} = \int_{\Gamma_l} \psi_n^{(l)}(P) v_m^*(P) d\Gamma(P) + \int_{\Gamma_{l-1}} \psi_n^{(l-1)}(P) v_m^*(P) d\Gamma(P)$$

при $n \neq m$, значения h_{mm} вычислены по формуле (12.26) и

$$g_{mn} = \int_{\Gamma_l} \psi_n^{(l)}(P) v_m(P) d\Gamma(P) + \int_{\Gamma_{l-1}} \psi_n^{(l-1)}(P) v_m(P) d\Gamma(P).$$

Для прямолинейного $\Gamma\mathcal{E}$ Γ_l длиной l_l с нулем отсчета координаты ξ_l в узле $P_n \in \Gamma_l$, т.е. при $\xi_l(P_n) = 0$ и $\xi_l(P_{n+1}) = l_l$, находим

$$\begin{aligned} \int_{\Gamma_l} \psi_n^{(l)}(P) v_m(P) d\Gamma(P) &= - \int_0^{l_l} \Gamma_l \left(1 - \frac{\xi_l}{l_l}\right) \ln \frac{\xi_l}{R_0} d\xi_l = \\ &= l_l \left(1 - \frac{l_l}{4R_0} - \left(1 - \frac{l_l}{2R_0}\right) \ln \frac{l_l}{R_0}\right). \end{aligned}$$

Тогда для узла P_m , в котором стыкуются два прямолинейных $\Gamma\mathcal{E}$ с номерами $l-1$ и l и длинами l_{l-1} и l_l соответственно,

получим

$$g_{mm} = l_{l-1} \left(1 - \frac{l_{l-1}}{4R_0} - \left(1 - \frac{l_{l-1}}{2R_0} \right) \ln \frac{l_{l-1}}{R_0} \right) + \\ + l_l \left(1 - \frac{l_l}{4R_0} - \left(1 - \frac{l_l}{2R_0} \right) \ln \frac{l_l}{R_0} \right). \quad (12.33)$$

Если же в узле P_m стыкуются два криволинейных ГЭ с номерами $l-1$ и l , то в окрестности узла эти ГЭ можно приближенно представить прямолинейными участками Γ'_{l-1} и Γ'_l длиной l'_{l-1} и l'_l соответственно и написать

$$g_{mm} = g'_{mm} + \int_{\Gamma''_{l-1}} \psi_n^{(l-1)}(P) v_m(P) d\Gamma(P) + \int_{\Gamma''_l} \psi_n^{(l)}(P) v_m(P) d\Gamma(P),$$

где $\Gamma''_{l-1} = \Gamma_{l-1} \setminus \Gamma'_{l-1}$ и $\Gamma''_l = \Gamma_l \setminus \Gamma'_l$. Значение g'_{mm} находим, заменяя в (12.33) l_{l-1} на l'_{l-1} и l_l на l'_l , а интегралы можно вычислить по обычным квадратурным формулам, поскольку функция $v_m(P)$ не имеет особенностей при $P \in \Gamma''_{l-1} \cup \Gamma''_l$.

При задании граничных условий (12.15) снова из (12.23) следует СЛАУ (12.29), решение которой дает недостающие узловые значения u_n и q_n , $n = \overline{1, N_\Gamma}$. После этого, подставляя (12.32) в (12.14) при $\varphi(M_0) = 2\pi$, можно вычислить значение $u(M_0)$ искомой функции u в любой точке $M_0 \in F$:

$$u(M_0) = \frac{1}{2\pi} \sum_{n=1}^{N_\Gamma} q_n \int_{\Gamma_{l-1} \cup \Gamma_l} \psi_n(P) v(P, M_0) d\Gamma(P) - \\ - \frac{1}{2\pi} \sum_{n=1}^{N_\Gamma} u_n \int_{\Gamma_{l-1} \cup \Gamma_l} \psi_n(P) v^*(P, M_0) d\Gamma(P) + \\ + \frac{1}{2\pi} \int_F f(M) v(M, M_0) dF(M), \quad (12.34)$$

где $\psi_n(P) = \psi_n^{(l-1)}(P)$ при $P \in \Gamma_{l-1}$ и $\psi_n(P) = \psi_n^{(l)}(P)$ при $P \in \Gamma_l$.

Распределение функций $u(P)$ и $q(P)$ в пределах ГЭ можно аппроксимировать и более сложными, чем линейные, зависи-

мостями (например, многочленами второй и более высоких степеней), однако это приводит к увеличению числа узлов в каждом ГЭ и арифметических операций при вычислении интегралов, входящих в выражения для элементов матриц H и G . При необходимости более точного представления этих функций на границе области может оказаться, что рациональнее использовать большее число простых ГЭ с линейной аппроксимацией или даже с постоянными значениями функций в пределах каждого элемента.

Решение при помощи МГЭ трехмерной задачи в области V для уравнения Пуассона (12.18) с граничными условиями (12.20) связано с представлением ограничивающей эту область кусочно гладкой поверхности S совокупностью N_S двумерных ГЭ. Эти элементы целесообразно выбирать в виде треугольников или четырехугольников (плоских или криволинейных) с аппроксимацией функций $u(P)$ и $q(P)$ в пределах каждого ГЭ с номером $n = \overline{1, N_S}$ постоянными значениями u_n и q_n или же зависимостями от координат в форме многочленов.

Если в пределах n -го ГЭ S_n принять $u(P) \approx u_n$ и $q(P) \approx q_n$, $P \in S_n$, $n = \overline{1, N_S}$, то соответствующее задаче (12.18), (12.20) ГИУ (12.21) переходит в матричное уравнение вида (12.23) с элементами u_n, q_n и

$$b_m = \int_V f(M) w_m(M) dV(M), \quad m = \overline{1, N_S}, \quad (12.35)$$

матриц-столбцов U, Q и B соответственно. При этом элементами матрицы H порядка N_S будут значения

$$h_{mn} = \int_{S_n} w_m^*(P) dS(P), \quad m \neq n, \quad \text{и} \quad h_{mm} = - \sum_{\substack{n=\overline{1, N_S} \\ n \neq m}} h_{mn},$$

а элементами матрицы G порядка N_S — значения

$$g_{mn} = \int_{S_n} w_m(P) dS(P), \quad m, n = \overline{1, N_S}. \quad (12.36)$$

В (12.35) $w_m(M) = w(M, P_m) = \frac{1}{r(M, P_m)}$, где $r(M, P_m)$ — расстояние между точкой $P_m \in S_m$, выбранной в середине $\Gamma \cap S_m$, и точкой $M \in V$. Тот же смысл имеет $w_m(P)$ в выражении для g_{mn} в (12.36), а в выражении для h_{mn} через $w_m^*(P)$ обозначена производная $(\nabla w_m(P)) \mathbf{n}(P)$ функции $w_m(P)$ в направлении единичного вектора $\mathbf{n}(P)$ внешней нормали к поверхности S в точке $P \in S_n$. Отметим, что нахождение диагонального элемента h_{mm} через внедиагональные элементы в m -й строке матрицы H позволяет избежать вычисления телесного угла $\omega(P_m)$ в случае, если в точке $P_m \in S_m \subset S$ поверхность S не является гладкой.

Трудности, возникающие при вычислении g_{mm} , можно преодолеть так же, как это было сделано для двумерной задачи, представив $\Gamma \cap S_m$ в окрестности точки $P_m \in S_m$ участком плоскости. При вычислении интеграла в (12.35) целесообразно, как и в случае двумерной задачи, применить квадратурные формулы Гаусса или же получить в аналитической форме оценку вклада в этот интеграл в малой окрестности точки $P_m \in S_m$.

Используя (12.20), можно преобразовать (12.23) в СЛАУ вида (12.29) относительно N_S неизвестных значений q_n на участках $S_* \subset S$ поверхности S и u_n на участках $S_* = S \setminus S^*$. При этом в (12.30) и (12.31) Γ^* следует заменить на S^* , а Γ_* — на S_* .

После нахождения из решения СЛАУ вида (12.29) недостающих значений u_n и q_n , $n = \overline{1, N_S}$, можно, используя (12.19) при $\omega(M_0) = 4\pi$, вычислить значение $u(M_0)$ искомой функции в любой точке $M_0 \in V$:

$$u(M_0) \approx \frac{1}{4\pi} \sum_{n=1}^{N_S} \left(q_n \int_{S_n} w(P, M_0) dS(P) - u_n \int_{S_n} w^*(P, M_0) dS(P) \right) + \frac{1}{4\pi} \int_V f(M) w(M, M_0) dV(M). \quad (12.37)$$

Точность вычислений по (12.37) зависит от погрешности аппроксимации функций $u(P)$ и $q(P)$ их постоянными значениями в пределах каждого ГЭ и от точности вычисления входящих в (12.37) интегралов. По аналогии с (12.28) можно воспользоваться и менее точной формулой

$$u(M_0) \approx \frac{1}{4\pi} \sum_{n=1}^{N_S} (q_n w(P_n, M_0) - u_n w^*(P_n, M_0)) S_n + \\ + \frac{1}{4\pi} \int_V f(M) w(M, M_0) dV(M), \quad (12.38)$$

где P_n — точка, выбранная в середине n -го ГЭ площадью S_n .

Несобственный интеграл по области V в (12.37) и (12.38) сходится равномерно в окрестности точки $M_0 \in V$, если в этой окрестности функция $f(M)$ непрерывна [XII]. Поэтому его можно дифференцировать по координатам точки M_0 как по параметрам, что позволяет вычислять производные искомой функции $u(M_0)$.

Аппроксимации ГЭ функций $u(P)$ и $q(P)$ многочленами в пределах двумерных ГЭ S_n , $n = \overline{1, N_S}$, на поверхности S и вычисление возникающих при этом интегралов аналогичны соответствующим процедурам, используемым в *методе конечных элементов* (см. 10.2). Тогда неизвестными в СЛАУ вида (12.29) будут значения этих функций в узлах таких ГЭ.

Замечание 12.1. Известно*, что решение краевой задачи для уравнения Лапласа можно представить в виде *потенциалов простого или двойного слоя* с плотностями, удовлетворяющими ГИУ. Эти ГИУ также могут быть решены при помощи МГЭ. Однако плотности потенциалов обычно не имеют определенного физического смысла, а их нахождение является промежуточным звеном, дающим возможность затем вычислить значения

*См.: Мизлин С.Г., 1968.

искомых функций, имеющих конкретную физическую интерпретацию в содержательной постановке решаемой задачи. В связи с этим способы, позволяющие найти плотности потенциалов в качестве вспомогательных функций, относят к непрямым МГЭ в отличие от прямых МГЭ, которые направлены непосредственно на вычисление значений функций, имеющих определенный физический смысл. В частности, к прямым следует отнести вариант МГЭ, рассмотренный в этом параграфе.

12.3. Учет анизотропии и неоднородности

Уравнение Пуассона (12.18) описывает широкий класс физических процессов в изотропной однородной среде (см. 1), т.е. в среде, свойства которой не зависят ни от направления, ни от координат точки $M \in V$ в рассматриваемой области V . Однако при решении прикладных задач довольно часто возникает необходимость учета такой зависимости, что в некоторых случаях удается осуществить в рамках *метода граничных элементов* (МГЭ). Рассмотрим эти возможности на примере решения при помощи МГЭ задачи стационарной теплопроводности в трехмерной области V .

Пусть в области V , ограниченной кусочно гладкой поверхностью S , находится неподвижная однородная анизотропная среда. Свойство этой среды передавать тепловую энергию характеризует *тензор $\tilde{\lambda}$ теплопроводности с компонентами $\lambda_{ij} = \lambda_{ji} = \text{const}$, $i, j = \overline{1, 3}$* , составляющими симметрическую матрицу (2.95) третьего порядка. Если ориентацию осей Ox_i декартовой прямоугольной системы координат выбрать так, чтобы они совпадали с *главными осями тензора $\tilde{\lambda}$* , то процесс стационарной теплопроводности в области V будет описывать дифференциальное уравнение с частными производными

$$\sum_{i=1}^3 \lambda_{(i)} \frac{\partial^2 T(\Lambda i)}{\partial x_i^2} + I_V^{(q)}(M) = 0, \quad M \in V, \quad (12.39)$$

где $T(M)$ и $I_V^{(q)}(M)$ — температура среды и объемная мощность внутренних источников энерговыделения, зависящие от координат точки $M \in V$, а $\lambda^{(i)} = \text{const}$ — главные коэффициенты теплопроводности анизотропной среды. Отметим, что (12.39) вытекает из равенства нулю левой части формул (1.31), (1.32) (см. замечание 2.1).

Изменим масштабы по осям координат в соответствии с заменой переменных $\hat{x}_i = x_i \sqrt{\frac{\lambda^\circ}{\lambda^{(i)}}}$, где $\lambda^\circ = \text{const}$ — некоторый масштаб измерения главных коэффициентов теплопроводности, в качестве которого можно выбрать одно из значений $\lambda^{(i)}$. Тогда вместо (12.39) получим

$$\lambda^\circ \sum_{i=1}^3 \frac{\partial^2 T(\bar{M})}{\partial \hat{x}_i^2} + I_V^{(q)}(\bar{M}) = 0, \quad \bar{M} \in \bar{V},$$

или

$$\hat{\nabla}^2 T(\bar{M}) + \frac{I_V^{(q)}(\bar{M})}{\lambda^\circ} = 0, \quad \bar{M} \in \bar{V}, \quad (12.40)$$

где $\hat{\nabla}^2$ — оператор Лапласа в системе координат $\hat{O}\hat{x}_1\hat{x}_2\hat{x}_3$ (рис. 12.8), а $\bar{M}$ — точка, соответствующая точке $M \in V$ после изменения масштабов по осям координат и преобразования области V в область $\bar{V}$ (на рис. 12.8 принято $\lambda^\circ = \lambda^{(1)}$ и $\lambda^{(2)} < \lambda^{(1)} < \lambda^{(3)}$).

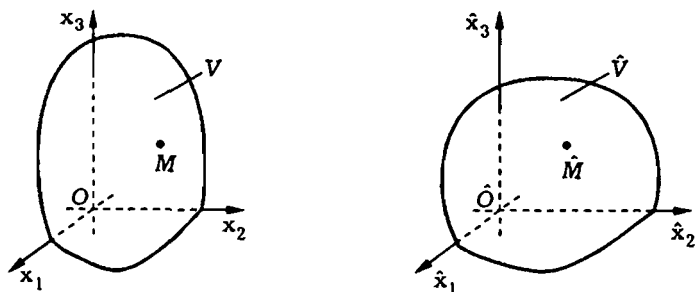


Рис. 12.8

Получившееся в итоге преобразования уравнение Пуассона (12.40) в сочетании с соответствующим образом преобразованными граничными условиями можно решить в области $\hat{V}$ при помощи МГЭ (см. 12.2), а затем результаты решения перенести в исходную область V . Однако такой путь не всегда удобен, особенно в случае области сложной конфигурации. Но для решения (12.39) в исходной области V необходимо располагать *фундаментальным решением* в пространстве $\mathbb{R}^3$ однородного уравнения

$$\sum_{i=1}^3 \lambda^{(i)} \frac{\partial^2 T(M)}{\partial x_i^2} = 0, \quad M \in V. \quad (12.41)$$

Убедимся, что функция

$$\tilde{w}(M, M_0) = \left(\sum_{j=1}^3 (x_j(M) - x_j(M_0))^2 \frac{\lambda^\circ}{\lambda^{(j)}} \right)^{-1/2} \quad (12.42)$$

является таким фундаментальным решением, т.е. удовлетворяет (12.41) во всех точках $M \in \mathbb{R}^3$, за исключением точки $M_0 \in \mathbb{R}^3$, в которой $w(M, M_0) \rightarrow \infty$. Действительно, поместим начало системы координат $Ox_1x_2x_3$ для упрощения выкладок в точку M_0 , т.е. положим $x_j(M_0) = 0$, $j = \overline{1, 3}$, и вычислим

$$\frac{\partial \tilde{w}}{\partial x_i} = -\frac{1}{2} \left(\sum_{j=1}^3 x_j^2 \frac{\lambda^\circ}{\lambda^{(j)}} \right)^{-3/2} 2x_i \frac{\lambda^\circ}{\lambda^{(i)}} = -\tilde{w}^3 x_i \frac{\lambda^\circ}{\lambda^{(i)}},$$

$$\frac{\partial^2 \tilde{w}}{\partial x_i^2} = -3\tilde{w}^2 \frac{\partial \tilde{w}}{\partial x_i} x_i \frac{\lambda^\circ}{\lambda^{(i)}} - \tilde{w}^3 \frac{\lambda^\circ}{\lambda^{(i)}} = \tilde{w}^3 \frac{\lambda^\circ}{\lambda^{(i)}} \left(3\tilde{w}^2 x_i^2 \frac{\lambda^\circ}{\lambda^{(i)}} - 1 \right),$$

$$\sum_{i=1}^3 \lambda^{(i)} \frac{\partial^2 \tilde{w}}{\partial x_i^2} = \tilde{w}^3 \lambda^\circ \left(3\tilde{w}^2 \sum_{i=1}^3 x_i^2 \frac{\lambda^\circ}{\lambda^{(i)}} - 3 \right) = 0$$

при $\tilde{w} < \infty$.

Отметим, что совмещение координатных осей и главных осей тензора $\tilde{\lambda}$ не является обязательным условием построения фундаментального решения однородного уравнения стационарной теплопроводности в анизотропной среде. При произвольной ориентации главных осей тензора $\tilde{\lambda}$ относительно координатных осей из (2.96) получим

$$\sum_{i=1}^3 \sum_{j=1}^3 \frac{\partial}{\partial x_i} \left(\lambda_{ij} \frac{\partial T(M)}{\partial x_j} \right) + I_V^{(q)}(M) = 0, \quad M \in V. \quad (12.43)$$

Можно показать, что функция

$$\hat{w}(M, M_0) = \left(\sum_{i=1}^3 \sum_{j=1}^3 \rho_{ij} (\Delta x_i(M) \Delta x_j(M))^2 \right)^{-1/2}, \quad (12.44)$$

где $\Delta x_i(M) = x_i(M) - x_i(M_0)$ и $\rho_{ij} = \rho_{ji} = \text{const}$ — компоненты симметрического тензора коэффициентов термического сопротивления среды, обратного к тензору $\tilde{\lambda}$, удовлетворяет (12.43) при $I_V^{(q)}(M) = 0$ во всех точках $M \in \mathbb{R}^3$, кроме точки M_0 , в которой $\hat{w}(M, M_0) \rightarrow \infty$.

В формуле Грина вида (2.99) положим $m = 3$, $a_{ij} = \lambda_{ij}$, $u = T(M)$, $v = \hat{w}(M, M_0)$, а Ω и $\partial\Omega$ отождествим с V и S соответственно. Тогда (2.99) можно использовать аналогично процедуре, рассмотренной выше (см. 12.1), для перехода от (12.43) к *граничному интегральному уравнению* (ГИУ)

$$\begin{aligned} \omega(M_0) u(M_0) = & \int_S (\hat{q}(P) \hat{w}(P, M_0) - T(P) \hat{w}^*(P, M_0)) dS(P) + \\ & + \int_V I_V^{(q)}(M) \hat{w}(M, M_0) dV(M), \quad M \in V, \quad P \in S, \quad M_0 \in \bar{V} = V \cup S. \end{aligned}$$

Здесь $\omega(M_0)$ — телесный угол, под которым из точки $M_0 \in \bar{V}$ видна внутренность области V (см. 12.2),

$$\hat{q}(P) = \sum_{i=1}^3 \sum_{j=1}^3 \lambda_{ij} \frac{\partial T(P)}{\partial x_j} n_i(P), \quad P \in S,$$

$$\hat{w}^*(P, M_0) = \sum_{i=1}^3 \sum_{j=1}^3 \lambda_{ij} \frac{\partial \hat{w}(P, M_0)}{\partial x_j} n_i(P), \quad P \in S,$$

где $n_i(P)$ — проекция на координатную ось Ox_i единичного вектора $\mathbf{n}(P)$ внешней нормали к поверхности S в точке $P \in S$.

Как и ранее (см. 12.2), от этого ГИУ можно перейти к системе линейных алгебраических уравнений (СЛАУ) (12.29) при заданных на участках S^* и S_* поверхности S граничных условиях вида (12.20)

$$\begin{cases} T(P) = f_1(P), & P \in S^* \subset S; \\ \hat{q}(P) + \beta(P)T(P) = f_2(P), & P \in S_* = S \setminus S^*, \end{cases} \quad (12.45)$$

где f_1, f_2, β — известные функции положения точек на границе S . Решение СЛАУ относительно неизвестных *узловых значений* T_n и $\hat{q}_n$, $n = \overline{1, N_S}$, при разбиении поверхности S на N_S *граничных элементов* (ГЭ) позволит затем по формуле вида (12.37) вычислить значение температуры в любой точке $M_0 \in V$:

$$\begin{aligned} T(M_0) \approx \frac{1}{4\pi} \sum_{n=1}^{N_S} \left(\hat{q}_n \int_{S_n} \hat{w}(P, M_0) dS(P) - T_n \int_{S_n} \hat{w}^*(P, M_0) dS(P) \right) + \\ + \frac{1}{4\pi} \int_V I_V^{(q)}(M) \hat{w}(M, M_0) dV(M). \end{aligned} \quad (12.46)$$

Отсюда можно перейти к более простой, но менее точной формуле вида (12.38).

Пусть теперь требуется найти решение уравнения Пуассона вида (12.18)

$$\nabla^2 T(M) + \frac{I_V^{(q)}(M)}{\lambda} = 0, \quad M \in V, \quad (12.47)$$

в трехмерной области V , состоящей из двух подобластей V' и V'' , имеющих общую границу в виде кусочно гладкой поверхности S^* (рис. 12.9). В каждой из подобластей среду будем считать изотропной и однородной, имеющей постоянные, но различные коэффициенты теплопроводности $\lambda = \lambda'$ и $\lambda = \lambda''$ (далее одним и двумя штрихами отмечены величины, относящиеся к подобластям V' и V'' соответственно). Это означает, что S^* является в данном случае *поверхностью разрыва*, причем *сильного*, если тепловой контакт между подобластями отличается в общем случае от идеального, т.е. в соответствии с (2.65)

$$\lambda' q'(P) = \alpha_k (T'(P) - T''(P)) = -\lambda'' q''(P), \quad P \in S^*, \quad (12.48)$$

где $\alpha_k \geq 0$ — коэффициент контактной проводимости (см. 2.3),

$$q'(P) = (\nabla T'(P)) \mathbf{n}'(P), \quad q''(P) = (\nabla T''(P)) \mathbf{n}''(P), \quad P \in S^*,$$

$\mathbf{n}'(P)$ и $\mathbf{n}''(P)$ — единичные векторы внешней по отношению к соответствующим подобластям нормали к поверхности S^* в точке $P \in S^*$. Напомним (см. пример 2.4), что при $\alpha_k = 0$

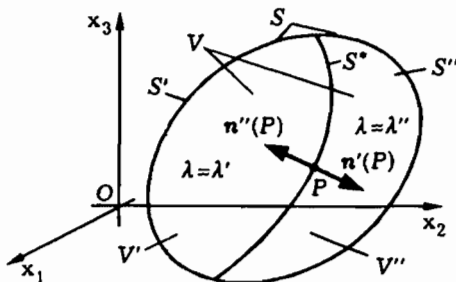


Рис. 12.9

участки поверхности S^* будут идеально теплоизолированными, т.е. $q'(P) = q''(P) = 0$, а при $\alpha_k \rightarrow \infty$ тепловой контакт будет идеальным, т.е. $T'(P) = T''(P)$, и соответствующие участки S^* составят поверхность слабого разрыва.

Используя в каждой из подобластей фундаментальное решение уравнения Лапласа в $\mathbb{R}^3$, имеющее вид $w(M, M_0) = \frac{1}{r(M, M_0)}$, где $r(M, M_0)$ — расстояние между точками M и M_0 в $\mathbb{R}^3$, можно перейти к ГИУ вида (12.19)

$$\begin{aligned} \omega'(M_0)T'(M_0) = \lambda' \int_{\tilde{S}'} (q'(P)w(P, M_0) - T'(P)w^*(P, M_0)) dS(P) + \\ + \int_{V'} I_V^{(q)}(M)w(M, M_0) dV(M), \quad M \in V', \quad P \in \tilde{S}', \quad M_0 \in \bar{V}', \quad (12.49) \end{aligned}$$

$$\begin{aligned} \omega'(M_0)T''(M_0) = \lambda'' \int_{\tilde{S}''} (q''(P)w(P, M_0) - T''(P)w^*(P, M_0)) dS(P) + \\ + \int_{V''} I_V^{(q)}(M)w(M, M_0) dV(M), \quad M \in V'', \quad P \in \tilde{S}'', \quad M_0 \in \bar{V}'', \quad (12.50) \end{aligned}$$

где $\tilde{S}' = S' \cup S^*$, $\tilde{S}'' = S'' \cup S^*$, причем S' и S'' — участки внешней границы подобластей V' и V'' соответственно, составляющие поверхность $S = S' \cup S''$, $S' \cap S'' = \emptyset$ (см. рис. 12.9).

Разобьем поверхности S' , S'' и S^* на N' , N'' и N^* граничных элементов соответственно с постоянными в пределах каждого ГЭ S_n значениями T_n и q_n , $n = \overline{1, N}$, $N = N' + N'' + N^*$. При переходе с использованием заданных на поверхности S граничных условий от ГИУ (12.49) и (12.50) к двум СЛАУ вида (12.29) общее число неизвестных в каждой СЛАУ будет на N^* превышать число уравнений. Для объединения обеих СЛАУ в одну квадратную необходимо использовать условия (12.48) на поверхности S^* , что даст недостающие $2N^*$ уравнений. После

решения квадратной СЛАУ относительно неизвестных значений T_n и q_n можно вычислить значение температуры в любой точке $M_0 \in V$. Например, для точки $M_0 \in V'$ получим

$$T'(M_0) \approx \frac{1}{4\pi} \int_{V'} I_V^{(q)}(M) w(M, M_0) dV(M) + \\ + \frac{\lambda'}{4\pi} \sum_{n=1}^{N'+N^*} \left(q'_n \int_{S_n} w(P, M_0) dS(P) - T'_n \int_{S_n} w^*(P, M_0) dS(P) \right).$$

Отсюда можно перейти к более простой формуле вида (12.38).

Ясно, что рассмотренную процедуру учета неоднородности области V можно использовать и в случае, если V состоит из более чем двух однородных подобластей с различными коэффициентами теплопроводности. Однако в прикладных задачах коэффициент теплопроводности $\lambda(M)$ может быть кусочно непрерывной функцией положения точки $M \in V$. Такая ситуация характерна, например, для нелинейной задачи теплопроводности в неоднородной области, когда коэффициент теплопроводности среды зависит от температуры и при решении задачи последовательными приближениями на очередной итерации может быть представлен кусочно непрерывной функцией $\lambda(M)$, $M \in \bar{V}$. В этом случае вместо (12.47) будем иметь уравнение

$$\nabla(\lambda(M)\nabla T(M)) + I_V^{(q)}(M) = 0, \quad M \in V. \quad (12.51)$$

Важно отметить, что функция $I_V^{(q)}(M)$ в (12.51) при решении нелинейной задачи последовательными приближениями может отражать на очередной итерации зависимость объемной мощности внутренних источников энергоснабжения не только от координат точки $M \in V$, но и от значений искомой функции $T(M)$ температуры.

Применить МГЭ непосредственно к решению (12.51) не удастся, так как для произвольной функции $\lambda(M)$ неизвестно фундаментальное решение соответствующего однородного уравнения $\nabla(\lambda(M)\nabla T(M)) = 0$. Но при условии, что функция $\lambda(M)$

имеет в области V кусочно непрерывные производные, (12.51) можно преобразовать, выделив в явном виде оператор Лапласа:

$$\lambda(M) \nabla^2 T(M) + (\nabla \lambda(M)) \nabla T(M) + I_V^{(q)}(M) = 0, \quad M \in V.$$

Отсюда, учитывая, что $\lambda(M) > 0$, получаем

$$\nabla^2 T(M) + f(M, T) = 0, \quad M \in V, \quad (12.52)$$

где

$$f(M, T) = \frac{1}{\lambda(M)} \left((\nabla \lambda(M)) \nabla T(M) + I_V^{(q)}(M) \right), \quad M \in V.$$

Отличие уравнения (12.52) от уравнения (12.18) состоит лишь в зависимости от температуры функции f . Эту зависимость при использовании МГЭ можно учесть последовательными приближениями.

Пусть $T^{(k)}(M)$, $M \in V$, — распределение температуры, найденное на k -й итерации (при $k=0$ оно соответствует начальному приближению), а $f_k(M) = f(M, T^{(k)}(M))$. Тогда для нахождения следующего, $(k+1)$ -го приближения используем ГИУ вида (12.19)

$$\begin{aligned} \omega(M_0) T^{(k+1)}(M_0) = \\ = \int_S (q^{(k+1)}(P) w(P, M_0) - T^{(k+1)}(P) w^*(P, M_0)) dS(P) + \\ + \int_V f_k(M) w(M, M_0) dV(M), \quad M \in V, \quad P \in S, \quad M_0 \in \bar{V}. \end{aligned} \quad (12.53)$$

Процедура перехода от уравнения (12.53) к СЛАУ вида (12.29) описана выше (см. 12.2). Сходимость последовательных приближений проще всего контролировать по максимальному значению $|T^{(k+1)}(M) - T^{(k)}(M)|$, $M \in V$.

12.4. Нестационарные задачи

Рассмотрим особенности использования *метода граничных элементов* (МГЭ) для решения нестационарных задач на примере задачи нестационарной теплопроводности в однородной трехмерной области V , ограниченной кусочно гладкой поверхностью S (рис. 12.10). Пусть функция $T(t, M)$, описывающая изменение во времени t температуры среды в точке $M \in \bar{V} = V \cup S$, удовлетворяет уравнению вида (2.54)

$$c \frac{\partial T(t, M)}{\partial t} = \lambda \nabla^2 T(t, M) + I_V^{(q)}(t, M), \quad M \in V, \quad (12.54)$$

где $c = \text{const}$ — объемная теплоемкость среды; $\lambda = \text{const}$ — коэффициент теплопроводности среды; $I_V^{(q)}$ — объемная мощность внутренних источников энерговыделения. Кроме того, заданы начальное условие $T(0, M) = T^0(M)$ в момент времени $t = 0$, принимаемый за начальный, и граничные условия

$$\begin{cases} T(t, P) = f_1(t, P), & P \in S^* \subset S; \\ \lambda q(t, P) + \beta(t, P) T(t, P) = f_2(t, P), & P \in S_* = S \setminus S^*, \end{cases} \quad (12.55)$$

на участках S^* и S_* поверхности S , где f_1 , f_2 и β — известные функции времени и положения точки P на поверхности, а

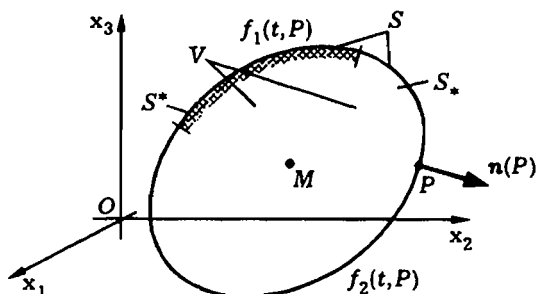


Рис. 12.10

$q(t, P) = (\nabla T(t, M)) \mathbf{n}(P)$ — проекция градиента температуры на направление единичного вектора $\mathbf{n}(P)$ внешней нормали к поверхности S .

Однородному уравнению (12.54) (при $I_V^{(q)}(t, M) \equiv 0$) во всех точках $M \in \mathbb{R}^3$, кроме точки $M_0 \in \mathbb{R}^3$, удовлетворяет при $t > \tau$ функция [XII]

$$w(t - \tau, M, M_0) = \frac{1}{8(\pi \tilde{a}(t - \tau))^{3/2}} \exp\left(-\frac{r(M, M_0)}{4\tilde{a}(t - \tau)}\right), \quad (12.56)$$

где $\tilde{a} = \lambda/c$, $r(M, M_0)$ — расстояние между точками M и M_0 . Эту функцию называют *фундаментальным решением уравнения теплопроводности* в пространстве. Она описывает изменение температуры при $t > \tau$ в точке M , вызванное мгновенным выделением в момент времени $t = \tau$ в точке M_0 количества теплоты, численно равное c .

Используя (12.56), задачу (12.54), (12.55) можно свести к *граничному интегральному уравнению* (ГИУ)*.

$$\begin{aligned} \omega(M_0) T(t, M_0) = & \int_V T^\circ(M) w(t, M, M_0) dV + \\ & + \tilde{a} \int_0^t \int_S (q(\tau, P) w(t - \tau, P, M_0) - T(\tau, P) w^*(t - \tau, M, M_0)) d\tau dS(P) + \\ & + \tilde{a} \int_0^t \int_V I_V^{(q)}(\tau, M) w(t - \tau, M, M_0) d\tau dV, \quad (12.57) \end{aligned}$$

где $w^*(t - \tau, P, M_0) = (\nabla w(t - \tau, P, M_0)) \mathbf{n}(P)$. Если промежуток времени, в течение которого рассматривается процесс нестационарной теплопроводности, разбить на интервалы $\Delta t_k = t_k - t_{k-1}$ и в пределах каждого k -го интервала принять функ-

*См., например: Карташов Э.М.

ции $T(t, M)$, $q(t, P)$ и $I_V^{(q)}(t, M)$ не зависящими от времени, заменив их на $T_k(M) = T(t_k, M)$, $q_k(P) = q(t_k, P)$ и $I_V^{(q)}(t_k, M)$ соответственно, то вместо (12.57) получим ГИУ

$$\begin{aligned} \omega(M_0) T_k(M_0) = & \int_V T(t_{k-1}, M) w(\Delta t_k, M, M_0) dV + \\ & + \tilde{a} \int_S q_k(P) \left(\int_{t_{k-1}}^{t_k} w(t_k - \tau, P, M_0) d\tau \right) dS(P) - \\ & - \tilde{a} \int_S T_k(P) \left(\int_{t_{k-1}}^{t_k} w^*(t_k - \tau, P, M_0) d\tau \right) dS(P) + \\ & + \int_V I_V^{(q)}(t_k, M) \left(\int_{t_{k-1}}^{t_k} w(t_k - \tau, M, M_0) d\tau \right) dV(M). \quad (12.58) \end{aligned}$$

Для численного решения ГИУ (12.58) при помощи МГЭ поверхность S разобьем на *граничные элементы* (ГЭ) с общим числом N_S граничных узлов $P_n \in S$, $n = \overline{1, N_S}$, а область V — на *конечные элементы* с общим числом N_V внутренних узлов $M_l \in V$, $l = \overline{1, N_V}$. Тогда (12.58) можно свести к матричному уравнению

$$H_k U_k = G_k Q_k + D_k W_{k-1} + B_k, \quad (12.59)$$

где U_k и Q_k — матрицы-столбцы размера $N_S \times 1$ с элементами $T_k(P_n)$ и $q_k(P_n)$ соответственно; H_k и G_k — матрицы порядка N_S ; W_{k-1} — матрица-столбец размера $(N_S + N_V) \times 1$, элементами которой являются известные значения температур в граничных и внутренних узлах в момент времени t_{k-1} в начале k -го интервала (при $k = 1$ эти значения соответствуют начальному распределению температуры); D_k — матрица раз-

мера $N_S \times (N_S + N_V)$ с элементами

$$d_{ml}^{(k)} = \int_V \varphi_l(M) w(\Delta t_k, M, P_m) dV(M),$$

где $\varphi_l(M)$ — функция формы, соответствующая внутреннему или граничному узлу с номером l и использованная при аппроксимации функции $T(t_{k-1}, M)$, $M \in V$; B_k — матрица-столбец размера $N_S \times 1$, элементы которой

$$b^{(k)} = \int_V I_V^{(q)}(t_k, M) \left(\int_{t_{k-1}}^{t_k} w(t_k - \tau, M, P_m) d\tau \right) dV(M)$$

отражают влияние внутренних источников энерговыделения на значения $T_k(P_n)$ и $q_k(P_n)$, $n = \overline{1, N_S}$.

Значения элементов матриц H_k , G_k и D_k зависят, как это следует из (12.58), от значения Δt_k , конфигурации области V , расположения граничных и внутренних узлов, а также от способа аппроксимации функций $T_k(P)$, $q_k(P)$ в пределах каждого ГЭ и функции $T(t_{k-1}, M)$ в пределах каждого конечного элемента. При $\Delta t_k = \text{const}$ элементы этих матриц не будут зависеть от номера интервала, так что нижний индекс k в их обозначениях может быть опущен.

Как и в случае решения уравнения Пуассона (см. 12.2) интегралы, входящие в выражения диагональных элементов $h_{mm}^{(k)}$ и $g_{mm}^{(k)}$ матриц H_k и G_k , являются несобственными. В малой окрестности m -го узла ГЭ можно заменить плоским участком и вычислить вклад в элемент $g_{mm}^{(k)}$ аналитически, а для преодоления трудностей, возникающих при вычислении $h_{mm}^{(k)}$, рассмотрим случай $I_V^{(q)}(t, M) \equiv 0$, $S^* = S$ и $f_1(t, P) = T^0(M) = T_0 = \text{const}$. Тогда температура в области V и на ее поверхности будет неизменной во времени и равной T_0 , а матрицы Q_k и B_k станут нулевыми. Тогда из (12.59) получим

$H_k U_k = D_k W_{k-1}$, причем все элементы матриц U_k и W_{k-1} равны T_0 . Отсюда следует, что

$$h_{mm}^{(k)} = - \sum_{\substack{n=\overline{1, N_S} \\ n \neq m}} h_{mn}^{(k)} + \sum_{l=1}^{N_S+N_V} d_{ml}^{(k)}, \quad m = \overline{1, N_S},$$

для любого k -го интервала.

В простейшем варианте МГЭ, когда в пределах каждого ГЭ S_n , $n = \overline{1, N_S}$, функции $T_k(P)$ и $q_k(P)$ совпадают со значениями $T_k(P_n)$ и $q_k(P_n)$, $P_n \in S_n$, соответственно, из (12.58) получим

$$h_{mn}^{(k)} = \tilde{a} \int_S \int_{t_{k-1}}^{t_k} w^*(t_k - \tau, P, M_0) d\tau dS(P),$$

$$g_{mn}^{(k)} = \tilde{a} \int_S \int_{t_{k-1}}^{t_k} w(t_k - \tau, P, M_0) d\tau dS(P).$$

В частном случае задания на всей поверхности S температуры $T(t, P) = f_1(t, P)$ все элементы матрицы U_k известны и из (12.59) можно найти

$$Q_k = G_k^{-1} (H_k U_k - D_k W_{k-1} - B_k),$$

т.е. вычислить значения $q_k(P_n)$, $P_n \in S$, $n = \overline{1, N_S}$. Это позволит, используя (12.58) при $\omega(M_0) = 4\pi$, вычислить температуру $T_k(M_0)$ в любой точке $M_0 \in V$ в момент времени t_k в конце k -го интервала.

Другой путь применения МГЭ к решению нестационарной задачи (12.54), (12.55) состоит в предварительной замене в (12.54) производной $\frac{\partial T(t, M)}{\partial t}$ конечноразностным соотношением $\frac{T_k(M) - T_{k-1}(M)}{\Delta t_k}$, т.е. связан с обращением к *методу прямых*. Тогда из (12.54) для момента времени t_k получим уравнение

$$\nabla^2 T_k(M) - \frac{1}{\tilde{a} \Delta t_k} T_k(M) = -f_k(M), \quad M \in V, \quad (12.60)$$

где

$$f_k(M) = \frac{T_{k-1}(M)}{\tilde{a}\Delta t_k} + \frac{I_V^{(g)}(t_k, M)}{\lambda},$$

а из (12.55) — граничные условия

$$\begin{cases} T_k(P) = f_1(t_k, P), & P \in S^* \subset S; \\ \lambda q_k(P) + \beta(t_k, P) T_k(P) = f_2(t_k, P), & P \in S_* = S \setminus S^*. \end{cases} \quad (12.61)$$

Непосредственной проверкой можно убедиться, что фундаментальным решением однородного уравнения (12.60) при $f_k(M) \equiv 0$ будет функция

$$w_k(M, M_0) = \frac{1}{r(M, M_0)} \exp\left(-\frac{r(M, M_0)}{\sqrt{\tilde{a}\Delta t_k}}\right). \quad (12.62)$$

Теперь задачу (12.60), (12.61) можно свести к ГИУ

$$\begin{aligned} \omega(M_0) T_k(M_0) = & \int_S (q_k(P) w_k(P, M_0) - T_k(P) w_k^*(P, M_0)) dS(P) + \\ & + \int_V f_k(M) w_k(M, M_0) dV(M), \quad M \in V, \quad P \in S, \quad M_0 \in \bar{V}, \end{aligned} \quad (12.63)$$

где $w_k^*(P, M_0) = (\nabla w_k(P, M_0)) \mathbf{n}(P)$. Как и выше (см. 12.2), можно перейти от (12.63) к системе линейных алгебраических уравнений (СЛАУ) вида (12.29) относительно неизвестных значений $T_k(P_n)$ и $q_k(P_n)$ в узлах $P_n \in S$, $n = \overline{1, N_S}$, ГЭ, на которые следует разбить поверхность S .

Еще одна возможность решения задачи (12.54), (12.55) при помощи МГЭ связана с использованием *интегрального преобразования Лапласа*. Применяя это преобразование к (12.54), (12.55) и учитывая начальное условие, получаем уравнение в изображениях

$$\nabla^2 T(s, M) - \frac{s}{\tilde{a}} T(s, M) = -f(s, M), \quad M \in V, \quad (12.64)$$

где

$$T(s, M) = \int_0^{\infty} T(t, M) e^{-st} dt,$$

$$f(s, M) = \frac{T^0(M)}{\tilde{a}} + \frac{1}{\lambda} \int_0^{\infty} I_V^{(q)}(t, M) e^{-st} dt,$$

с граничными условиями (в случае неизменного во времени коэффициента теплообмена $\beta(P)$)

$$\begin{cases} T(s, P) = \int_0^{\infty} f_1(t, P) e^{-st} dt, & P \in S^*, \\ \lambda q(s, P) + \beta(P) T(s, P) = \int_0^{\infty} f_2(t, P) e^{-st} dt, & P \in S_*. \end{cases} \quad (12.65)$$

Здесь

$$q(s, P) = \int_0^{\infty} q(t, P) e^{-st} dt.$$

По аналогии с (12.62) фундаментальным решением однородного уравнения (12.64) (при $f(s, M) \equiv 0$) будет функция

$$\bar{w}(M, M_0) = \frac{1}{r(M, M_0)} \exp\left(-\sqrt{\frac{s}{\tilde{a}}} r(M, M_0)\right),$$

что позволяет от задачи (12.64), (12.65) перейти к ГИУ

$$\begin{aligned} \omega(M_0) T(s, M_0) = & \int_S (q(s, P) \bar{w}(P, M_0) - T(s, P) \bar{w}^*(P, M_0)) dS(P) + \\ & + \int_V f(s, M) \bar{w}(M, M_0) dV(M), \quad M \in V, \quad P \in S, \quad M_0 \in \bar{V}, \end{aligned} \quad (12.66)$$

где $\bar{w}^*(P, M_0) = (\nabla \bar{w}(P, M_0)) \mathbf{n}(P)$.

При выделении N_S граничных узлов $P_n \in S$, $n = \overline{1, N_S}$, (12.66) можно свести к матричному уравнению вида (12.23)

$$H_s U_s = G_s Q_s + B_s, \quad (12.67)$$

в котором элементы всех матриц будут зависеть от параметра s интегрального преобразования Лапласа. С учетом (12.65) из (12.67) следует в матричной форме зависимость от s изображений $T_n(s)$ и $q_n(s)$ в граничных узлах, а затем с использованием (12.66) — и зависимость от s изображений $T(s, M_0)$ во внутренних точках $M_0 \in V$. Для перехода к оригиналам $T_n(t)$ и $T(t, M_0)$ вследствие матричной формы зависимости изображений от s приходится применять численные способы обращения преобразования Лапласа*.

12.5. Статическая задача теории упругости

Применение *метода граничных элементов* (МГЭ) к решению задач теории упругости также основано на возможности перехода от формулировки краевой задачи, содержащей дифференциальные уравнения и граничные условия, к *граничному интегральному уравнению* (ГИУ). Однако в отличие от уравнения Пуассона для такого перехода вместо второй формулы Грина в случае линейно упругой среды необходимо использовать *теорему взаимности работ*, суть которой сводится к следующему.

Пусть линейно упругая среда занимает область $V \subset \mathbb{R}^3$, ограниченную кусочно гладкой поверхностью S . Свойства среды описывают константы Ламе $\tilde{\lambda}$ и $\tilde{\mu}$, устанавливающие связь между компонентами σ_{ij} и ε_{ij} , $i, j = \overline{1, 3}$, тензоров напряжений и деформаций соответственно в виде обобщенного закона Гука (3.31). Рассмотрим два состояния равновесия этой среды. Первое состояние характеризуется проекциями u_i вектора u перемещения частиц среды на оси Ox_i декартовой прямоугольной

*См., например. Крылов В.И., Скобля Н.С.

системы координат, компонентами ε_{ij} и σ_{ij} , $i, j = \overline{1, 3}$, тензоров деформаций и напряжений соответственно, а второе — величинами $\bar{u}_i$, $\bar{\varepsilon}_{ij}$ и $\bar{\sigma}_{ij}$.

Если эти состояния вызваны действием двух систем объемных и поверхностных сил с проекциями b_i , p_i и $\bar{b}_i$, $\bar{p}_i$ соответственно, то должны быть выполнены уравнения равновесия (здесь и далее в этом параграфе принято правило суммирования по повторяющимся нижним индексам $i, j = \overline{1, 3}$)

$$\frac{\partial \sigma_{ji}(M)}{\partial x_j} + b_i(M) = 0, \quad \frac{\partial \bar{\sigma}_{ji}(M)}{\partial x_j} + \bar{b}_i(M) = 0, \quad M \in V, \quad (12.68)$$

и граничные условия

$$\sigma_{ij}(P) n_j(P) = p_i(P), \quad \bar{\sigma}_{ij}(P) n_j(P) = \bar{p}_i(P), \quad P \in S, \quad (12.69)$$

где n_{jk} — проекции единичного вектора $\mathbf{n}$ внешней нормали к поверхности S на координатные оси Ox_i , $i = \overline{1, 3}$. Кроме того, в соответствии с соотношениями Коши (3.30)

$$\varepsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \quad \bar{\varepsilon}_{ij} = \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \quad (12.70)$$

и в соответствии с обобщенным законом Гука (3.31)

$$\sigma_{ij} = \tilde{\lambda} \Theta \delta_{ij} + 2\tilde{\mu} \varepsilon_{ij}, \quad \bar{\sigma}_{ij} = \tilde{\lambda} \bar{\Theta} \delta_{ij} + 2\tilde{\mu} \bar{\varepsilon}_{ij}, \quad (12.71)$$

где $\Theta = \varepsilon_{ij} \delta_{ij}$ — объемная деформация; δ_{ij} — компоненты единичного тензора второго ранга, для которого $\delta_{ij} = 1$ при $i = j$ и $\delta_{ij} = 0$ при $i \neq j$.

Используя (12.71), можно установить, что с учетом $\delta_{ij} \delta_{ij} = 3$

$$\begin{aligned} \sigma_{ij} \bar{\varepsilon}_{ij} &= (\tilde{\lambda} \Theta \delta_{ij} + 2\tilde{\mu} \varepsilon_{ij}) \frac{1}{3} (\bar{\Theta} \delta_{ij} + 3\bar{\varepsilon}_{ij} - \bar{\Theta} \delta_{ij}) = \\ &= \tilde{\lambda} \Theta \bar{\Theta} + 2\tilde{\mu} \varepsilon_{ij} \bar{\varepsilon}_{ij} = \tilde{\lambda} \Theta \bar{\Theta} + (\bar{\sigma}_{ij} - \tilde{\lambda} \bar{\Theta}) \varepsilon_{ij} = \bar{\sigma}_{ij} \varepsilon_{ij}. \end{aligned}$$

Отсюда интегрированием по области V получаем

$$\int_V \sigma_{ij} \bar{\epsilon}_{ij} dV = \int_V \bar{\sigma}_{ij} \epsilon_{ij} dV. \quad (12.72)$$

Учитывая (12.68)–(12.70) и симметрию тензора напряжений, преобразуем интеграл в левой части (12.72):

$$\begin{aligned} \int_V \sigma_{ij} \bar{\epsilon}_{ij} dV &= \int_V \sigma_{ij} \frac{1}{2} \left(\frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) dV = \int_V \sigma_{ij} \frac{\partial \bar{u}_i}{\partial x_j} dV = \\ &= \int_S \sigma_{ij} n_j \bar{u}_i dS - \int_V \frac{\partial \sigma_{ij}}{\partial x_j} \bar{u}_i dV = \int_S p_i \bar{u}_i dS + \int_V b_i \bar{u}_i dV. \end{aligned}$$

Аналогично можно преобразовать интеграл в правой части (12.72), в результате вместо (12.72) получим итоговое соотношение

$$\int_S p_i \bar{u}_i dS + \int_V b_i \bar{u}_i dV = \int_S \bar{p}_i u_i dS + \int_V \bar{b}_i u_i dV \quad (12.73)$$

упомянутой теоремы: работа первой системы сил на перемещениях, вызванных второй системой сил, равна работе второй системы на перемещениях, вызванных первой системой.

Чтобы избавиться в правой части (12.73) от интеграла по области V и перейти к ГИУ, необходимо располагать решением задачи для перемещений в неограниченной линейно упругой среде, вызванных приложением в точке $M_0 \in \mathbb{R}^3$ сосредоточенной силы $\mathbf{f}$. Для этого рассмотрим уравнение (3.40) теории упругости в перемещениях, положив в нем вектор плотности объемных сил равным $\mathbf{b}(M) = \mathbf{f}(M_0) \delta_3(M, M_0)$, где $\delta_3(M, M_0)$ — дельта-функция, обладающая в случае функции $f(M)$, $M \in \bar{V} = V \cup S$, непрерывной на замыкании $\bar{V}$ области V , свойством

$$\int_V f(M) \delta_3(M, M_0) dV(M) = \begin{cases} \frac{\omega(M_0) f(M_0)}{4\pi}, & M_0 \in \bar{V}; \\ 0, & M_0 \notin \bar{V}, \end{cases}$$

где $\omega(M_0)$ — телесный угол, под которым из точки $M_0 \in \bar{V}$ видна „внутренность“ области V (в частности, $\omega(M_0) = 4\pi$ при $M_0 \in V$ и $\omega(M_0) = 2\pi$, если точка M_0 лежит на гладком участке поверхности S). Тогда (3.40) примет вид

$$\nabla^2 \mathbf{u} + \frac{1}{1-2\nu} \nabla(\nabla \mathbf{u}) = -\frac{1}{\tilde{\mu}} f \delta_3, \quad (12.74)$$

где $\nu = \frac{\tilde{\lambda}}{2(\tilde{\lambda} + \tilde{\mu})}$ — коэффициент Пуассона упругой среды.

Решение (12.74) будем искать в виде $\mathbf{u} = \mathbf{u}^* + \tilde{\mathbf{u}}$, причем $\mathbf{u}^*$ удовлетворяет векторному уравнению Пуассона

$$\nabla^2 \mathbf{u}^* = -\frac{1}{\tilde{\mu}} f \delta_3, \quad (12.75)$$

или в проекциях на координатные оси Ox_i , $i = \overline{1, 3}$,

$$\nabla^2 u_i^* = -\frac{1}{\tilde{\mu}} f_i \delta_3. \quad (12.76)$$

Известно [XII], что $u_i = \frac{f_i}{4\pi\tilde{\mu}r}$, где r — расстояние между точками $M, M_0 \in \mathbb{R}^3$, удовлетворяет (12.76). Тогда

$$\mathbf{u}^* = \frac{\mathbf{f}}{4\pi\tilde{\mu}r} \quad (12.77)$$

будет удовлетворять (12.75).

Подставляя теперь $\mathbf{u} = \mathbf{u}^* + \tilde{\mathbf{u}}$ в (12.74), получаем уравнение

$$(1-2\nu)\nabla^2 \tilde{\mathbf{u}} + \nabla(\nabla \tilde{\mathbf{u}}) = -\nabla(\nabla \mathbf{u}^*) \quad (12.78)$$

относительно искомой векторной функции $\tilde{\mathbf{u}}$. Применяя к нему операцию ротора и учитывая, что $\nabla \times (\nabla \varphi) = \mathbf{0}$ для любой дважды непрерывно дифференцируемой действительной функции φ , получаем $\nabla^2(\nabla \times \tilde{\mathbf{u}}) = \mathbf{0}$, где $\mathbf{0}$ — нулевой вектор. Таким образом, проекции $g_i = (\nabla \times \tilde{\mathbf{u}})_i$ векторной функции $\mathbf{g} = \nabla \times \tilde{\mathbf{u}}$

на координатные оси Ox_i , $i = \overline{1, 3}$, с ортами e_i удовлетворяют уравнениям Лапласа $\nabla^2 g_i = 0$, т.е. являются гармоническими функциями.

Особенность рассматриваемой задачи состоит в том, что по мере удаления от точки $M_0 \in \mathbb{R}^3$ перемещения, деформации и напряжения, вызванные приложением в этой точке силы f , стремятся к нулю, в том числе и $g_i \rightarrow 0$, $i = \overline{1, 3}$. Но функция g_i , являясь гармонической в пространстве $\mathbb{R}^3$ и обращаясь в нуль на бесконечности, равна тождественно нулю [XII]. Таким образом, $g = \nabla \times \tilde{u} = 0$ и поэтому существует скалярный потенциал Φ искомого векторного поля, задаваемого функцией $\tilde{u}$, т.е. $\tilde{u} = \nabla \Phi$. Тогда из (12.78) получим

$$(1 - 2\nu)\nabla^2 \nabla \Phi + \nabla(\nabla(\nabla \Phi)) = -\nabla(\nabla u^*),$$

или

$$\nabla(2(1 - \nu)\nabla^2 \Phi + \nabla u^*) = \nabla h = 0.$$

Отсюда следует, что функция $h \equiv \text{const}$. Но на бесконечности она должна обращаться в нуль. Поэтому с учетом (12.77) имеем

$$\nabla^2 \Phi = -\frac{1}{2(1 - \nu)} \nabla u^* = -\frac{f}{8\pi\tilde{\mu}(1 - \nu)} \nabla\left(\frac{1}{r}\right). \quad (12.79)$$

Правая часть (12.79) определена всюду в $\mathbb{R}^3$, кроме точки M_0 приложения силы f .

Пусть функция ψ удовлетворяет уравнению $\nabla^2 \psi = 1/r$. Тогда вместо (12.79) запишем

$$\nabla^2 \Phi = -\frac{f}{8\pi\tilde{\mu}(1 - \nu)} \nabla(\nabla^2 \psi) = -\frac{f}{8\pi\tilde{\mu}(1 - \nu)} \nabla^2(\nabla \psi).$$

Этому равенству удовлетворяет функция

$$\Phi = -\frac{f}{8\pi\tilde{\mu}(1 - \nu)} \nabla \psi. \quad (12.80)$$

Убедимся, что функция $\psi = r/2$ является решением уравнения $\nabla^2 \psi = 1/r$. Для этого поместим точку M_0 в начало сферических координат, в которых это уравнение примет вид [XII]

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\psi}{dr} \right) = \frac{1}{r}.$$

Последовательным дифференцированием находим

$$\frac{d\psi}{dr} = \frac{1}{2}, \quad \frac{d}{dr} \left(r^2 \frac{d\psi}{dr} \right) = \frac{d}{dr} \left(r^2 \frac{1}{2} \right) = r,$$

так что действительно $\nabla^2 \left(\frac{r}{2} \right) = \frac{1}{r}$. Подставляя $\psi = \frac{r}{2}$ в (12.80), получаем

$$\tilde{u} = \nabla \Phi = - \frac{1}{16\pi\tilde{\mu}(1-\nu)} \nabla(f \nabla r).$$

В итоге, учитывая (12.77), имеем

$$u = u^* + \tilde{u} = \frac{f}{4\pi\tilde{\mu}r} - \frac{1}{16\pi\tilde{\mu}(1-\nu)} \nabla(f \nabla r). \quad (12.81)$$

Это решение получено в 1848 году У. Томсоном*. Оно является *фундаментальным решением* однородного уравнения (3.40) теории упругости в перемещениях (при $b \equiv 0$).

Для перехода от (12.72) к ГИУ удобно представить (12.81) в проекциях $u_i(M) = u_i^{(l)}(M, M_0) f_l(M_0)$, $i, l = \overline{1, 3}$. Тогда значение функции $u_i^{(l)}(M, M_0)$ будет указывать перемещение в точке $M \in \mathbb{R}^3$ в направлении оси Ox_i , вызванное приложением в точке M_0 единичной силы в направлении оси Ox_l . Положим $f = f_l e_l$ и, учитывая, что $e_i e_l = \delta_{il}$, запишем

$$\nabla(f \nabla r) = e_i \frac{\partial}{\partial x_i} \left(f_l e_l \frac{\partial r}{\partial x_j} e_j \right) = f_l \delta_{il} \frac{\partial^2 r}{\partial x_i \partial x_j} e_j = f_l \frac{\partial^2 r}{\partial x_i \partial x_l} e_i.$$

*У Томсон (лорд Кельвин) (1824–1907) — английский физик и математик.

В итоге вместо (12.81) получим

$$u_i e_i = \frac{f_l e_l}{4\pi\tilde{\mu}r} - \frac{f_l e_i}{16\pi\tilde{\mu}(1-\nu)} \frac{\partial^2 r}{\partial x_i \partial x_l} = \frac{e_i}{4\pi\tilde{\mu}} \left(\frac{\delta_{il}}{r} - \frac{1}{4(1-\nu)} \frac{\partial^2 r}{\partial x_i \partial x_l} \right) f_l.$$

Отсюда находим искомую функцию

$$u_i^{(l)}(M, M_0) = \frac{1}{4\pi\tilde{\mu}} \left(\frac{\delta_{il}}{r(M, M_0)} - \frac{1}{4(1-\nu)} \frac{\partial^2 r(M, M_0)}{\partial x_i \partial x_l} \right). \quad (12.82)$$

При помощи (12.82), используя соотношения Коши (12.70) и обобщенного закона Гука (12.71), можно найти деформации и напряжения в точке $M \in \mathbb{R}^3$, вызванные приложением в точке M_0 единичной силы, направленной вдоль оси Ox_l :

$$\varepsilon_{ij}^{(l)}(M, M_0) = \frac{1}{2} \left(\frac{\partial u_i^{(l)}(M, M_0)}{\partial x_j} + \frac{\partial u_j^{(l)}(M, M_0)}{\partial x_i} \right), \quad i, j, l = \overline{1, 3};$$

$$\sigma_{ij}^{(l)}(M, M_0) = \tilde{\lambda} \varepsilon_{kk}^{(l)}(M, M_0) \delta_{ij} + 2\tilde{\mu} \varepsilon_{ij}^{(l)}(M, M_0), \quad i, j, l = \overline{1, 3},$$

а затем вычислить проекции $p_i^{(l)}(M, M_0) = \sigma_{ij}^{(l)}(M, M_0) n_j(M)$ вектора напряжений на площадку, проходящую через точку M и имеющую направляющие косинусы $n_j(M)$ единичного вектора $\mathbf{n}(M)$ нормали к этой площадке.

Подставляя в (12.73) $u_i^{(l)}(M, M_0)$ вместо $\bar{u}_i$, а $p_i^{(l)}(M, M_0)$ вместо $\bar{p}_i$ и полагая $\bar{b}_i = \delta_3 \delta_{il}$, с учетом указанного выше свойства дельта-функции получаем ГИУ

$$\begin{aligned} \frac{\omega(M_0)}{4\pi} u_l(M_0) = & \int_S (p_i(P) u_i^{(l)}(P, M_0) - u_i(P) p_i^{(l)}(P, M_0)) dS(P) + \\ & + \int_V b_i(M) u_i^{(l)}(M, M_0) dV(M), \quad M \in V, P \in S, M_0 \in \bar{V}. \end{aligned} \quad (12.83)$$

Это ГИУ можно решить при помощи МГЭ. Для этого кусочно гладкую поверхность S разобьем на N_S граничных элементов (ГЭ) S_n , $n = \overline{1, N_S}$, причем угловые точки и ребра, если они есть на поверхности, расположим на стыке элементов. Тогда в пределах каждого ГЭ будет определен единичный вектор внешней нормали к поверхности.

В простейшем случае в каждом ГЭ S_n функции $u_i(P)$ и $p_i(P)$ аппроксимируем значениями $u_i(P_n)$ и $p_i(P_n)$ соответственно в точке $P_n \in S_n$, помещенной в середине ГЭ, т.е. приближенно заменим эти функции их узловыми значениями в n -м узле. Совместим точку $M_0 \in \Gamma$ с m -м узлом, приняв $M_0 = P_m \in S_m$. Для этого узла, лежащего на гладком участке границы, в (12.83) имеем $\omega(M_0) = \omega(P_m) = 2\pi$. Тогда вместо (12.83), используя свойство аддитивности определенного интеграла, получаем

$$\begin{aligned} \frac{1}{2} u_l(P_m) = & \sum_{n=1}^{N_S} \int_{S_n} (p_i(P_n) u_i^{(l)}(P, P_m) - u_i(P_n) p_i^{(l)}(P, P_m)) dS(P) + \\ & + \int_V b_i(M) u_i^{(l)}(M, P_m) dV(M), \quad i, l = \overline{1, 3}, \end{aligned}$$

или в матричной записи, учитывая, что точку M_0 можно совместить с любым m -м узлом,

$$\bar{H}\tilde{U} = \bar{G}\tilde{Q} + D, \quad (12.84)$$

где $\tilde{U}$, $\tilde{Q}$ и D — матрицы-столбцы размера $3N_S \times 1$ узловых значений $u_{3(n-1)+1} = u_1(P_n)$, $p_{3(n-1)+1} = p_1(P_n)$, $n = \overline{1, N_S}$, и

$$d_{3(m-1)+l} = \int_F b_i(M) u_i^{(l)}(M, P_m) dV(M), \quad m = \overline{1, N_S}, \quad (12.85)$$

соответственно, $\bar{H}$ и $\bar{G}$ — матрицы порядка $3N_S$ с элементами

$$h_{3(m-1)+l, 3(n-1)+1} = \frac{\delta_{mn}}{2} + \int_{S_n} p_i^{(l)}(P, P_m) dS(P), \quad m, n = \overline{1, N_S}$$

$$g_{3(m-1)+l, 3(n-1)+1} = \int_{S_n} u_i^{(l)}(P, P_m) dS(P), \quad m, n = \overline{1, N_S},$$

$\delta_{mn} = 1$ при $m = n$ и $\delta_{mn} = 0$ при $m \neq n$.

Подынтегральные функции в интегралах по $\Gamma \Delta S_n$, $n = \overline{1, N_S}$, входящих в выражения для элементов матриц $\bar{H}$ и $\bar{G}$, при $n \neq m$ ограничены. Поэтому такие интегралы можно вычислить при помощи обычных *квадратурных формул* [VII]. При $n = m$ несобственные интегралы в выражениях для элементов матрицы $\bar{G}$ можно вычислить аналитически, если $\Gamma \Delta S_m$ является плоским*. В случае криволинейного $\Gamma \Delta$ в окрестности точки $P_m \in S_m$ целесообразно $\Gamma \Delta$ аппроксимировать плоским участком S'_m , вычислить аналитически вклад в несобственный интеграл на этом участке, а на остальной части $\Gamma \Delta$, где отсутствуют особенности в подынтегральных функциях, использовать для вычислений обычные квадратурные формулы.

Выражения для элементов матрицы $\bar{H}$ при $n = m$ также содержат несобственные интегралы. Можно избежать их вычисления, если учесть, что матричное уравнение (12.84) справедливо и для частного случая перемещения области V как твердого тела при отсутствии поверхностных и объемных сил и заданных значениях $u_i(P_n) = u_i^0 = \text{const}$. В этом случае матрицы $\tilde{Q}$ и D будут нулевыми и вместо (12.84) получим $\bar{H}\tilde{U} = O_{3N_S}$, где O_{3N_S} — нулевая матрица-столбец размера $3N_S \times 1$. Теперь заданные значения $u_{3(n-1)+1} = u_i(P_n) = u_i^0$, $n = \overline{1, N_S}$, будут удовлетворять матричному уравнению $\bar{H}\tilde{U} = O_{3N_S}$ лишь при

*См., например: Бенерджи П., Баттерфилд Р.

условии

$$\sum_{n=1}^{N_S} h_{3(m-1)+l, 3(n-1)+1} = 0, \quad i, l = \overline{1, 3}. \quad (12.86)$$

Отсюда можно вычислить значения $h_{3(m-1)+l, 3(n-1)+1}$, не прибегая к интегрированию по $\Gamma \in S_m$.

Интеграл в (12.85) также является несобственным. Можно показать, что он сходится в окрестности точки $P_m \in \Gamma_n$, если функция $b_i(M)$ ограничена в этой окрестности. Его можно вычислить, например, при помощи соответствующей *квадратурной формулы Гаусса*, не использующей значения подынтегральной функции в граничных точках.

Если на границе S области V заданы значения перемещений $u_i(P) = f_i(P)$, $P \in S$, $i = \overline{1, 3}$, то в (12.84) будут известны все элементы матрицы $\tilde{U}$. Тогда (12.84) переходит в систему линейных алгебраических уравнений (СЛАУ) относительно значений $p_i(P_n)$ в точках $P_n \in S_n$. После решения этой СЛАУ, используя (12.83) при $\omega(M_0) = 4\pi$, в любой точке $M_0 \in V$ можно вычислить значения

$$\begin{aligned} u_l(M_0) = & \sum_{n=1}^{N_S} p_i(P_n) \int_{S_n} u_i^{(l)}(P, M_0) dS(P) - \\ & - \sum_{n=1}^{N_S} u_i(P_n) \int_{S_n} p_i^{(l)}(P, M_0) dS(P) + \\ & + \int_V b_i(M) u_i^{(l)}(M, M_0) dV(M), \quad M \in V, \quad P \in S, \quad M_0 \in \bar{V}. \quad (12.87) \end{aligned}$$

Приняв, что $u_i^{(l)}(P, M_0) \approx u_i^{(l)}(P_n, M_0)$ и $p_i^{(l)}(P, M_0) \approx p_i^{(l)}(P_n, M_0)$, $P_n \in S_n$, $n = \overline{1, N_S}$, в пределах каждого $\Gamma \in S$ с площадью S_n , вме-

сто (12.87) получим менее точную, но более простую формулу

$$u_i(M_0) = \sum_{n=1}^{N_S} (p_i(P_n) u_i^{(l)}(P_n, M_0) - u_i(P_n) p_i^{(l)}(P_n, M_0)) S_n + \int_V b_i(M) u_i^{(l)}(M, M_0) dV(M). \quad (12.88)$$

Можно показать, что несобственный интеграл по области V в (12.87) и (12.88) сходится равномерно в окрестности точки $M_0 \in V$, если в этой окрестности функция $b_i(M)$ ограничена. Поэтому его можно дифференцировать по координатам точки M_0 как по параметрам. Это позволяет вычислить в любой точке $M_0 \in V$ компоненты тензоров деформаций и напряжений.

В более общем случае на участках $S_u \subset S$ поверхности S могут быть заданы проекции $u_i(P)$ вектора u перемещения, а на участках $S_p = S \setminus S_u$ — проекции $p_i(P)$ вектора напряжения. В этом случае разбиение поверхности S на N_S ГЭ целесообразно провести так, чтобы каждый ГЭ целиком находился либо на участке S_u , либо на участке S_p . Тогда (12.84) можно привести к квадратной СЛАУ вида (12.29) с $3N_S$ неизвестными значениями $u_i(P_n)$, $P_n \in S_p$, и $p_i(P_n)$, $P_n \in S_u$, $i = \overline{1, 3}$, $n = \overline{1, N_S}$. Возможен и более сложный вариант задания граничных условий, когда на некоторых участках $S_{up} = S \setminus (S_u \cup S_p)$ известны линейные комбинации указанных проекций. Тем не менее с учетом этих линейных комбинаций общее число неизвестных граничных значений составит $3N_S$, причем формирование элементов матриц A и B° в СЛАУ вида (12.29) аналогично рассмотренному выше (см. 12.2).

Если область V состоит из нескольких подобластей, в каждой из которых упругие свойства однородной среды различны, то матричные уравнения вида (12.84), составленные для каждой из однородных подобластей, можно объединить в одну СЛАУ при помощи условий, заданных на поверхностях контакта этих подобластей (см. 12.3). Для учета анизотропии

упругих свойств среды и ее неоднородности в области V применимы те же приемы, что и при использовании метода конечных элементов для решения задач теплопроводности (см. 12.3).

12.6. Сравнение методов граничных и конечных элементов

Метод граничных элементов (МГЭ) основан на использовании *фундаментальных решений* линейных дифференциальных уравнений с постоянными коэффициентами и поэтому в своем исходном виде применим к решению линейных задач в областях, содержащих однородные среды. Если область состоит из нескольких смежных подобластей, содержащих однородные, но различные среды, то в рамках МГЭ несложно состыковать между собой решения для отдельных подобластей. Однако учет неоднородности среды и (или) ее анизотропии требует, как правило, применения процедуры последовательных приближений (см. 12.3). Аналогичная ситуация возникает и при использовании МГЭ для решения нелинейных задач.

В *методе конечных элементов* (МКЭ) учет неоднородности и анизотропии среды не вызывает принципиальных трудностей. Однако в случае нелинейных задач при матричном представлении уравнений относительно *узловых значений* искомых функций элементы матриц зависят от этих значений, что также приводит к итерациям.

Ясно, что формулировка исходных задач в виде *граничных интегральных уравнений* (ГИУ) приводит к уменьшению размерности на единицу, т.е. для двумерных задач имеем одномерные ГИУ вдоль контура области, а для трехмерных — двумерные ГИУ по поверхности, ограничивающей область. Отметим, что некоторые осесимметричные задачи удастся свести к одномерным ГИУ вдоль контура осевого сечения рассматриваемой области. Поэтому последующее применение МГЭ для решения ГИУ позволяет обычно решать системы линейных

алгебраических уравнений (СЛАУ) с матрицами меньшего порядка, чем при использовании МКЭ для решения той же задачи. Однако формируемые на основе МГЭ матрицы не содержат нулевых элементов (в случае разбиения области на однородные подобласти матрица является блочно-ленточной, причем отдельные блоки состоят только из ненулевых элементов).

В противоположность этому матрица, сформированная на основе МКЭ, имеет большое число нулевых элементов. Оптимизируя нумерацию узлов КЭ, такую матрицу можно свести к ленточной с минимально возможной шириной ленты. Это позволяет уменьшить затраты памяти ЭВМ и число арифметических операций при решении СЛАУ с ленточной матрицей. Кроме того, вычисление каждого элемента такой матрицы требует меньшего числа операций, чем для элемента матрицы при ее формировании на основе МГЭ, хотя при отсутствии в ГИУ интеграла по области объем вычислений также уменьшается.

Оценки показывают, что при решении МГЭ трехмерных задач затраты машинного времени в четыре — десять раз меньше, чем при их решении МКЭ с той же точностью*. Выигрыш растет при решении задач, благоприятных для МГЭ. Например, в случае неограниченной области подбором фундаментального решения можно автоматически удовлетворить условиям на бесконечности, тогда как в МКЭ эти условия придется удовлетворять приближенно введением большого числа удаленных КЭ. При наличии у области участка границы с заданным нулевым значением производной искомой функции в направлении внешней нормали (например, идеально теплоизолированный участок в задаче теплопроводности или свободная поверхность в задаче движения жидкости) также можно подобрать фундаментальное решение, которое автоматически удовлетворит такое граничное условие, что позволит не разбивать этот участок на граничные элементы (ГЭ).

*См., например: Бенерджи П., Баттерфилд Р.

В МГЭ погрешности связаны лишь с аппроксимацией границы при помощи ГЭ и приближенном представлении на границе искомых и заданных функций, тогда как ГИУ является точной формулировкой искомого решения во всей области. В МКЭ же возникают дополнительные погрешности при представлении искомого решения во всей области. Если при решении задачи представляет интерес не все поле в области, а лишь его фрагменты или же единственное значение искомой функции в какой-либо характерной точке области, то МГЭ позволяет получить эту информацию с более высокой точностью и при меньшем объеме вычислений.

Проведенное сравнение показывает, что ни МКЭ, ни МГЭ не имеют абсолютного преимущества при решении широкого класса задач, но существуют типы задач, для которых применение одного из этих методов предпочтительнее. В некоторых случаях удобным может оказаться одновременное применение этих методов, когда в части области вводят *сетку* КЭ, а решение в остальной части представляют в виде ГИУ с последующим применением МГЭ. Возможно комбинировать МГЭ и с другими численными методами*.

Дополнение 12.1. Особенности решения осесимметричных задач

Рассмотрим область V , полученную вращением вокруг оси Ox_3 плоской области F (не обязательно односвязной), ограниченной кусочно гладкой границей Γ , образующей при вращении кусочно гладкую поверхность S (рис. 12.11). Введем цилиндрическую систему координат ρ, φ, z , начало которой совпадает с началом декартовой прямоугольной системы координат $Ox_1x_2x_3$, а ось Oz направлена по оси вращения Ox_3 .

В *граничном интегральном уравнении* (ГИУ) (12.19) для функции $u(M)$, удовлетворяющей в области V уравнению Пуас-

*См., например: Бреббия К., Теллес Ж., Вроубелл Л.

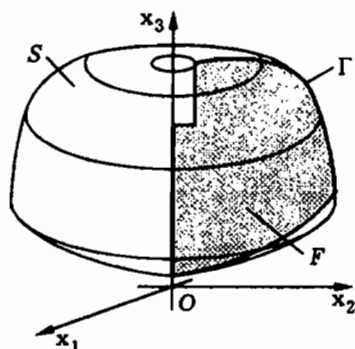


Рис. 12.11

сона (12.18), перейдем к цилиндрическим координатам, положив

$$dS(P) = \rho(P) d\varphi(P) d\Gamma(P),$$

$$dV(M) = \rho(M) d\varphi(M) d\rho(M) dz(M) = \rho(M) d\varphi(M) dF(M).$$

Тогда вместо (12.17) получим

$$\begin{aligned} \omega(M_0) u(M_0) = & \int_{\Gamma} \int_0^{2\pi} (q(P) w(P, M_0) - u(P) w^*(P, M_0)) \rho(P) d\varphi(P) d\Gamma(P) + \\ & + \int_F \int_0^{2\pi} f(M) w(M, M_0) \rho(M) d\varphi(M) dF(M), \quad (12.89) \end{aligned}$$

$$M \in V, M_0 \in \bar{V}, P \in S.$$

Пусть в осесимметричной области V требуется найти решение уравнения Пуассона (12.18) с граничными условиями (12.20), причем функции $f(M)$ в (12.18) и $f_1(P)$, $f_2(P)$, $\beta(P)$ не зависят от угловой координаты φ . При этом (12.18) можно

представить в виде

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial u(M)}{\partial \rho} \right) + \frac{\partial^2 u(M)}{\partial z^2} + f(M) = 0, \quad M \in F, \quad (12.90)$$

а (12.20) заменить на (12.15). Ясно, что в этом случае искомая функция $u(M)$ и ее производные по ρ и по z также не будут зависеть от φ . Тогда из (12.89) следует ГИУ

$$\begin{aligned} \omega(M_0) u(M_0) = & \int_{\Gamma} (q(P) \zeta(P, M_0) - u(P) \zeta^*(P, M_0)) \rho(P) d\Gamma(P) + \\ & + \int_F f(M) \zeta(M, M_0) \rho(M) dF(M), \quad M \in F, \quad M_0 \in \bar{F}, \quad P \in \Gamma, \end{aligned} \quad (12.91)$$

где

$$\zeta(M, M_0) = \int_0^{2\pi} w(M', M_0) d\varphi(M'), \quad M, M_0 \in \bar{F}, \quad M' \in \bar{V}; \quad (12.92)$$

$$\zeta^*(P, M_0) = \int_0^{2\pi} w^*(P', M_0) d\varphi(P'), \quad M_0 \in \bar{V}, \quad P \in \Gamma, \quad P' \in S. \quad (12.93)$$

Для вычисления интегралов в этих равенствах расстояние $r(M', M_0)$ между точками $M_0 \in \bar{F}$ и $M' \in \bar{V}$, входящее в *фундаментальное решение* $w(M', M_0) = \frac{1}{r(M', M_0)}$ уравнения Лапласа в пространстве, представим в цилиндрических координатах (рис. 12.12):

$$\begin{aligned} r^2(M', M_0) = & \rho^2(M) + \rho^2(M_0) - \\ & - 2\rho(M)\rho^2(M_0)\cos(\varphi(M') - \varphi(M_0)) + (z(M) - z(M_0))^2. \end{aligned} \quad (12.94)$$

Тогда, принимая в (12.94) $\varphi(M_0) = 0$ и учитывая, что

$$w^*(P', M_0) = \frac{\partial w(P', M_0)}{\partial n(P')} = -\frac{1}{r^2(P', M_0)} \frac{\partial r(P', M_0)}{\partial n(P')},$$

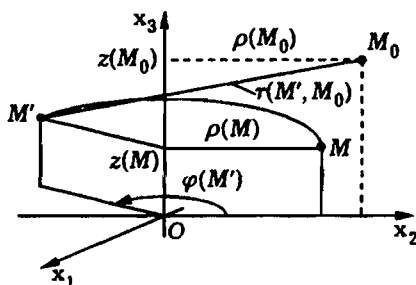


Рис. 12.12

где $n(P')$ обозначает направление единичного вектора $n(P)$ внешней нормали к поверхности S в точке $P' \in S$, из (12.92) получаем*

$$\zeta(M, M_0) = \int_0^{2\pi} \frac{d\varphi(M')}{r(M', M_0)} = \frac{4K(k)}{\mu(M, M_0)}, \quad M, M_0 \in \bar{F},$$

а из (12.93) находим

$$\begin{aligned} \zeta^*(P, M_0) &= \int_0^{2\pi} \frac{\partial r(P', M_0)}{\partial n(P')} \frac{d\varphi(P')}{r^2(P', M_0)} = \\ &= \frac{2}{\rho(P)\mu(P, M_0)} \left(K(k) + \frac{2 \frac{\rho^2(P)}{\mu(P, M_0)} + \frac{k^2}{2} - 1}{1 - k^2} E(k) \right) \frac{\partial \rho(P)}{\partial n(P)} + \\ &+ \frac{4}{\mu^3(P, M_0)} \frac{z(P) - z(M_0)}{1 - k^2} E(k) \frac{\partial z(P)}{\partial n(P)}, \quad P \in \Gamma, \quad M_0 \in \bar{F}, \end{aligned}$$

где $K(k)$ и $E(k)$ — полные эллиптические интегралы первого и второго рода соответственно

$$K(k) = \int_0^{\pi/2} \frac{d\varphi}{\sqrt{1 - k^2 \sin^2 \varphi}}, \quad E(k) = \int_0^{\pi/2} \sqrt{1 - k^2 \sin^2 \varphi} d\varphi$$

*См.: Зарубин В.С., 1985.

с модулем $k = k(M, M_0) = 4\rho(M)\rho(M_0)/\mu^2(M, M_0)$ и

$$\mu^2(M, M_0) = (\rho(M) + \rho(M_0))^2 + (z(M) - z(M_0))^2.$$

Разбиению поверхности S на N кольцевых *граничных элементов* (ГЭ) S_n , $n = \overline{1, N}$, соответствует разбиение контура Γ также на N участков, которые аналогичны ГЭ при решении *методом граничных элементов* (МГЭ) двумерной задачи. Поэтому эти ГЭ будем обозначать не S_n , а Γ_n . Если теперь искомую функцию и ее производную по направлению нормали аппроксимировать в пределах каждого ГЭ их постоянными значениями, т.е. положить $u(P) \approx u(P_n) = u_n$ и $q(P) = (\nabla u(P)) \mathbf{n}(P) \approx q(P_n) = q_n$, где $P_n \in \Gamma_n$ — узел в середине ГЭ Γ_n , то ГИУ (12.91) перейдет в матричное уравнение вида (12.23), в котором элементами матриц U и Q размера $N \times 1$ являются значения u_n и q_n , элементами матрицы B размера $N \times 1$ — значения

$$b_m = \int_F f(M) \zeta(M, P_m) \rho(M) dF(M), \quad m = \overline{1, N},$$

элементами квадратной матрицы H порядка N — значения

$$h_{mn} = \int_{\Gamma_n} \zeta^*(P, P_m) \rho(P) d\Gamma(P), \quad m \neq n,$$

и h_{mm} , вычисленные по формуле (12.26), а элементами матрицы G порядка N — значения

$$g_{mm} = \int_{\Gamma_n} \zeta(P, P_m) \rho(P) d\Gamma(P).$$

Интеграл в выражении для элемента g_{mm} является несобственным, так как при сближении точек $P, P_m \in \Gamma_m$ модуль $k \rightarrow 1$ и $K(k) \rightarrow \infty$. При

$$\nu(P, P_m) = \frac{(\rho(P) - \rho(P_m))^2 + (z(P) - z(P_m))^2}{2\rho(P)\rho(P_m)} \ll 1$$

можно приближенно принять*

$$\zeta(P, P_m) \approx -\frac{\ln \frac{\nu(P, P_m)}{32}}{\sqrt{\rho(P)\rho(P_m)}}.$$

Тогда для прямолинейного ГЭ Γ_m длиной l_m (т.е. для кольцевого ГЭ S_m с прямолинейной образующей) несобственный интеграл можно вычислить аналитически:

$$\begin{aligned} g_{mm} &= \int_{\Gamma_m} \zeta(P, P_m) \rho(P) d\Gamma(P) \approx \\ &\approx \kappa_m \left(\frac{2}{3} - \ln \frac{l_m}{256\rho_m} \right) ((\rho_m + b_m)^{3/2} - (\rho_m - b_m)^{3/2}) - \\ &- 2\kappa_m \rho_m^{3/2} \left(\ln \frac{(\rho_m + b_m)^{1/2} + \rho_m^{1/2}}{(\rho_m + b_m)^{1/2} - \rho_m^{1/2}} - \ln \left| \frac{((\rho_m - b_m)^{1/2} + \rho_m^{1/2})}{((\rho_m - b_m)^{1/2} - \rho_m^{1/2})} \right| \right) - \\ &- \kappa_m ((\rho_m + b_m)^{3/2} \ln(\rho_m + b_m) - (\rho_m - b_m)^{3/2} \ln(\rho_m - b_m)) + \\ &+ 4\kappa_m \rho_m ((\rho_m + b_m)^{1/2} - (\rho_m - b_m)^{1/2}), \quad (12.95) \end{aligned}$$

где

$$\kappa_m = \frac{2}{3\sqrt{\rho_m} \sin \alpha_m}, \quad \rho_m = \rho(P_m), \quad b_m = \frac{l_m}{2} \sin \alpha_m \geq 0,$$

α_m — угол между радиальным направлением и внешней нормалью $\mathbf{n}(P_m)$ в узле $P_m \in \Gamma_m$ (рис. 12.13). Если прямолинейный элемент ГЭ Γ_m параллелен оси Oz , то $\sin \alpha_m = 0$ и

$$g_{mm} \approx 2l_m \left(1 - \ln \frac{l_m}{16\rho_m} \right).$$

*См.: Зарубин В.С., 1985.

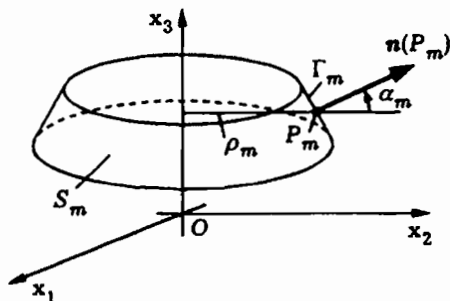


Рис. 12.13

Для ГЭ с малым значением ρ_m значение $\nu(P, P_m)$ может оказаться большим и поэтому вычисление несобственного интеграла в аналитической форме становится возможным лишь в малой окрестности точки $P_m \in \Gamma_m$ в пределах той части $\Gamma'_m \subset \Gamma_m$ длиной $l'_m \ll 2\sqrt{\frac{\rho(P)\rho_m}{50}}$, в середине которой остается точка $P_m \in \Gamma_m$. За пределами этого участка подынтегральная функция не содержит особенности и можно использовать обычные квадратурные формулы. Поэтому

$$g_{mm} = g'_{mm} + \int_{\Gamma_m \setminus \Gamma'_m} \zeta(P, P_m) \rho(P) d\Gamma(P),$$

где g'_{mm} находят, заменяя в (12.95) l_m на l'_m .

Аналогично можно найти значение g_{mm} для криволинейного ГЭ Γ_m , если в окрестности точки $P_m \in \Gamma_m$ представить ГЭ небольшим прямолинейным участком длиной l'_m .

При задании на контуре Γ граничных условий вида (12.15) и в случае осесимметричной задачи из (12.23) при помощи (12.30) и (12.31) можно перейти к системе линейных алгебраических уравнений (СЛАУ) (12.29). Решение этой СЛАУ дает недостающие узловые значения u_n и q_n , $n = \overline{1, N}$. После этого, используя (12.91) при $\omega(M_0) = 4\pi$, можно вычислить значение

$u(M_0)$ искомой функции u в любой точке $M_0 \in F$:

$$u(M_0) = \frac{1}{4\pi} \int_F f(M) \zeta(M, M_0) \rho(M) dF(M) + \\ + \frac{1}{4\pi} \sum_{n=1}^N \left(q_n \int_{\Gamma_n} \zeta(P, M_0) \rho(P) d\Gamma(P) - u_n \int_{\Gamma_n} \zeta^*(P, M_0) \rho(P) d\Gamma(P) \right).$$

Отметим, что при помощи МГЭ аналогичным образом можно решать нестационарные осесимметричные задачи и задачи теории упругости при осесимметричном напряженно-деформированном состоянии среды*.

Вопросы и задачи

12.1. Показать, что на прямолинейном в окрестности точки $M_0 \in \Gamma$ участке границы Γ первый интеграл в правой части (12.8) является собственным.

12.2. Доказать, что интеграл в выражении для диагонального элемента матрицы H в (12.23) равен нулю, если соответствующий граничный элемент является прямолинейным.

12.3. Проверить, что функция (12.44) является фундаментальным решением уравнения (12.43) при $I_V^{(q)}(M) \equiv 0$.

12.4. Убедиться, что (12.56) удовлетворяет однородному уравнению (12.54) (при $I_V^{(q)}(t, M) \equiv 0$) во всех точках $M \in \mathbb{R}^3$, кроме точки $M_0 \in \mathbb{R}^3$.

12.5. Проверить, что (12.62) является фундаментальным решением однородного уравнения (12.60) (при $f_k(M) \equiv 0$).

*См., например: Зарубин В.С., 1985.

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

Учебники и учебные пособия

Амосов А.А., Дубинский Ю.А., Копченкова Н.В. Вычислительные методы для инженеров. М.: Высш. шк., 1994. 544 с.

Батзалов Н.С., Жидков Н.П., Кобельков Г.М. Численные методы. М.: Наука, 1987. 600 с.

Беляев Н.М., Рядно А. Методы теории теплопроводности: В 2 т. Т. 1. М.: Высш. шк., 1982. 328 с.

Березанский Ю.М., Ус Г.Ф., Шефтель З.Г. Функциональный анализ. Курс лекций. Киев: Вища школа, 1990. 600 с.

Власова Е.А., Феоктистов В.В., Чадов В.Б. Введение в прикладной функциональный анализ. М.: Изд-во МГТУ им. Н.Э. Баумана, 1994. 54 с.

Воеводин В.В. Вычислительные основы линейной алгебры. М.: Наука, 1977. 304 с.

Гавурин М.К. Лекции по методам вычислений. М.: Наука, 1971. 248 с.

Ефимов А.В., Золотарев Ю.Г., Терпигорева В.М. Математический анализ (специальные разделы): В 2 т. Т. 2. М.: Высш. шк., 1980. 296 с.

Калиткин Н.Н. Численные методы. М.: Наука, 1978. 512 с.

Канторович Л.В., Акилов Г.П. Функциональный анализ. М.: Наука, 1984. 752 с.

Карташов Э.М. Аналитические методы в теплопроводности твердых тел. М.: Высш. шк., 1979. 416 с.

Коллатц Л. Функциональный анализ и вычислительная математика: Пер. с нем. М.: Мир, 1969. 448 с.

Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа. М.: Наука, 1989. 624 с.

Крылов В.И., Бобков В.В., Моастырный П.И. Вычислительные методы: В 2 т. Т. 1. М.: Наука, 1976. 304 с.

Кудрявцев Л.Д. Курс математического анализа: В 3 т. Т. 3. М.: Высш. шк., 1989. 352 с.

Люстерник Л.А., Соболев В.И. Краткий курс функционального анализа. М.: Высш. шк., 1982. 272 с.

- Марчук Г.И., Агошков В.И.* Введение в проекционно-сеточные методы. М.: Наука, 1981. 416 с.
- Мизлин С.Г.* Курс математической физики. М.: Наука, 1968. 576 с.
- Мизлин С.Г.* Вариационные методы математической физики. М.: Наука, 1970. 512 с.
- Пирумов У.Г., Росляков Г.С.* Численные методы газовой динамики. М.: Высш. шк., 1987. 232 с.
- Садовничий В.А.* Теория операторов. М.: Изд-во Моск. ун-та, 1986. 368 с.
- Самарский А.А.* Теория разностных схем. М.: Наука, 1977. 656 с.
- Самарский А.А., Гулин А.В.* Численные методы. М.: Наука, 1989. 432 с.
- Самарский А.А., Николаев Е.С.* Методы решения сеточных уравнений. М.: Наука, 1978. 592 с.
- Самарский А.А., Попов Ю.П.* Разностные схемы газовой динамики. М.: Наука, 1975. 352 с.
- Седов Л.И.* Механика сплошной среды: В 2 т. Т. 1. М.: Наука, 1970. 492 с.
- Соболев С.Л.* Некоторые применения функционального анализа в математической физике. М.: Наука, 1988. 334 с.
- Талдыкин А.Т.* Элементы прикладного функционального анализа. М.: Высш. шк., 1982. 384 с.
- Треногин В.А.* Функциональный анализ. М.: Наука, 1980. 496 с.
- Турчак Л.И.* Основы численных методов. М.: Наука, 1987. 320 с.
- Федоренко Р.П.* Введение в вычислительную физику. М.: Изд-во МФТИ, 1994. 528 с.
- Феокистов В.В.* Бесконечномерные пространства и применение функционального анализа в математической физике. М.: Изд-во МГТУ им. Н.Э. Баумана, 1994. 54 с.
- Фиттенгольц Г.М.* Основы математического анализа: В 2 т. Т. 2. М.: Наука, 1968. 464 с.

Справочные издания и монографии

- Александрова Н.В.* Математические термины: Справочник. М.: Высш. шк., 1978. 190 с.
- Андерсон Д., Таннехилл Дж., Плетчер Р.* Вычислительная гидромеханика и теплообмен: Пер. с англ. В 2 т. М.: Мир, 1990. 728 с.

Белоцерковский О. М. Численное моделирование в механике сплошных сред. М.: Наука, 1984. 520 с.

Бенерджи П., Баттерфилд Р. Методы граничных элементов в прикладных науках: Пер. с англ. М.: Мир, 1984. 496 с.

Бреббия К., Теллес Ж., Врубелл Л. Методы граничных элементов: Пер. с англ. М.: Мир, 1987. 524 с.

Бронштейн И. Н., Семендяев К. А. Справочник по математике для инженеров и учащихся втузов. М.: Наука, 1986. 544 с.

Верлань А. Ф., Сизиков В. С. Интегральные уравнения: методы, алгоритмы, программы: Справочное пособие. Киев: Наук. думка, 1986. 546 с.

Воднев В. Т., Наумович А. Ф., Наумович Н. Ф. Математический словарь высшей школы / Под ред. Ю. С. Богданова. Минск: Вышэйш. шк. 1984. 528 с.

Галанин М. П., Попов Ю. П. Квазистационарные магнитные поля в неоднородных средах. Математическое моделирование. М.: Наука, 1995. 320 с.

Герасимович А. И., Рысюк Н. А. Математический анализ: Справочное пособие для студентов втузов и инженеров. В 2 т. Т. 1. Минск: Вышэйш. шк., 1989. 288 с.

Горинштейн А. М. Практика решения инженерных задач на ЭВМ. М.: Радио и связь, 1984. 232 с.

Градштейн И. С., Рыжик И. М. Таблицы интегралов, сумм, рядов и произведений. М.: Физматгиз, 1963. 1100 с.

Джордж А., Лю Дж. Численное решение больших разреженных систем уравнений: Пер. с англ. М.: Мир, 1985. 336 с.

Дороговцев А. Я. Математический анализ: Справочное пособие для преподавателей математики, инженерно-технических работников и студентов. Киев: Вища школа, 1985. 528 с.

Зарубин В. С. Инженерные методы решения задач теплопроводности. М.: Энергоатомиздат, 1983. 328 с.

Зарубин В. С. Прикладные задачи термостойкости элементов конструкций. М.: Машиностроение, 1985. 296 с.

Зарубин В. С., Селиванов В. В. Вариационные и численные методы механики сплошной среды. М.: Изд-во МГТУ им. Н. Э. Баумана, 1993. 360 с.

Зимо И. Е., Тропп Э. А. Асимптотические методы в задачах теории теплопроводности и термоупругости. Л.: Изд-во Ленингр. ун-та, 1978. 224 с.

Иванов В.В. Методы вычислений на ЭВМ: Справочное пособие. Киев: Наук. думка, 1986. 584 с.

Коул Дж. Методы возмущений в прикладной математике: Пер. с англ. М.: Мир, 1972. 276 с.

Крылов В.И., Скобля Н.С. Справочная книга по численному обращению преобразования Лапласа. Минск: Наука и техника, 1968. 296 с.

Кувыркин Г.Н. Термомеханика деформируемого твердого тела при высокоинтенсивном нагружении. М.: Изд-во МГТУ им. Н.Э. Баумана, 1993. 144 с.

Курбатов П.А., Аринчин С.А. Численный расчет электромагнитных полей. М.: Энергоатомиздат, 1984. 168 с.

Лыков А.В. Тепломассообмен: Справочник. М.: Энергия, 1972. 560 с.

Математический энциклопедический словарь / Гл. ред. *Ю.В. Проторов*. М.: Сов. энцикл., 1988. 848 с.

Мейз Дж. Теория и задачи механики сплошных сред: Пер. с англ. М.: Мир, 1974. 320 с.

Мизлим С.Г. Численная реализация вариационных методов. М.: Наука, 1966. 432 с.

Можен Ж. Механика электромагнитных сплошных сред: Пер. с англ. М.: Мир, 1991. 560 с.

Норри Д., де Фриз Ж. Введение в метод конечных элементов: Пер. с англ. М.: Мир, 1981. 304 с.

Писсанецки С. Технология разреженных матриц: Пер. с англ. М.: Мир, 1988. 412 с.

Ректорис К. Вариационные методы в математической физике и технике: Пер. с англ. М.: Мир, 1985. 590 с.

Решение краевых задач методом Монте-Карло / *Б.С. Еленов, А.А. Кромберг, Г.А. Михайлов, К.К. Сабельфельд*. Новосибирск: Наука, Сибирское отд., 1980. 176 с.

Риттмайер Р., Мортон К. Разностные методы решения краевых задач: Пер. с англ. М.: Мир, 1972. 420 с.

Роуч П. Вычислительная гидродинамика: Пер. с англ. М.: Мир, 1980. 616 с.

Рудин У. Функциональный анализ: Пер. с англ. М.: Мир, 1975. 448 с.

Сильвестер П., Феррари Р. Метод конечных элементов для радиоинженеров и инженеров-электриков: Пер. с англ. М.: Мир, 1986. 232 с.

Сиротин Ю.И., Шаскольская М.П. Основы кристаллофизики. М.: Наука, 1979. 640 с.

Справочное пособие по приближенным методам решения задач высшей математики / Л.И. Бородич, А.И. Герасимович, Н.П. Кеда, И.Н. Мелешко. Минск: Вышэйш. шк., 1986. 190 с

Стренг Г., Фикс Дж. Теория метода конечных элементов: Пер. с англ М.: Мир, 1977. 352 с.

Сьярле Ф. Метод конечных элементов для эллиптических задач: Пер. с англ. М.: Мир, 1980. 512 с.

Толмачев В.В., Головин А.М., Потапов В.С. Термодинамика и электродинамика сплошной среды. М.: Изд-во Моск ун-та, 1988. 232 с.

Трауб Дж. Итерационные методы решения уравнений: Пер. с англ М.: Мир, 1985. 264 с.

Уилкинсон, Райнш. Справочник алгоритмов на языке АЛГОЛ. Линейная алгебра: Пер. с англ. М.: Машиностроение, 1976. 390 с.

Форсайт Дж., Малькольм М., Моулер К. Машинные методы математических вычислений: Пер. с англ. М.: Мир, 1980. 280 с.

Хатсон В., Пим Дж. Приложения функционального анализа и теории операторов: Пер. с англ. М.: Мир, 1983. 432 с.

Шуп Т.Е. Прикладные численные методы в физике и технике: Пер. с англ. М.: Высш. шк., 1990. 256 с.

Задачники

Механика сплошных сред в задачах / Под ред. М.Э. Эглит. В 2 т. М.: Московский Лицей, 1996. Т. 1, 396 с.; Т. 2, 394 с.

Треногин В.А., Писаревский Б.М., Соболева Т.С. Задачи и упражнения по функциональному анализу. М.: Наука, 1984. 256 с.

Халмош П. Гильбертово пространство в задачах: Пер. с англ. М.: Мир, 1970. 352 с.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

Аксиомы скалярного умножения
IV, 209

Алгоритм неустойчивый 418

– устойчивый 418

Альтернатива Фредгольма 194

Аппроксимация гладкая 557

– производных 404

– функции VII, 538

Базис стандартный IV

– счетный IX, 196

Блок матрицы III

Вариация функционала первая XV

Вектор напряжения 74

– собственный квадратной
матрицы IV

Величина случайная XVI

Вероятность события XVI

Волны стоячие XII

Гиперплоскость XIV, 170

Девiator напряжений 92

– скоростей деформаций 92

Дельта-функция XII, 665

Дифференциал Гато XV, 598

Завихренность VII, 94

Задача Дирихле XII, 502

– на собственные значения 237

– Штурма — Лиувилля XI, 441

Закон Гука обобщенный 98

Запись СЛАУ матричная III

Значение регулярное оператора 186

– собственное IV, 186

– квадратной матрицы IV

– кратное 234

– операторного уравнения 234

– простое 234

– узловое 402

Изометрия 198

Интеграл Бернулли 87

– зависящий от параметра VI

– Коши — Лагранжа 87

– эллиптический второго рода VI,
679

– первого рода VI, 679

Интерполирование II

Интерполяция линейная
(двухточечная) II

Квадрат скалярный IV, 209

Класс эквивалентности XIX

Композиция операторов 146

Компонента тензора IV

Константы Ламе 98

Координаты точки

барицентрические 543

– элемента в базисе IX, 196

Коэффициент сжатия 159

Коэффициенты весовые VI

– теплопроводности главные 82

Кратность собственного значения
234

Лемма фундаментальная Коши 76

Матрица Грама IV

- жесткости глобальная 611
- -- конечного элемента 611
- квадратичной формы IV
- масс глобальная 614
- -- конечного элемента 615
- неразложимая 444
- ортогональная IV
- положительно определенная IV
- пятидиагональная 440
- разреженная 579
- с диагональным преобладанием III, 416
- -- -- -- частичным 416
- слабозаполненная 579
- теплоемкости глобальная 586
- -- конечного элемента 587
- теплопроводности глобальная 576
- -- конечного элемента 576
- трехдиагональная III, 415

Матрица-столбец III

Мера Лебега IX

Метод баланса 409

- Бубнова — Галеркина 351
- вариации постоянных VIII
- взвешенных невязок 326
- Гаусса III
- -- с выбором главного элемента III
- граничных элементов 635
- итераций I-100, 284
- Канторовича 386
- коллокации в подобластях 327
- -- -- -- точках 334
- конечных разностей 403
- -- -- -- элементов (МКЭ) 536
- локальных вариаций 583

Метод малого параметра 303

- моментов 326
- наименьших квадратов IV, 345
- ортогональных проекций 325
- последовательных приближений I-100, 284
- прогонки III, 416
- проекционный 389
- прямых 453
- разделения области 327
- Ритца 353
- Рунге II
- Рунге — Кутты VIII, 459
- сеток 402
- статистических испытаний 504
- стрельбы 448
- установления 501
- Фурье XII
- штрафа 383

Методы проекционно-сеточные 403

- сеточные 402

Метод Эйлера VIII, 459

Метрика I-177

Минор III

- главный IV, 446
- порядка k III
- угловой III

Миноры главные диагональные VIII, 446

Многообразие линейное IX

Многочлен интерполяционный II

- -- Лагранжа II, 404
- -- полный 539
- -- Эрмита кубический II, 557

Множество всюду плотное IX, 140

- измеримое IX
- объема нуль VII
- относительно компактное 153

Множество предкомпактное 153

- функций равномерно ограниченное 155
- равностепенно непрерывное 155

Моделирование математическое 19

Модель математическая II

Напряжение касательное 76

- нормальное 76

Невязка уравнения операторного 326

Неравенство Коши — Буняковского IV, 209

- Шварца 210

Норма IV, 139

- кубическая IV
- линейного оператора 172
- функционала 164
- спектральная IV, 455
- функции 139
- энергетическая 227

Область значений оператора 146

- определения оператора 146

Оболочка линейная IV, 138

Ожидание математическое XVI

Оператор 146

- бигармонический 95
- взаимно однозначный 146
- вполне непрерывный 153
- неотрицательный XV
- непрерывный 151
- в точке 151
- нулевой IV, 161
- обратный 146
- левый 149
- правый 149
- ограниченный 151
- положительно определенных 215

Оператор положительный 215

- проекционный 390
- разностный 493
- резольвентный 186
- самосопряженный IV, 268
- сжимающий 159
- симметрический 215
- сопряженный IV, 265
- тождественный IV, 146
- Штурма — Лиувилля XI, 278
- эллиптический 272

Оператор-функция 309

Орт III

Ортопроектор 394

Оси главные тензора 82

Отрезок разбиения частичный VI

Оценка погрешности

- апостериорная 294
- априорная 294

Параметр интегрального уравнения 185

- узловой 525

Плотность количества движения среды объемная 22

- момента количества движения среды объемная 22

- объемная 21

- потока 26

-- массы 26

-- энергии 26

- среды 21

- электрического заряда объемная 21

-- тока 26

- энергии объемная 21

Поверхность разрыва 47

-- сильного 66

-- слабого 66

- Погрешность аппроксимации 404
- ОДУ 421
 - квадратурной формулы VI, 289
 - разностной схемы 421
- Подмножество выпуклое XV, 32
- Подпоследовательность I-243
- Подпространство нормированного пространства IX, 140
- собственное операторного уравнения 235
- Покрывание множества открытое I-188
- Пополнение нормированного пространства 198
- Порядок аппроксимации 421
- погрешности аппроксимации 405
 - точности разностной схемы 422
- Последовательность
- минимизирующая XV, 225
 - сходящаяся поточечно 175
 - фундаментальная I-314, 140
- Потенциал векторного поля
- скалярный VII, 42
 - объемный XII, 125
 - скорости 42
 - слоя двойного XII, 125
 - простого XII, 125
- Предел последовательности, сходящейся по норме IX, 139
- Преобразование интегральное Лапласа XI
- линейное IV
 - ортогональное матрицы IV
- Принцип максимума 418
- равномерной ограниченности 177
- Прогонка встречная 477
- матричная 516
 - немонотонная 483
 - ортогональная 517
- Прогонка потоковая 478
- продольно-поперечная 494
 - циклическая 480
- Продолжение функционала 170
- Проектор 390
- Произведение скалярное энергетическое 227
- Пространства нормированные изометричные 198
- Пространство I-177, 137
- банахово IX
 - рефлексивное 264
 - гильбертово IX, 209
 - линейное бесконечномерное IV, 138
 - метрическое I-177
 - нормированное полное IX
 - сепарабельное IX, 140
 - сопряженное IV, 258
 - второе IV, 263
 - функций почти всюду ограниченных 261
 - суммируемых IX
 - с квадратом IX
 - p -й степенью 260
 - функциональное 137
 - нормированное 139
 - энергетическое 229
- Процесс ортогонализации Грама -- Шмидта IV
- Р**авенство Парсеваля IX
- Разложение матрицы мультипликативное IV
- Холецкого III
 - элемента по базису IX, 196
- Разность конечная II
- левая (назад) II, 405
 - правая (вперед) II, 405

Разность конечная центральная II,
405

Ранг тензора IV

Расширение оператора 146

Резольвента оператора 186

Рефлексивность XIX, 198

Решение классическое 197

– обобщенное 197

– операторного уравнения
приближенное XV, 527

– слабое 226

– тривиальное IV

– фундаментальное уравнения
Лапласа XII, 623

– – – теплопроводности XII, 657

Ряд мажорирующий IX

Свертка тензора IV

Сетка временная 401

– двумерная 402

– конечных элементов 525

– одномерная 402

– пространственная 401

– пространственно-временная 401

– трехмерная 402

Сила эйлерова VIII, 363

Симметричность XIX, 198

Система биортогональная 268

– биортонормированная 268

– замкнутая IX, 140

– координат сопутствующая 54

– минимальная 268

– ОДУ жесткая 458

– ортонормированная IX

– – – полная IX

– сопряженная 269

След матрицы IV

Слой сетки

пространственно-временной 401

Соотношения Коши 98

Спектр дискретный 245

– оператора IV, 187

– – непрерывный 187

– – остаточный 187

– – точечный 187

Среда сплошная 20

Степень оператора 159

Сужение оператора 146

Сумма квадратурная VI

Схема разностная 402

– – двухслойная неявная 464

– – – с весами 462

– – – симметричная 465

– – – явная 463

– – корректная 476

– – локально-одномерная 496

– – неявная трехслойная 467

– – типа предиктор-корректор 471

– – трехслойная симметричная 468

– – явная трехслойная 466

Сходимость в среднем

квадратичном IX, 211

– сильная 261

– слабая 262

Тензор IV

– второго ранга единичный 92

– деформаций 98

– – напряжений 76

– скоростей деформаций 92

– теплопроводности 81

– шаровой 86

Теорема Арцела 155

– Банаха об обратном операторе
179

– Банаха — Штейнгауза 177

– взаимности работ 663

– Гершгорина 443

Теорема Гильберта — Шмидта 252
 – о квадратичном функционале 220
 – Стокса VII, 122
 – Ф. Рисса 218
 – Хана — Банаха 171
 Тетраздр VII, 539
 Точка неподвижная оператора 160
 – стационарная фуикционала XV, 21
 – экстремума фуикционала XV
 Точки коллокации 334
 Транзитивность I-43, 198
 Триангуляция области VII, 543
 Тройка некомпланарных векторов
 правая III, 124
 – точек правая V, 540
 Трубка векторная VII

Узел квадратурной формулы VI

– конечного элемента 525
 – сетки 401
 – промежуточный 405
 Умножение операторов 182
 – скалярное IV
 – стандартное IV
 Уравнение бигармоническое 96
 – Гельмгольца 45
 – интегральное 125
 – граничное 622
 – II рода XI, 180
 – интегро-дифференциальное 125
 – неразрывности 39
 – операторное XV, 146
 – однородное 186
 – разностное 402
 – состояния 20
 – телеграфное XII, 110

Уравнение теории упругости в
 перемещениях 102

– функциональное 20

Уравнения движения 78

– Ламе 99

– Максвелла XII, 109

– Навье — Стокса 93

– равновесия 78

– Эйлера 86

Условие граиичное главное
 (предварительное) 382

– естественное 382

Условия граиичные связанные 433

– краевые 20

Устойчивость алгоритма
 вычислительная 418

– по входным данным 418

Форма уравнения неразрывности дивергентная 46

– обобщенная 50

Формула Гриаи вторая VII, 36

– интегральная XII, 125

– первая VII, 36

– квадратурная VI, 287

– Гаусса VI

– Остроградского 35

– Остроградского — Гаусса VII, 30

– парабол VI, 289

– трапеции VI

Формулировка задачи
 интегральная 377

Функции базисные 325

– финитные 536

– весовые 326

– проекционные 325

Функционал билинейный 219

Функционал выпуклый строго XV,
598

- квадратичный 219
- Лагранжа XV
- линейный 164
- ограниченный 164
- энергии 224

Функция аналитическая 305

- бигармоническая 96
- гармоническая XII, 87
- Дирихле I-107, 144
- измеримая IX
- интегрируемая по Лебегу IX, 144
- потенциальная VII, 27
- собственная оператора 186
- суммируемая IX, 144
- с квадратом и весом IX, 212
- тока 94
- формы конечного элемента 529

Число обусловленности матрицы
IV, 295

- характеристическое 186

Шаблон разностной схемы 463

Шаг сетки 404

Шар I-179

- замкнутый I-179

Эквивалентность XIX

Элемент граничный 635

- коичный (КЭ) 525
 - комплексный 538
 - лагранжевый 538
 - мультиплексный 538
 - симплексный 538
 - сингулярный 548
 - эрмитовый 538
 - ортогональный подпространству IX, 213
 - собственный оператора 186
 - операторного уравнения 234
- Элементы ортогональные IX, 213
- Эллипсоид трехосный III

Ядро интегрального уравнения 185

- оператора IV, 178