А. М. КОЛОБОВ, Г. С. НЕВЕРОВ

ИЗБРАННЫЕ ГЛАВЫ ВЫСШЕЙ МАТЕМАТИКИ

Издательство «Вышэйшая школа»

А. М. КОЛОБОВ, Г. С. НЕВЕРОВ

ИЗБРАННЫЕ ГЛАВЫ ВЫСШЕЙ МАТЕМАТИКИ

Часть 3

Методы математической физики (дифференциальные уравнения в частных производных второго порядка). Элементы математической логики

Издательство «Вышэйшая школа»
Минск 1971

517
К61
УДК 51(075.8)

$\frac{2-2-3}{6-71}$

ПРЕДИСЛОВИЕ

Учебное пособие по избранным главам высшей математики составлено для студентов радиотехнических, энергетических и машиностроительных специальностей втузов. Оно может быть использовано студентами младших курсов университетов и пединститутов, инженерами и аспирантами технических специальностей.

Книга состоит из двух разделов.

Первый раздел «Методы математической физики (дифференциальные уравнения в частных производных второго порядка)» посвящен введению в теорию дифференциальных уравнений в частных производных. Содержание его соответствует специальному курсу по уравнениям математической физики, рекомендованному программой по высшей математике МВ и ССО СССР для инженерно-технических специальностей высших учебных заведений.

Второй раздел «Элементы математической логики» содержит материал по специальному курсу высшей математики для радиотехнических и энергетических специальностей втузов.

Первый раздел написан А. М. Колобовым, второй — Г. С. Неверовым.

Пособие предназначено для лиц, самостоятельно изучающих курс высшей математики. В нем указаны пути по использованию изучаемых методов и понятий в радиотехнике, автоматике, вычислительной технике.

Третья часть «Избранных глав высшей математики» имеет самостоятельное значение, хотя в ней и затрагиваются вопросы, подробно обсуждаемые в первой и во второй частях, изданных в 1965 и 1967 гг. Это обстоятельство в большей степени касается первого раздела, в котором широко используются такие понятия, как ряд и интеграл Фурье, преобразование Лапласа, элементы теории поля и другие.

Второй раздел пособия посвящен изложению основ математической логики и ее применению в кибернетике. Вычислительная техника и дискретная математика, занявшие в настоящее время ведущее положение в науке, планировании и управлении, широчайшим образом используют понятия математической логики. Эти обстоятельства сделали необходимым введение в программу высшей школы элементов математической логики.

Авторы старались сохранить стиль и характер изложения таким же, который был избран при написании двух предыдущих частей пособия. Особое внимание уделяется не строгому доказательству теорем и отдельных положений, а скорее рассказу о том, как они доказаны, дается постановка задачи, указывается метод решения ее, подчеркиваются особенности этого метода и диапазон его использования. Во втором разделе пособия изложение материала и доказательства проведены не аксиоматическим методом, а с использованием так называемых таблиц истинности. Это позволяет сделать все изложение материала значительно проще и доступнее. В то же время сохраняется общность теории, применяемость ее в различных областях знаний.

С чувством глубокой признательности авторы выражают благодарность руководству Брестского инженерно-строительного института, чье внимательное и доброжелательное отношение к работе способствовало ее успешному завершению. Авторы благодарят также руководство Минского высшего инженерного зенитно-ракетного училища за внимание к этой работе. Затем мы хотим поблагодарить профессора Е. Н. Вавилова и доцента И. А. Кодачигова, а также руководимые ими коллективы, за ряд ценных замечаний, значительно улучшивших окончательную редакцию книги. Мы благодарны также доценту Б. К. Федюшину, прочитавшему с большим вниманием рукопись и сделавшему ряд ценных предложений. Наконец, мы считаем своим приятным долгом выразить искреннюю благодарность доценту П. А. Тревогину за ту неоценимую помощь, которую он оказал авторам в процессе работы над книгой.

Авторы

Раздел I

МЕТОДЫ МАТЕМАТИЧЕСКОЙ ФИЗИКИ (ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ ВТОРОГО ПОРЯДКА)

Этот раздел является заключительным разделом традиционного курса математики, посвященного методам исследования физических и технических систем, не связанных со случайными процессами. Своеобразие его заключается в том, что почти все понятия высшей математики находят здесь дальнейшее развитие и использование.

Материал раздела разбит на две главы.

В первой главе особое внимание уделено задачам, приводящим к простейшим уравнениям в частных производных. Дается математическая формулировка краевых задач, рассматривается важный вопрос о приведении уравнения второго порядка к каноническому виду.

Вторая глава содержит важнейшие методы интегрирования уравнений в частных производных второго порядка. Кроме традиционных методов: характеристик, разделения переменных и метода функции Грина, сюда включен метод интегральных преобразований, с которым читатель знаком по первой части пособия.

Раздел заканчивается изложением первоначальных понятий по приближенному интегрированию дифференциальных уравнений в частных производных методом конечных разностей (методом сеток).

Глава 1. ОСНОВНЫЕ УРАВНЕНИЯ МАТЕМАТИЧЕСКОЙ ФИЗИКИ. ПРИВЕДЕНИЕ К КАНОНИЧЕСКОМУ ВИДУ И КЛАССИФИКАЦИЯ

§ 1.1. ВВЕДЕНИЕ

Трудно переоценить то значение, которое имеют дифференциальные уравнения для каждого, кто сталкивается с изучением законов природы. Дело в том, что к решению таких уравнений может быть сведено исследование многих физических проблем и технических задач. Часто бывает и так, что сами физические законы, которым подчиняется то или иное явление, записываются в форме дифференциальных уравнений.

Способность математики описывать явления природы удивляла и не перестает удивлять исследователей всех времен и народов. Великий астроном XVII в. Кеплер, открывший законы движения планет, изумленно писал: «... осмеливаюсь думать, что вся природа и благословенное небо записаны на языке искусства геометрии». Немецкий физик Герц, подтвердивший опытом уравнения Максвелла, доказав существование радиоволн, отдавая дань идеалистическим течениям XIX в., говорил: «Невозможно избавиться от ощущения, что эти математические формулы существуют независимо от нас и обладают собственным разумом, что они мудрее нас, мудрее даже тех, кто их открыл, и что мы извлекаем из них больше, чем первоначально было заложено». Общность методов математики делает ее гибким и мощным орудием исследования законов природы.

В математике дифференциальные уравнения занимают особое место. В механике, астрономии, физике, технике с их помощью были решены многие важные задачи. Достаточно сказать, что Ньютон, исследуя дифференциальные уравнения движения небесных тел, получил законы движения планет, установленные Кеплером эмпирически, а Леверье и Адамс в 1846 г. предсказали существование планеты Нептун и определили ее положение на небе на основании численного анализа тех же уравнений.

Дифференциальным уравнением мы называем такое уравнение, в которое входит не только сама неизвестная функция, но и ее производные некоторых порядков. Основной задачей теории дифференциальных уравнений является изучение функций, удовлетворяющих заданному дифференциальному уравнению.

Может оказаться, что неизвестная функция, входящая в дифференциальное уравнение вместе со своими производными, зависит от одного аргумента. Такое уравнение называют обыкновенным дифференциальным уравнением. С элементами теории обыкновенных дифференциальных уравнений знакомятся обычно в общем курсе высшей математики. Если неизвестная функция зависит от нескольких аргументов и в уравнение входят производные от нее по нескольким аргументам, дифференциальное уравнение называют уравнением в частных производных. С элементами теории дифференциальных уравнений в частных производных мы познакомимся в этом пособии.

Из теории поля нам известно (см. ч. 2, стр. 12), что физические явления, физические процессы, происходящие во времени и в пространстве, всегда заключаются в изменении каких-либо физических величин, которые могут быть описаны при помощи функций нескольких переменных. Поэтому, изучая какое-либо явление природы, мы встречаемся с уравнениями в частных производных даже чаще, чем с обыкновенными дифференциальными уравнениями. К тому же проблемы техники и прикладной физики часто приводят к рассмотрению таких задач, которые являются задачами теории поля, т. е. приводят к необходимости рассматривать уравнения в частных производных.

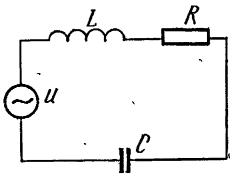

Рис. 1.1

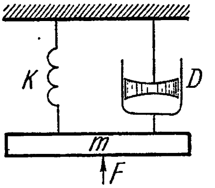

Рис. 1.2

Формально эти проблемы можно разделить на две большие категории. Проблемы, которые относятся к системам с сосредоточенными параметрами (см. ч. 1, стр. 197), и проблемы, которые относятся к системам с распределенными параметрами. В теории электрических цепей мы неоднократно встречались с примерами систем с сосредоточенными параметрами (см. рис. 1.1). Аналогичный пример можно привести и из механики (см. рис. 1.2). В электрической цепи индуктивность L и емкость C являются накопителями энергии, в то время как сопротивление R рассеивает ее. В механической же цепи накопителями энергии (потенциальной и кинетической) являются соответственно пружина K и масса m, а роль рассеивателя энергии играет поршень D.

Поведение как электрической, так и механической системы полностью определяется способом взаимодействия перечисленных выше элементов с источником напряжения U или с приложенной силой F. При составлении уравнения, описывающего работу такой системы, не принимается во внимание ни длина линий связи, ни порядок расположения элементов. Несущественным является не только расположение элементов, но и их размеры. Так, например, в случае механической системы принимается что весом пружины можно пренебречь. В действительности же под влиянием собственного веса верхняя часть пружины удлиняется несколько больше нижней части, что при составлении уравнения естественно не учитывается. Не учитывается также и то обстоятельство, что сила, приложенная к одному концу элемента, не может сразу же оказать влияние точно таким же образом на другой конец элемента и всюду внутри элемента. Другими словами, если учесть массу и внутреннее трение пружины, то сила, действующая на верхнюю часть пружины, будет отличаться от силы, приложенной к нижней ее части.

Такая идеализация при составлении уравнения, описывающего поведение механической системы, характерна для системы с сосредоточенными параметрами и при решении многих проблем динамики приводит к удовлетворительным результатам.

Если же исследователя интересует внутреннее поведение пружины, т. е. вопрос о том, насколько сжат или растянут каждый ее элемент, то такой подход к проблеме непригоден. Теперь уже приходится учитывать, что каждый участок пружины как бы он мал ни был, обладая свойствами идеальной пружины, обладает к тому же массой и внутренним трением. При таком изучении поведения пружины под действием силы все эти обстоятельства и свойства должны рассматриваться совместно. Так мы приходим к понятию о непрерывных или распределенных параметрах, когда все напряжения и деформации нужно связать с их распределением по длине пружины. Таким образом, задача для системы с сосредоточенными параметрами становится задачей для системы с распределенными параметрами, т. е. задачей теории поля.

Как мы уже указывали, отличительной чертой задач теории поля является то, что положение в пространстве характеризуется несколькими независимыми переменными. Другими словами, описывая поведение систем с распределенными параметрами при помощи дифференциального уравнения, мы приходим к уравнению в частных производных от искомой функции. В системах же с сосредоточенными параметрами независимой переменной является только время. Поэтому, описывая поведение такой системы уравнением, приходят к обыкновенному дифференциальному уравнению. Совершенно аналогично можно было бы рассмотреть переход от электрической системы с сосредоточенными параметрами к системам с распределенными параметрами. И вообще такие рассуждения применимы во многих других областях физики и техники. Необходимо помнить, что для полного описания и исследования таких явлений, как распространение тепла через теплопроводную среду, течение жидкости через пористую среду или по каналам, распространение звуковых и электромагнитных волн, гравитационное, электрическое и магнитное взаимодействие тел и других явлений требуется знание распределенных параметров и положения исследуемой точки в пространстве. Таким образом, все эти явления представляют собой задачи теории поля и могут быть сведены к исследованию дифференциальных уравнений в частных производных.

Между теорией обыкновенных дифференциальных уравнений и теорией дифференциальных уравнений в частных производных существует глубокое различие. Если первая, вообще говоря, достигла высокого уровня развития и представляет собой стройную и строгую систему, то вторая находится еще в процессе бурного развития и далека до завершения. Последнее скорее всего объясняется чрезвычайной трудностью многих проблем теории уравнений в частных производных. Даже при беглом знакомстве с простейшими методами интегрирования уравнений в частных производных бросается в глаза заметное повышение трудности проблемы по сравнению с обыкновенными дифференциальными уравнениями.

Оставляя в стороне вопросы общей теории дифференциальных уравнений в частных производных, в дальнейшем будем рассматривать лишь простейшие из таких уравнений, к которым приводят многочисленные задачи физики и техники. Оказывается, рассмотрение разнообразных задач естествознания приводит к сравнительно небольшому числу типов дифференциальных уравнений в частных производных, названных в связи с этим уравнениями математической физики. Это обстоятельство позволяет нам с некоторым оптимизмом приступить к изучению трудного и увлекательного раздела курса высшей математики, каковым являются дифференциальные уравнения в частных производных математической физики. Одними из основных понятий теории дифференциальных уравнений являются понятия частного и общего решения. Обратимся к рассмотрению этого вопроса в первую очередь.

§ 1. 2. ОБЫКНОВЕННЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ И ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ

В этом параграфе читателю предлагается вспомнить некоторые факты теории обыкновенных дифференциальных уравнений, без которых понимание последующего материала пособия будет затруднительным.

Рассмотрим обыкновенное дифференциальное уравнение n-го порядка с неизвестной функцией $u(x)$

$$F(x,\ u(x),\ u'(x),\ \ldots,\ u^{(n)}(x)) = 0. \tag{1.1}$$

Как известно, *порядком* дифференциального уравнения называется наивысший порядок производной от неизвестной функции, которая входит в это уравнение. Решить уравнение (1.1) это значит найти такую функцию $u = f(x)$, чтобы после замены в этом уравнении $u(x)$ на $f(x)$, $u'(x)$ на $f'(x)$, ..., $u^{(n)}(x)$ на $f^{(n)}(x)$ оно обращалось в тождество. Функция $u = f(x)$, обладающая этим свойством, называется *решением* рассматриваемого уравнения или его *интегралом*. Но задача состоит порой не только в том, чтобы находить отдельные, как их называют *частные решения*, но и в том, чтобы найти всю совокупность решений. Так приходят к понятию *общего решения* уравнения (1.1). В уравнении (1.1) порядка n совокупность решений может быть задана функцией, зависящей от независимой переменной x и, кроме того, от n произвольных постоянных $C_1, C_2, \ldots, C_n$

$$u = \varphi(x, C_1, C_2, \ldots, C_n). \tag{1.2}$$

Очевидно, эта функция должна являться решением уравнения (1.1) при всех, вообще говоря, значениях произвольных постоянных $C_1, C_2, \ldots, C_n$. Имея общее решение (1.2), можно ставить задачу о выделении из него частных решений, удовлетворяющих некоторым дополнительным условиям, которые обычно присоединяются к заданному уравнению. Задачу нахождения такого решения называют обычно *краевой задачей*. Наиболее изученной среди краевых задач является, пожалуй, задача Коши: среди бесконечного числа решений (1.2) уравнения (1.1) найти такое решение, которое удовлетворяет начальным условиям

$$u(x_0) = u_0;\ u'(x_0) = u'_0;\ \ldots;\ u^{(n-1)}(x_0) = u_0^{(n-1)},$$

где x_0 принадлежит промежутку значений x, на котором и отыскивается решение. В теории обыкновенных дифференциальных уравнений доказывается, что при весьма общих ограничениях, накладываемых на функцию F, задача Коши для уравнения (1.1) всегда имеет решение. Характерной особенностью задачи Коши является то, что значения искомой функции $u = u(x)$ и ее производных до $(n-1)$-го порядка включительно задаются в одной точке $x = x_0$, принадлежащей промежутку рассмотрения. Но значительно чаще появляется необходимость в нахождении такого решения $u(x)$ уравнения (1.1), когда начальные условия заданы не в одной точке $x = x_0$, а в нескольких точках (в этом случае их принято называть краевыми условиями). В принципе краевые условия для уравнения (1.1) могут быть самые разнообразные. Например, часто возникает задача о нахождении решения $u(x)$ уравнения (1.1), удовлетворяющего условиям

$$u(x_1) = A_1;\ u(x_2) = A_2;\ \ldots;\ u(x_n) = A_n.$$

Можно поставить задачу о нахождении решения $u(x)$ уравнения (1.1), удовлетворяющего краевым условиям вида

$$u(x_1)=A_1;\ u'(x_1)=A_2;\ \ldots;\ u^{(n-2)}(x_1)=A_{n-1};\ u(x_2)=B.$$

Можно ставить и другие краевые условия.

Решение краевой задачи является более сложной проблемой, чем, например, решение задачи Коши. Дело в том, что не всякая краевая задача имеет решение, а даже если и имеет решение, то не единственное. В связи с этим теория краевых задач для обыкновенных дифференциальных уравнений еще далека до завершения и ее успехи в значительной степени будут стимулировать решение многих задач физики и техники (см. ч. 2, стр. 168).

Дифференциальное уравнение в частных производных с m независимыми переменными $x_1, x_2, \ldots, x_m$ относительно неизвестной функции $u(x_1, x_2, \ldots, x_m)$ можно записать в виде

$$F\left(x_1, x_2, \ldots, x_m, \frac{\partial u}{\partial x_1}, \frac{\partial u}{\partial x_2}, \ldots, \frac{\partial u}{\partial x_m}, \frac{\partial^2 u}{\partial x_1^2}, \ldots, \right.$$
$$\left. \frac{\partial^2 u}{\partial x_1 \partial x_2}, \ldots, \frac{\partial^2 u}{\partial x_m^2}, \ldots, \frac{\partial^n u}{\partial x_m^n}\right)=0, \tag{1.3}$$

где F — заданная функция своих аргументов.

Наивысший порядок n производной от неизвестной функции, входящей в дифференциальное уравнение, называется *порядком* этого уравнения.

Решением или *интегралом* уравнения (1.3) называется такая функция $u=f(x_1, x_2, \ldots, x_m)$, которая, будучи подставлена в уравнение вместо неизвестной функции, обращает это уравнение в тождество по независимым переменным $x_1, x_2, \ldots, x_m$.

Задача интегрирования уравнения (1.3) состоит не только в том, чтобы находить частные его решения, но и в том, чтобы получить всю совокупность решений. Другими словами, и в теории уравнений с частными производными естественно ставить вопрос об общем решении, т. е. таком решении, из которого можно выделить индивидуальное решение с помощью дополнительных условий, присоединяемых к дифференциальному уравнению.

Каков же характер общего решения дифференциального уравнения в частных производных? Будет ли оправданным предполагать, что и в этом случае совокупность всех решений уравнения является функцией, зависящей не только от независимых переменных, но и от произвольных постоянных как в случае обыкновенного дифференциального уравнения? А может быть роль произвольных постоянных будут выполнять другие произвольные элементы? На все эти вопросы нам поможет ответить рассмотрение простейших уравнений в частных производных.

Пример 1. Найти решение (проинтегрировать) уравнения

$$\frac{\partial u}{\partial x} = y - x. \tag{1.4}$$

Решение. Данное уравнение есть уравнение первого порядка в частных производных относительно неизвестной функции $u(x, y)$. Оно не содержит производной $\frac{\partial u(x, y)}{\partial y}$, а потому мы можем, фиксируя y, рассматривать его как обыкновенное дифференциальное уравнение. Интегрируя по x, получим

$$u(x, y) = xy - \frac{x^2}{2} + \varphi(y), \tag{1.5}$$

где $\varphi(y)$ — произвольная функция y.

Вас не должно смущать, что в правой части равенства (1.5) в качестве одного из слагаемых после интегрирования обеих частей уравнения (1.4) по x появилась произвольная функция $\varphi(y)$. Функция (1.5) будет удовлетворять уравнению (1.4) при любых $\varphi(y)$, в чем легко убедиться дифференцированием обеих частей равенства (1.5) по x.

Итак, уравнение в частных производных первого порядка (1.4) имеет бесчисленное множество решений (1.5), зависящих от одной произвольной функции $\varphi(y)$.

Пример 2. Найти решение уравнения

$$\frac{\partial u(x, y)}{\partial y} = y^3. \tag{1.6}$$

Решение. Заданное уравнение есть уравнение первого порядка в частных производных относительно неизвестной функции $u(x, y)$, в котором отсутствует производная $\frac{\partial u(x, y)}{\partial x}$. Интегрируя по y, получим

$$u(x, y) = \frac{y^4}{4} + \psi(x), \tag{1.7}$$

где $\psi(x)$ — произвольная функция x.

Простой проверкой убеждаемся, что полученная совокупность функций (1.7), зависящих от одной произвольной функции $\psi(x)$, удовлетворяет уравнению (1.6).

Рассмотрение приведенных и других аналогичных примеров заставляет нас сделать ряд выводов.

Во-первых, дифференциальное уравнение в частных производных имеет бесчисленное множество решений. Во-вторых, уравнение в частных производных первого порядка имеет решение, содержащее одну произвольную функцию. Второе обстоятельство указывает на коренное отличие решения обыкновенного дифференциального уравнения первого порядка, содержащего лишь произвольную постоянную, от решения уравнения в частных производных первого порядка.

По аналогии с общим решением обыкновенного дифференциального уравнения решения (1.5) и (1.7) соответственно уравнений (1.4) и (1.6) будем называть общими решениями этих уравнений.

Итак, общее решение уравнения в частных производных первого порядка содержит одну произвольную функцию.

Пример 3. Проверить, что функция

$$u(x, y) = x + \varphi(x - y), \tag{1.8}$$

где φ — произвольная дифференцируемая функция, является общим решением уравнения

$$\frac{\partial u(x, y)}{\partial x} + \frac{\partial u(x, y)}{\partial y} = 1. \tag{1.9}$$

Решение. Достаточно убедиться, что функция (1.8) удовлетворяет уравнению (1.9). Для этого из (1.8) с учетом правила дифференцирования сложной функции, найдем

$$\frac{\partial u}{\partial x} = 1 + \frac{\partial \varphi}{\partial v}\,\frac{\partial v}{\partial x} \text{ и } \frac{\partial u}{\partial y} = \frac{\partial \varphi}{\partial v}\,\frac{\partial v}{\partial y},$$

где $v = x - y$. Окончательно получим

$$\frac{\partial u}{\partial x} = 1 + \frac{\partial \varphi}{\partial v}; \quad \frac{\partial u}{\partial y} = -\frac{\partial \varphi}{\partial v}.$$

Отсюда уже видно, что функция (1.8) действительно является общим решением уравнения (1.9).

Пример 4. Проинтегрировать уравнение

$$\frac{\partial^2 u(x, y)}{\partial x\, \partial y} = 0. \tag{1.10}$$

Решение. Данное уравнение есть уравнение в частных производных второго порядка. Его интегрирование не представляет труда, так как, используя определение частной производной второго порядка, уравнение (1.10) можно переписать в виде

$$\frac{\partial}{\partial x}\left(\frac{\partial u(x, y)}{\partial y}\right) = 0.$$

Производная по x некоторой функции $\dfrac{\partial u(x, y)}{\partial y}$ равна нулю. Значит эта функция от x не зависит. Отсюда следует, что

$$\frac{\partial u(x, y)}{\partial y} = f(y),$$

где f — произвольная функция y.

Общим решением уравнения (1.10) будет функция

$$u(x, y) = \int f(y)\, dy + \psi(x). \tag{1.11}$$

Так как $f(y)$ — произвольная функция y, то и интеграл от нее будет произвольной функцией y. Введем обозначение

$$\int f(y)\,dy = \varphi(y).$$

Окончательно получим

$$u(x, y) = \varphi(y) + \psi(x). \qquad (1.12)$$

Итак, общее решение (1.12) уравнения в частных производных второго порядка (1.10) содержит две произвольные функции $\varphi(y)$ и $\psi(x)$.

Пример 5. Проверить, что функция

$$u(x, y) = \varphi(x + ay) + \psi(x - ay), \qquad (1.13)$$

где φ и ψ — две произвольные дифференцируемые функции, является общим решением уравнения в частных производных второго порядка

$$\frac{\partial^2 u}{\partial y^2} - a^2 \frac{\partial^2 u}{\partial x^2} = 0. \qquad (1.14)$$

Решение. Введя обозначения $x + ay = v$ и $x - ay = w$ и помня правила дифференцирования сложной функции, из (1.13) последовательно находим

$$\frac{\partial u}{\partial x} = \frac{\partial \varphi}{\partial v} + \frac{\partial \psi}{w}; \quad \frac{\partial u}{\partial y} = a\frac{\partial \varphi}{\partial v} - a\frac{\partial \psi}{\partial w};$$

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 \varphi}{\partial^2 v} + \frac{\partial^2 \psi}{\partial w^2}; \quad \frac{\partial^2 u}{\partial y^2} = a^2\frac{\partial^2 \varphi}{\partial v^2} + a^2\frac{\partial^2 \psi}{\partial w^2}.$$

Подставляя выражения для производных $\frac{\partial^2 u}{\partial x^2}$ и $\frac{\partial^2 u}{\partial y^2}$ в левую часть уравнения (1.14), убеждаемся, что она обращается в нуль.

Пример 6. Проинтегрировать уравнение

$$\frac{\partial^2 u(x, y)}{\partial x\,\partial y} - \frac{1}{2x}\frac{\partial u(x, y)}{\partial y} = 0, \quad x > 0; \; y > 0.$$

Решение. Данное уравнение представляет собой уравнение в частных производных второго порядка. Его можно переписать в виде

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right) - \frac{1}{2x}\frac{\partial u}{\partial y} = 0.$$

Новое представление уравнения позволяет предположить, что с введением новой функции $v(x, y)$ по формуле $\frac{\partial u}{\partial y} = v(x, y)$ мы сведем его к обыкновенному дифференциальному уравнению относительно функции v, где y будет играть роль параметра

$$\frac{\partial v}{\partial x} - \frac{1}{2x}v = 0.$$

Разделяя переменные и интегрируя, получим

$$\frac{\partial v}{v} = \frac{\partial x}{2x}; \; \ln|v| = \frac{1}{2}\ln|x| + \ln|f(y)|,$$

где $f(y)$ — произвольная функция y. После потенцирования находим

$$v = f(y)\sqrt{x}.$$

Но $\frac{\partial u}{\partial y} = v$, поэтому $\frac{\partial u}{\partial y} = f(y)\sqrt{x}$. Отсюда

$$u(x, y) = \sqrt{x}\int f(y)\,dy + \psi(x),$$

где $\psi(x)$ — произвольная функция x.

Вводя обозначение

$$\int f(y)\,dy = \varphi(y),$$

окончательно получим

$$u(x, y) = \sqrt{x}\,\varphi(y) + \psi(x),$$

где $\varphi(y)$ и $\psi(x)$ — произвольные функции y и x соответственно.

Таким образом, общее решение уравнения в частных производных второго порядка содержит две произвольные функции.

Следует ожидать, что общее решение уравнения в частных производных n-го порядка будет содержать n произвольных функций, т. е. будет зависеть от числа произвольных функций, равного порядку уравнения.

Естественно было бы теперь поставить задачу об использовании общего решения заданного уравнения для определения частного его решения, удовлетворяющего некоторым дополнительным условиям. Оставляя выяснение этого вопроса на более позднее время, заметим, что упомянутые дополнительные условия не будут теперь совпадать по характеру с теми, которые уже встречались в теории обыкновенных дифференциальных уравнений: ведь в общее решение соответствующего уравнения входят не произвольные постоянные, а произвольные функции. Это обстоятельство в корне меняет дело. Наиболее ярко это различие можно проиллюстрировать сопоставлением геометрических интерпретаций решений обыкновенного дифференциального уравнения и уравнения в частных производных.

Рассмотрим уравнение в частных производных первого порядка

$$F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) = 0. \qquad (1.15)$$

Решение $z = z(x, y)$ этого уравнения можно истолковать как поверхность — интегральную поверхность в пространстве x, y, z.

Наряду с уравнением (1.15) рассмотрим обыкновенное дифференциальное уравнение первого порядка

$$F\left(x, y, \frac{dy}{dx}\right) = 0,$$

которое может быть переписано в виде, разрешенном относительно производной,

$$\frac{dy}{dx} = f(x, y). \tag{1.16}$$

Обыкновенное дифференциальное уравнение (1.16) устанавливает зависимость между координатами точки плоскости xOy и угловым коэффициентом касательной $\frac{dy}{dx}$ к графику решения в той же точке. Зная x и y, можно вычислить $\frac{dy}{dx}$. Таким образом, уравнение (1.16) определяет поле направлений и задача интегрирования этого уравнения заключается в том, чтобы найти такую кривую, направление касательных к которой в каждой ее точке совпадает с направлением поля. Эта кривая и будет интегральной кривой заданного уравнения. При весьма общих условиях, накладываемых на функцию $f(x, y)$, через каждую точку плоскости xOy будет проходить своя интегральная кривая. Эти кривые образуют, таким образом, бесконечное семейство кривых, зависящее от одного параметра — произвольной постоянной. Из предыдущего следует, что отдельная интегральная кривая (частное решение) вполне определяется заданием какой-либо точки x_0, y_0 (начальные условия).

Попробуем для уравнения в частных производных (1.15) установить нечто аналогичное, но уже теперь, конечно, в трехмерном пространстве x, y, z. Место направления, о котором шла речь выше, соответствующего некоторой точке $P_0(x_0, y_0, z_0)$, должен занять наклон плоскости

$$\frac{\partial z}{\partial x}(x - x_0) + \frac{\partial z}{\partial y}(y - y_0) - (z - z_0) = 0, \tag{1.17}$$

где $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$ — пара значений частных производных, удовлетворяющих уравнению

$$F\left(x_0, y_0, z_0, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) = 0, \tag{1.18}$$

так как интегральная поверхность уравнения (1.15) $z = z(x, y)$, проходящая через точку $P_0(x_0, y_0, z_0)$, имеет в качестве касательной плоскости в точке P_0 как раз плоскость (1.17) со значениями коэффициентов $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$, удовлетворяющих уравнению (1.18). В связи с тем, что уравнению (1.18) удовлетворяет бесчисленное множество пар $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$ (ведь уравнение (1.18) определяет не сами значения $\frac{\partial z}{\partial x}$ и $\frac{\partial z}{\partial y}$, а только некоторую определенную зависимость между ними), касательная плоскость в любой точке P_0 определяется не однозначно. Отсюда следует, что одной точки недостаточно для однозначного определения интегральной поверхности, т. е. множество интегральных поверхностей уравнения в частных производных оказывается значительно богаче множества интегральных кривых обыкновенного дифференциального уравнения. В этом сказывается различие между общим решением обыкновенного дифференциального уравнения (1.16), зависящим от одного произвольного элемента — постоянной величины, и общим решением уравнения в частных производных (1.15), зависящим тоже от одного произвольного элемента, но уже функции. Теперь для однозначного определения интегральной поверхности нужно задать не точку в пространстве, а целую пространственную кривую, через которую эта поверхность и должна проходить.

§ 1. 3. ЛИНЕЙНЫЕ ДИФФЕРЕНЦИАЛЬНЫЕ УРАВНЕНИЯ В ЧАСТНЫХ ПРОИЗВОДНЫХ. ПРИНЦИП СУПЕРПОЗИЦИИ

Особое место в общей теории дифференциальных уравнений в частных производных занимают линейные уравнения. Это объясняется тем, что, с одной стороны, они составляют наиболее разработанную часть этой теории, а с другой стороны, описывая реальные физические процессы, находят многочисленные приложения в физике и технике. Прежде чем приступить к построению методов решения таких уравнений нам придется познакомиться с целым рядом замечательных свойств их решений, напоминающих соответствующие свойства решений обыкновенных дифференциальных уравнений.

Дифференциальное уравнение называется *линейным*, если неизвестная функция и ее производные входят в уравнение в первой степени и если оно не содержит членов с произведением этих величин.

Ограничиваясь здесь случаем линейного уравнения в частных производных второго порядка с неизвестной функцией, зависящей от двух переменных, мы будем помнить, что все приведенные ниже свойства будут справедливыми и в самом общем случае линейного уравнения в частных производных.

Дифференциальное уравнение в частных производных второго порядка, линейное и однородное относительно неизвестной функции и ее производных имеет вид

$$L[u] \equiv a_{11}\frac{\partial^2 u}{\partial x^2} + 2a_{12}\frac{\partial^2 u}{\partial x\,\partial y} + a_{22}\frac{\partial^2 u}{\partial y^2} + b_1\frac{\partial u}{\partial x} + b_2\frac{\partial u}{\partial y} + cu = 0, \quad (1.19)$$

где a_{11}, a_{12}, a_{22}, b_1, b_2, c — заданные функции от x и y — коэффициенты этого уравнения. (Не напоминает оно вам общее уравнение кривой второго порядка, где вместо текущих координат подставлены частные производные от неизвестной функции $u(x, y)$ соответствующего порядка?)

Если правая часть уравнения отлична от нуля, т. е.

$$L[u] = f(x, y),$$

где $f(x, y)$ — известная функция от x и y, то линейное уравнение называется *неоднородным*.

В основе важнейших методов интегрирования линейных уравнений в частных производных лежит общая для всех них идея. Она заключается в свойстве решений, которое называется *принципом суперпозиции (наложения)*. Его можно высказать в виде следующей теоремы.

Теорема. *Если каждая из функций*

$$u_1(x, y),\ u_2(x, y),\ \ldots,\ u_n(x, y)$$

является решением однородного уравнения (1.19), то и их линейная комбинация

$$u(x, y) = C_1u_1(x, y) + C_2u_2(x, y) + \ldots + C_nu_n(x, y),$$

где C_1, C_2, ..., C_n — произвольные постоянные, также является решением этого уравнения.

Доказательство. Достаточно проверить, что $L[u] = 0$, если $L[u_1] = 0$; $L[u_2] = 0$; ... ; $L[u_n] = 0$. Используя правила дифференцирования, легко получаем цепочку равенств

$$L[u] = C_1L[u_1] + C_2L[u_2] + \ldots + C_nL[u_n] = 0,$$

что и требовалось доказать.

Принцип суперпозиции допускает следующее важное обобщение. Предположим, что мы располагаем не конечным числом частных решений уравнения (1.19), а бесконечным их числом

$$u_1(x, y),\ u_2(x, y),\ \ldots,\ u_n(x, y),\ \ldots$$

Составим ряд из этой бесконечной последовательности решений

$$C_1u_1(x, y) + C_2u_2(x, y) + \ldots + C_nu_n(x, y) + \ldots = \sum_{i=1}^{\infty} C_iu_i(x, y).$$

Этот ряд будет решением уравнения (1.19), если, во-первых, он сходится к некоторой функции $u(x, y)$ и, во-вторых, допускает почленное дифференцирование.

Нетрудно вспомнить, что принцип суперпозиции, аналогичный сформулированному нами в виде теоремы для уравнения (1.19), имеет место и для обыкновенных линейных дифференциальных уравнений. Но из теории этих уравнений известно также, что линейное уравнение n-го порядка имеет не более n линейно независимых решений. Поэтому в этом случае имеет смысл говорить лишь о комбинировании этих решений в числе не более n.

В случае же линейного уравнения в частных производных можно выбрать бесконечное множество линейно независимых частных решений. (Бесконечное множество функций называется линейно независимым, если любое конечное множество этих функций будет линейно независимым.) Это обстоятельство совместно с обобщенным принципом суперпозиции играет в образовании решения уравнения в частных производных чрезвычайно важную роль. Именно большой набор частных решений такого уравнения даст возможность выделить из общего решения то, которое должно удовлетворять, вообще говоря, произвольным дополнительным условиям.

Обобщение принципа суперпозиции идет еще дальше. Может оказаться, что решение уравнения в частных производных $L[u]=0$ имеет вид $u(x, y, \Theta)$, т. е. зависит от некоторого параметра, изменяющегося в конечном или бесконечном промежутке. Эта зависимость понимается так, что при всех значениях Θ функция $u(x, y, \Theta)$ является решением этого уравнения, т. е. $L[u(x, y, \Theta)]\equiv 0$. Из предыдущего следует, что в свою очередь функция $C(\Theta)\,u(x, y, \Theta)$, где $C(\Theta)$ — произвольная функция, будет также решением уравнения $L[u]=0$: ведь мы попадаем в условия доказанной выше теоремы, где роль произвольной постоянной играет $C(\Theta)$. Как и ранее, интерес представляет бесконечная сумма выписанных решений. Но так как $C(\Theta)$ зависит от непрерывно меняющегося параметра Θ, то это суммирование должно осуществляться при помощи интеграла. Итак, мы приходим к рассмотрению интеграла

$$\int C(\Theta)\,u(x, y, \Theta)\,d\Theta, \tag{1.20}$$

где интегрирование производится в пределах изменения параметра Θ. Читатель без особого труда убедится в том, что интеграл (1.20) является решением уравнения в частных производных $L[y]=0$, конечно, при условиях его существования и законности выполнения операции дифференцирования по x и y под знаком интеграла (см. ч. 1, стр. 65).

§ 1. 4. О ЗАДАЧАХ, ПРИВОДЯЩИХ К УРАВНЕНИЯМ В ЧАСТНЫХ ПРОИЗВОДНЫХ

Основные уравнения математической физики, которыми мы намерены заниматься в дальнейшем, могут быть получены при помощи законов механики и физики, как правило, записанных в виде интегральных уравнений или уравнений в частных производных.

В предлагаемом параграфе будет рассмотрен ряд задач, каждая из которых приводит к уравнению в частных производных определенного типа. Для инженера представляет большой интерес не только сам тип получившегося уравнения, но и особенности физической системы, описываемой уравнением именно этого типа, а не другого. От знания физической природы системы во многом зависит успех применения основных законов механики и физики.

Перед тем как приступить к чтению этого параграфа рекомендуется повторить основные положения теории поля (см. ч. 2, гл. 1).

Отдавая дань традиции, начнем рассмотрение примеров на применение законов механики и физики в дифференциальной форме с уравнений движения.

1. 4. 1. Уравнения движения

Уравнение колебания струны. Струной называют тонкую сильно натянутую нить, сделанную из упругого вещества. Предполагают, что струна обладает абсолютной гибкостью, т. е. не сопротивляется изгибу. Это значит, что если мысленно провести в струне сечение, разделив ее на две части, то со стороны каждого из отрезков струны на другой будет действовать сила, равная натяжению и направленная по касательной к линии струны. Упругость струны при достаточно малых ее отклонениях от положения равновесия гарантирует отсутствие удлинения струны. Отсюда натяжение струны остается одним и тем же во всех сечениях в течение всего колебания. Под малыми колебаниями струны, которые здесь и будут рассматриваться, понимают такие ее колебания, когда струна в произвольный момент времени имеет столь малый угловой коэффициент, что квадратом его можно пренебречь. И еще одно условие. Нас будут интересовать лишь плоские поперечные колебания струны, т. е. такое движение струны, когда колебание происходит в одной плоскости и каждая точка струны перемещается только по перпендикуляру к положению ее равновесия.

Пусть в состоянии равновесия (покоя) струна расположена вдоль оси Ox. Обозначим отклонение точки струны от положения равновесия через u. Очевидно в процессе колебания величина u — функция абсциссы точки струны x и момента времени t, т. е. $u=u(x, t)$. Действительно, при фиксированном $t=t_0$ уравнение $u=u(x, t_0)$ определяет форму струны в момент времени t_0 (рис. 1.3). При фиксированном $x=x_0$

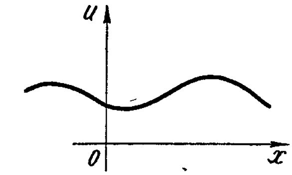

Рис. 1.3

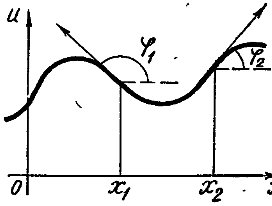

Рис. 1.4

уравнение $u=u(x_0, t)$ определяет закон движения точки струны с абсциссой $x=x_0$. Итак, функция $u=u(x, t)$ исчерпывающим образом описывает движение всех точек струны около положения ее равновесия.

Теперь наша задача состоит в том, чтобы составить дифференциальное уравнение движения струны, которому удовлетворяет функция $u(x, t)$.

Рассмотрим участок струны между точками x_1 и x_2 (рис. 1.4). Как это следует из предыдущего, в этих точках будут действовать две силы, равные натяжению T и направленные по касательным к струне.

Согласно закону механики равнодействующая этих двух сил равна скорости изменения количества движения выделенного участка струны. Очевидно, эта равнодействующая будет отлична от нуля, если рассматриваемый участок струны имеет некоторую кривизну.

Обозначая линейную плотность струны через ρ, запишем, чему равна скорость изменения количества движения участка струны между x_1 и x_2:

$$\int_{x_1}^{x_2} \rho \frac{\partial^2 u}{\partial t^2}\, dx.$$

Если через φ обозначить угол, составляемый касательной к струне с осью Ox, то равнодействующая сил натяжения, действующих в точках x_1 и x_2, будет равна

$$T \sin \varphi_2 - T \sin \varphi_1$$

(Так как колебания струны поперечные, то составляющая по оси Ox суммы всех сил, приложенных к участку струны, будет равна нулю, а поэтому нужно учесть лишь алгебраическую сумму составляющих по оси Ou.)

Итак, уравнение, выражающее второй закон механики, будет иметь вид

$$T\sin\varphi_2 - T\sin\varphi_1 = \int_{x_1}^{x_2} \rho \frac{\partial^2 u}{\partial t^2}\, dx. \tag{1.21}$$

Уравнение (1.21) страдает существенным недостатком: левая его часть не выражена через неизвестную функцию $u(x, y)$. Если нам удасться связать $T\sin\varphi$ с $u(x, y)$, тогда задача будет решена окончательно: получим уравнение, описывающее движение струны при сделанных выше предположениях.

Вспоминая в чем состоит геометрический смысл производной, получим $\operatorname{tg}\varphi = \frac{\partial u}{\partial x}$. По известной формуле тригонометрии выразим $\sin\varphi$ через $\operatorname{tg}\varphi$:

$$\sin\varphi = \frac{\operatorname{tg}\varphi}{\sqrt{1+\operatorname{tg}^2\varphi}},$$

откуда

$$\sin\varphi = \frac{\frac{\partial u}{\partial x}}{\sqrt{1+\left(\frac{\partial u}{\partial x}\right)^2}} \approx \frac{\partial u}{\partial x},$$

так как по предположению квадратом углового коэффициента можно пренебречь. Окончательно уравнение (1.21) можно переписать в следующем виде

$$T\frac{\partial u}{\partial x}\bigg|_{x=x_2} - T\frac{\partial u}{\partial x}\bigg|_{x=x_1} = \int_{x_1}^{x_2} \rho \frac{\partial^2 u}{\partial t^2}\, dx$$

или с учетом формулы Лейбница — Ньютона

$$\int_{x_1}^{x_2} \frac{\partial}{\partial x}\left(T\frac{\partial u}{\partial x}\right) dx = \int_{x_1}^{x_2} \rho \frac{\partial^2 u}{\partial t^2}\, dx.$$

Отсюда получим уравнение

$$T\frac{\partial^2 u}{\partial x^2} = \rho\frac{\partial^2 u}{\partial t^2}, \tag{1.22}$$

которое называется *уравнением колебания струны*.

Можно получить уравнение колебания струны с учетом сил, действующих на нее параллельно оси Ou в плоскости колебания. Эти силы могут меняться вдоль струны со временем, т. е. являться функциями абсциссы x и времени t. Если через $g(x, t)$ обозначить плотность распределения этих сил, отнесенную к единице массы, то равнодействующая их на участке струны от x_1 до x_2 в момент t будет равна

$$\int_{x_1}^{x_2} \rho\, g(x, t)\, dx.$$

Очевидно, при этих условиях уравнение (1.22) в левой части приобретет еще одно слагаемое $\rho\, g(x, t)$, т. е. будет иметь вид

$$T\frac{\partial^2 u}{\partial x^2} + \rho\, g(x, t) = \rho\frac{\partial^2 u}{\partial t^2}.$$

Поделив обе части выписанного равенства на ρ и введя обозначение $\frac{T}{\rho} = a$, окончательно получим

$$\frac{\partial^2 u}{\partial t^2} - a^2\frac{\partial^2 u}{\partial x^2} = g(x, t). \tag{1.23}$$

Уравнение (1.23) представляет собой линейное уравнение в частных производных второго порядка с постоянными коэффициентами. Его часто называют также *одномерным волновым уравнением*.

Если на струну действуют только силы тяжести и натяжение струны T велико, то естественно рассматривать уравнение струны в виде (1.22), где $g(x, t) \equiv 0$. В этом случае уравнение (1.22), будучи однородным, описывает *свободные колебания струны*. Если же $g(x, t) \neq 0$, то уравнение колебаний струны, будучи неоднородным, описывает так называемые *вынужденные колебания струны*, т. е. колебания, происходящие под действием внешних сил.

Уравнение движения среды. Мы не можем обойти чрезвычайно важный для приложений случай выражения физического закона при помощи дифференциальных уравнений в частных производных, каковым является уравнение движения жидкой или газообразной среды.

Пусть среда состоит из материальных частиц, движущихся каждая со своей скоростью. Математическое описание состояния движущейся среды осуществляется с помощью функций $v_x(x, y, z, t)$, $v_y(x, y, z, t)$, $v_z(x, y, z, t)$, представляющих компоненты вектора скорости $\vec{v}(x, y, z, t)$ среды в каждой данной точке x, y, z пространства в момент t, т. е. относится к определенным точкам пространства, а не к определенным частицам среды, передвигающимся со временем в пространстве. Характер движения среды будет зависеть также от плотности среды $\rho(x, y, z, t)$, давления $p(x, y, z, t)$ и внешних действующих сил $\vec{F}(x, y, z, t)$, приложенных к единице объема.

Прежде чем приступить к выводу уравнений движения среды, получим уравнение, выражающее закон сохранения вещества.

Рассмотрим некоторый объем V пространства. Подсчитаем массу вещества, находящегося в этом объеме. Она будет выражаться интегралом

$$m = \iiint\limits_V \rho(x, y, z, t)\, dV. \tag{1.24}$$

Эта масса не будет постоянной, так как в процессе колебаний плотность в каждой точке будет изменяться за счет того, что частицы вещества будут то входить в фиксированный объем, то выходить из него. Скорость изменения массы можно получить, дифференцируя по времени интеграл (1.24):

$$\frac{dm}{dt} = \iiint\limits_V \frac{\partial \rho}{\partial t}\, dV.$$

Если внутри объема V нет источников и стоков, то изменение в единицу времени массы вещества, заключенной внутри V, равно потоку этого вещества через поверхность S, ограничивающую этот объем (см. ч. 2, стр. 29). Другими словами имеет место равенство

$$\iiint\limits_V \frac{\partial \rho}{\partial t}\, dV = -\iint\limits_S \rho v_n\, dS, \tag{1.25}$$

где v_n — проекция вектора $\vec{v}(x, y, z, t)$ на направление внешней нормали $\vec{n}$ к поверхности S.

Интегральное соотношение (1.25), выражающее закон сохранения вещества, называют также *уравнением неразрывности в интегральной форме.*

Формула Остроградского поможет нам получить уравнение неразрывности и в дифференциальной форме. Действительно (см. ч. 2, стр. 33),

$$\iint\limits_S \rho v_n\, dS = \iiint\limits_V \operatorname{div}(\rho \vec{v})\, dV,$$

где

$$\operatorname{div}(\rho \vec{v}) = \frac{\partial (\rho v_x)}{\partial x} + \frac{\partial (\rho v_y)}{\partial y} + \frac{\partial (\rho v_z)}{\partial z}.$$

Подставляем полученное выражение в формулу (1.25):

$$\iiint\limits_V \frac{\partial \rho}{\partial t} dV = -\iiint\limits_V \operatorname{div}(\rho \vec{v})\, dV$$

или

$$\iiint\limits_V \left[\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{v}) \right] = 0.$$

Это равенство имеет место для любого объема V, а это означает, что оно возможно только в том случае, если подынтегральная функция тождественно равна нулю:

$$\frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{v}) = 0$$

или

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho v_x)}{\partial x} + \frac{\partial (\rho v_y)}{\partial y} + \frac{\partial (\rho v_z)}{\partial z} = 0. \tag{1.26}$$

Уравнение (1.26) называется *уравнением неразрывности* и является типичным примером выражения физического закона (закона сохранения вещества) на языке уравнений в частных производных.

Вернемся к выводу уравнений движения среды.

Выделим в нашей среде некоторый объем V, ограниченный поверхностью S. Наша задача состоит в том, чтобы записать второй закон Ньютона для частиц, находящихся в рассматриваемом объеме. Закон Ньютона говорит, что при движении среды скорость изменения суммарного количества движения частиц некоторого объема равна сумме всех сил, приложенных к данному объему.

Подсчитаем сначала скорость изменения количества движения. Как известно, количество движения частиц вещества, находящихся в объеме V, представляет собой векторную величину и выражается интегралом

$$\iiint\limits_V \rho \vec{v} dV.$$

Очевидно, частицы, заполнившие с плотностью ρ достаточно малый объем dV, через промежуток времени t будут заполнять с некоторой новой плотностью ρ' новый объем dV', но их общая масса будет прежней, т. е.

$$\rho' dV' = \rho dV.$$

При этом изменение скорости $\vec{v}$ до некоторого значения $\vec{v}'$, т. е. на величину

$$\Delta\vec{v} = \vec{v}' - \vec{v}$$

повлечет изменение количества движения на величину

$$\rho'\vec{v}'dV' - \rho\vec{v}dV = \rho\vec{v}'dV - \rho\vec{v}dV = \rho\Delta\vec{v}dV.$$

Подсчитаем измененное количество движения в единицу времени

$$\rho\frac{\Delta\vec{v}}{\Delta t}dV \approx \rho\frac{d\vec{v}}{dt}dV.$$

Для подсчета скорости изменения количества движения всех частиц вещества, находившихся в объеме V, нужно взять интеграл

$$\iiint\limits_V \rho\frac{d\vec{v}}{dt}dV. \tag{1.27}$$

Стоящая под знаком интеграла прозводная $\frac{d\vec{v}}{dt}$ определяет не изменение скорости вещества в данной неподвижной точке пространства, а изменение скорости определенной передвигающейся в пространстве частицы вещества.

Интегралу (1.27) в скалярной форме соответствует три интеграла

$$\iiint\limits_V \rho\frac{dv_x}{dt}dV;\quad \iiint\limits_V \rho\frac{dv_y}{dt}dV;\quad \iiint\limits_V \rho\frac{dv_z}{dt}dV. \tag{1.28}$$

В силу сделанного выше замечания, производные $\frac{d\vec{v}_x}{dt}$, $\frac{d\vec{v}_y}{dt}$, $\frac{d\vec{v}_z}{dt}$ обозначают скорость изменения той или иной составляющей $\vec{v}$ для данной двигающейся частицы. Если $x = x(t)$, $y = y(t)$, $z = z(t)$ — уравнение траектории частицы вещества, то

$$\frac{dv_x}{dt} = \frac{\partial v_x}{\partial t} + \frac{\partial v_x}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial v_x}{\partial y}\cdot\frac{dy}{dt} + \frac{\partial v_x}{\partial z}\cdot\frac{dz}{dt} =$$

$$= \frac{\partial v_x}{\partial t} + \frac{\partial v_x}{\partial x}v_x + \frac{\partial v_x}{\partial y}v_y + \frac{\partial v_x}{\partial z}v_z.$$

Аналогично получаем еще два равенства

$$\frac{dv_y}{dt} = \frac{\partial v_y}{\partial t} + \frac{\partial v_y}{\partial x}v_x + \frac{\partial v_y}{\partial y}v_y + \frac{\partial v_y}{\partial z}v_z;$$

$$\frac{dv_z}{dt} = \frac{\partial v_z}{\partial t} + \frac{\partial v_z}{\partial x} v_x + \frac{\partial v_z}{\partial y} v_y + \frac{\partial v_z}{\partial z} v_z.$$

Итак, скорость изменения количества движения вещества, находящегося в объеме V, выражается интегралом (1.27) или в скалярной эквивалентной форме интегралами (1.28). Теперь подсчитаем силы, приложенные к выделенному объему вещества. Для определенности будем полагать, что рассматриваемая среда представляет собой идеальную жидкость. Под идеальной жидкостью понимают сплошную среду, в которой внутренние силы сводятся к нормальному давлению. Таким образом, действие на выделенный объем остальной части жидкости приводится к силе, направленной в каждой точке поверхности S по внутренней нормали. Итак, поверхностные силы, действующие на элемент поверхности dS, будут иметь величину $p(x, y, z, t)\,dS$ и будут направлены в сторону, обратную внешней нормали. Если единичный вектор внешней нормали к поверхности S обозначить через $\vec{n}$, то силы, приложенные к участку поверхности dS, будут равны

$$-p\vec{n}\,dS,$$

а равнодействующая сил давления, приложенных к поверхности S, равна

$$-\iint_S p\vec{n}\,dS. \tag{1.29}$$

Осталось учесть равнодействующую внешних сил $\vec{F}(x, y, z, t)$, приложенных к объему V. Она выразится интегралом

$$\iiint_V \vec{F}\,dV. \tag{1.30}$$

Но скорость изменения количества движения вещества равна сумме всех сил, приложенных к выделенному объему. С учетом (1.27), (1,29) и (1.30) составляем равенство

$$\iiint_V \rho\frac{d\vec{v}}{dt}\,dV = \iiint_V \vec{F}\,dV - \iint_S p\vec{n}\,dS. \tag{1.31}$$

Уравнение (1.31) и представляет собой уравнение движения в интегральной форме. Преобразуем его в уравнение в дифференциальной форме. Очевидно, это преобразование будет легко выполнимо, если удастся выразить поверхностный интеграл в правой части формулы (1.31) через объемный интеграл. Проделаем предварительные преобразования

$$p\vec{n} = p\cos(\vec{n}, \vec{x})\,\vec{i} + p\cos(\vec{n}, \vec{y})\,\vec{j} + p\cos(\vec{n}, \vec{z})\,\vec{k},$$

но

$$\cos(\vec{n}, \vec{x})\, dS = dy\, dz;\ \cos(\vec{n}, \vec{y})\, dS = dx\, dz;\ \cos(\vec{n}, \vec{z})\, dS = dx\, dy,$$

тогда

$$\iint_S p\cos(\vec{n}, \vec{x})\, dS = \iint_S p\, dy\, dz;$$

$$\iint_S p\cos(\vec{n}, \vec{y})\, dS = \iint_S p\, dx\, dz;$$

$$\iint_S p\cos(\vec{n}, \vec{z})\, dS = \iint_S p\, dx\, dy.$$

Формула Остроградского (см. ч. 2, стр. 35) даст возможность преобразовать поверхностный интеграл в объемный. Действительно,

$$\iint_S p\, dy\, dz = \iiint_V \frac{\partial p}{\partial x}\, dV;$$

$$\iint_S p\, dx\, dz = \iiint_V \frac{\partial p}{\partial y}\, dV;$$

$$\iint_S p\, dx\, dy = \iiint_V \frac{\partial p}{\partial z}\, dV.$$

Умножая каждое из этих равенств соответственно на $\vec{i}$, $\vec{j}$, $\vec{k}$ и складывая, получим

$$\iint_S p\vec{n}\, dS = \iiint_V \operatorname{grad} p\, dV,$$

где

$$\operatorname{grad} p = \frac{\partial p}{\partial x}\vec{i} + \frac{\partial p}{\partial y}\vec{j} + \frac{\partial p}{\partial z}\vec{k}.$$

Уравнение (1.31) может быть переписано в виде

$$\iiint_V \rho\frac{d\vec{v}}{dt}\, dV = \iiint_V \vec{F}\, dV - \iiint_V \operatorname{grad} p\, dV$$

или

$$\iiint_V \left(\rho\frac{d\vec{v}}{dt} - \vec{F} + \operatorname{grad} p\right) dV = 0,$$

что возможно только в том случае, если подынтегральная функция равна нулю тождественно

$$\rho \frac{d\vec{v}}{dt} - \vec{F} + \operatorname{grad} p = 0. \tag{1.32}$$

Векторное уравнение (1.32) эквивалентно системе трех скалярных уравнений

$$\left.\begin{aligned} \rho \frac{dv_x}{dt} + \frac{\partial p}{\partial x} &= F_x; \\ \rho \frac{dv_y}{dt} + \frac{\partial p}{\partial y} &= F_y; \\ \rho \frac{dv_z}{dt} + \frac{\partial p}{\partial z} &= F_z. \end{aligned}\right\} \tag{1.33}$$

Итак, второй закон Ньютона выражен нами в виде системы трех уравнений в частных производных.

Уравнения движения (1.33) могут быть использованы для вывода малых колебаний газовой среды, например звуковых колебаний.

Предположим, что внешние силы отсутствуют, т. е. $F_x = F_y = F_z = 0$. Тогда уравнения движения перепишутся в виде

$$\rho \frac{dv_x}{dt} + \frac{\partial p}{\partial x} = 0; \ \rho \frac{dv_y}{dt} + \frac{\partial p}{\partial y} = 0; \ \rho \frac{dv_z}{dt} + \frac{\partial p}{\partial z} = 0.$$

При малых колебаниях газа около положения равновесия плотность меняется мало и величины $\frac{\partial \rho}{\partial x}$, $\frac{\partial \rho}{\partial y}$, $\frac{\partial \rho}{\partial z}$, $\frac{\partial \rho}{\partial t}$ — достаточно малые, произведениями которых на v_x, v_y и v_z можно пренебречь. Тогда из уравнения неразрывности (1.26) получим

$$\frac{\partial \rho}{\partial t} + \rho \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \right) = 0.$$

Дифференцируя это уравнение по времени и пренебрегая произведениями $\frac{\partial \rho}{\partial t}$ на $\frac{\partial v_x}{\partial x}$, $\frac{\partial v_y}{\partial y}$, $\frac{\partial v_z}{\partial z}$, получим уравнение типа уравнения неразрывности

$$\frac{\partial^2 \rho}{\partial t^2} + \rho \frac{\partial}{\partial x} \left(\frac{dv_x}{dt} \right) + \rho \frac{\partial}{\partial y} \left(\frac{dv_y}{dt} \right) + \rho \frac{\partial}{\partial z} \left(\frac{dv_z}{dt} \right) = 0.$$

По закону Бойля—Мариотта $p = a^2 \rho$. Это позволит нам из полученных уравнений сначала исключить плотность ρ, а затем и v_x, v_y, v_z. Элементарные преобразования приводят к уравнению

$$\frac{\partial^2 p}{\partial t^2} = a^2 \left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2} \right), \tag{1.34}$$

описывающему малые колебания газовой среды. Уравнение (1.34) представляет собой линейное однородное уравнение в частных производных второго порядка от неизвестной функции $p(x, y, z, t)$ с постоянными коэффициентами. Его часто называют также трехмерным волновым уравнением.

1. 4. 2. Уравнения процессов выравнивания

Уравнение сохранения тепловой энергии. Вывод уравнения сохранения тепловой энергии почти дословно повторяет вывод уравнения неразрывности, который приведен выше. Не останавливаясь на деталях, обращаем ваше внимание на ряд специфических моментов, присущих процессу теплопередачи.

Как и всякое физическое явление процесс теплопередачи происходит в пространстве и во времени. Поэтому исследование теплопроводности сводится к изучению изменения основной физической величины — температуры, характерной для данного явления

$$T = T(x, y, z, t).$$

Зная эту функцию, мы сможем найти распределение температуры в поле в любой момент времени t, определить температуру в любой точке поля.

Опыт показывает, что передача тепла происходит от мест с большей температурой к местам с меньшей температурой.

Пусть мы имеем некоторую среду, в которой выделим объем V, ограниченный поверхностью S. Попробуем выразить в виде дифференциального уравнения закон сохранения тепловой энергии в этом объеме. Для этого в первую очередь нужно подсчитать количество тепла, находящегося в выделенном объеме. Известно, что объемная плотность Q тепловой энергии выражается соотношением

$$Q = c_v \rho T,$$

где ρ — плотность среды, c_v — удельная теплоемкость при постоянном объеме.

Тогда тепловая энергия, аккумулированная в нашем объеме, будет выражаться интегралом

$$\iiint_V \rho c_v T(x, y, z, t)\, dV.$$

Но тепловая энергия, содержащаяся в объеме V, не будет постоянной, она будет изменяться за счет движения частиц от мест более нагретых к менее нагретым. Поэтому скорость изменения тепловой энергии можно получить, дифференцируя по времени полученный выше интеграл

$$\iiint\limits_V \rho c_v \frac{\partial T}{\partial t} dV. \tag{1.35}$$

Скорость изменения тепла, содержащегося в объеме V, можно подсчитать и другим способом. В теории теплопроводности одной из основных величин является вектор $\vec{q}(x, y, z, t)$, названный *вектором теплового потока*. Он характеризует количество тепла, проходящего в единицу времени и отнесенного к единице площади некоторой поверхности S, помещенной в поле. Поэтому скорость изменения тепловой энергии, заключенной внутри V, равна потоку вектора $\vec{q}(x, y, z, t)$ через поверхность S, ограничивающую этот объем, т. е.

$$-\iint\limits_S q_n dS,$$

где q_n — проекция вектора теплового потока $\vec{q}(x, y, z, t)$ на направление внешней нормали $\vec{n}$ поверхности S.

Итак, закон сохранения тепловой энергии в интегральной форме может быть записан в виде равенства

$$\iiint\limits_V \rho c_v \frac{\partial T}{\partial t} dV = -\iint\limits_S q_n dS. \tag{1.36}$$

Желая его получить в дифференциальной форме, преобразуем поверхностный интеграл правой части равенства (1.36) в объемный по формуле Остроградского

$$\iint\limits_S q_n dS = \iiint\limits_V \operatorname{div} \vec{q} dV,$$

где

$$\operatorname{div} \vec{q} = \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} + \frac{\partial q_z}{\partial z}.$$

Отсюда с учетом равенства (1.36) получим

$$\iiint\limits_V \rho c_v \frac{\partial T}{\partial t} dV = -\iiint\limits_V \operatorname{div} \vec{q} \, dV$$

или

$$\iiint\limits_V \left[\rho c_v \frac{\partial T}{\partial t} + \operatorname{div} \vec{q}\right] dV = 0.$$

Окончательно,

$$\rho c_v \frac{\partial T}{\partial t} + \operatorname{div} \vec{q} = 0.$$

В скалярной форме закон сохранения тепловой энергии можно выразить дифференциальным уравнением

$$\rho c_v \frac{\partial T}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} + \frac{\partial q_z}{\partial z} = 0.$$

Оно получено в предположении, что внутри V нет источников тепла. Если же в рассматриваемой среде происходит тепловыделение, то уравнение должно быть изменено. Обозначим через $\varphi(x, y, z, t)$ количество тепловой энергии, выделенное за единицу времени в единице объема среды. (Функцию $\varphi(x, y, z, t)$ называют плотностью тепловыделения.) Тогда уравнение сохранения тепловой энергии примет следующий вид

$$\rho c_v \frac{\partial T}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} + \frac{\partial q_z}{\partial z} = \varphi, \qquad (1.37)$$

т. е. станет неоднородным дифференциальным уравнением.

Уравнение теплопроводности. Уравнение сохранения тепловой энергии (1.37) позволит нам составить дифференциальное уравнение в частных производных относительно искомой функции $T(x, y, z, t)$, определяющей распределение температуры среды.

Введем в рассмотрение вектор — градиент температуры

$$\operatorname{grad} T = \frac{\partial T}{\partial x}\vec{i} + \frac{\partial T}{\partial y}\vec{j} + \frac{\partial T}{\partial z}\vec{k}.$$

Как известно (см. ч. 2, стр. 19), $\operatorname{grad} T$ направлен в сторону возрастания функции $T(x, y, z, t)$. Если учесть, что тепло течет в направлении от нагретых частей тела к холодным, то вектор $\operatorname{grad} T$ прямо противоположен вектору теплового потока $\vec{q}$. Какая же существует зависимость между этими двумя векторами? Она дается основным законом теплопроводности, который может быть сформулирован так: плотность теплового потока прямо пропорциональна градиенту температуры, т. е.

$$\vec{q} = -k \operatorname{grad} T \qquad \left(\vec{q} = -k\frac{\partial T}{\partial n}\vec{n}^0\right), \qquad (1.38)$$

где k — коэффициент теплопроводности.

В скалярной форме это равенство эквивалентно трем уравнениям

$$q_x = -k\frac{\partial T}{\partial x};\ q_y = -k\frac{\partial T}{\partial y};\ q_z = -k\frac{\partial T}{\partial z}. \tag{1.39}$$

Решив полученные уравнения (1.39) совместно с уравнением сохранения тепловой энергии (1.37) и исключив q_x, q_y, q_z, получим

$$\rho c_v\frac{\partial T}{\partial t} = k\left[\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right] + \varphi(x, y, z, t). \tag{1.40}$$

Уравнение (1.40) представляет собой линейное неоднородное уравнение в частных производных второго порядка с постоянными коэффициентами. Оно носит название трехмерного уравнения теплопроводности.

В задачах на теплопередачу встречаются также одномерные и двумерные уравнения теплопроводности.

1. 4. 3. Уравнения установившихся процессов

Нам осталось познакомиться с одним типом линейного уравнения в частных производных второго порядка. Оно лежит в основе теории потенциала и используется для описания установившихся условий фактически во всех основных разделах физики.

а) Рассмотрим установившееся безвихревое движение жидкости. Как известно (см. ч. 2, стр. 56), в этом случае векторное поле скоростей $\vec{v}(x, y, z)$ движения частиц жидкости является потенциальным, т. е. существует такая скалярная функция $u(x, y, z)$, называемая потенциалом скоростей, что

$$\vec{v} = -\operatorname{grad} u$$

или

$$-\vec{v} = \frac{\partial u}{\partial x}\vec{i} + \frac{\partial u}{\partial y}\vec{j} + \frac{\partial u}{\partial z}\vec{k}. \tag{1.41}$$

Для несжимаемой жидкости плотность ρ есть величина постоянная. Поэтому уравнение неразрывности (1.26) в рассматриваемом случае будет иметь вид

$$\rho\left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) = 0,$$

ибо $\frac{\partial \rho}{\partial t} = 0$.

Присоединив уравнения потенциального движения из (1.41)

$$v_x = -\frac{\partial u}{\partial x}; \quad v_y = -\frac{\partial u}{\partial y}; \quad v_z = -\frac{\partial u}{\partial z}$$

и исключив v_x, v_y, v_z, получим уравнение для определения потенциала $u(x, y, z)$ скоростей движения несжимаемой жидкости

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0. \tag{1.42}$$

Если потенциал $u(x, y, z)$ будет найден, то скорость $\vec{v}$ определится из равенства (1.41).

Уравнение (1.42) есть линейное однородное уравнение в частных производных второго порядка с постоянными коэффициентами. Оно называется *однородным уравнением Лапласа.*

б) К уравнению Лапласа приводит и задача о стационарном (установившемся) тепловом поле. Как известно температура нестационарного теплового поля удовлетворяет уравнению теплопроводности

$$\rho c_v \frac{\partial T}{\partial t} = k\left[\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right] + \varphi.$$

Если процесс установился, т. е. распределение температуры не меняется с течением времени, то искомая функция T от времени уже не зависит, т. е. $T = T(x, y, z)$. А тогда $\frac{\partial T}{\partial t} = 0$ и уравнение теплопроводности переходит в неоднородное уравнение Лапласа

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = f(x, y, z), \tag{1.43}$$

где $f = \frac{\varphi}{k}$.

Неоднородное уравнение Лапласа называют также уравнением Пуассона.

Таким образом, при решении ряда физических задач мы пришли к трем типам линейных уравнений в частных производных: волновому уравнению, уравнению теплопроводности и уравнению Лапласа. Каждая из этих задач имеет свои специфические особенности, что сказывается на форме характеризующих их уравнений. Известно, что уравнения для большинства задач теории поля, интересные для физики и техники, имеют форму или волнового уравнения, или уравнения теплопроводности, или уравнения Лапласа. Поэтому изучению этих уравнений и будут посвящены последующие параграфы.

§ 1. 5. НАЧАЛЬНЫЕ И КРАЕВЫЕ УСЛОВИЯ. О КОРРЕКТНОСТИ ПОСТАНОВКИ ЗАДАЧ МАТЕМАТИЧЕСКОЙ ФИЗИКИ

В предыдущем параграфе мы довольно подробно знакомились с разнообразными задачами физики и механики, которые приводят к уравнениям в частных производных. Можно было бы привести еще много примеров подобного рода. Но основная наша цель состоит не в выводе уравнений в частных производных, а в исследовании и решении их.

Прежде чем говорить о методах решений уравнений в частных производных еще раз обратимся к постановке задач математической физики, остановимся на целом ряде специфических черт таких задач, носящих для дальнейшего принципиальный характер.

С примерами постановки простейших задач для обыкновенных дифференциальных уравнений мы уже встречались. В силу физических особенностей системы, описываемой дифференциальным уравнением, отыскивалось то его решение, которое и интересовало исследователя. Но для определения такого решения недостаточно было решить полученное уравнение, т. е. найти его общее решение. Необходимо было из всего множества решений выбрать нужное. Так как само уравнение никакой информации об этом «нужном» решении не содержит, то исследователь должен располагать некоторой дополнительной информацией об этом решении. Эта дополнительная информация в силу физической (технической) задачи как правило имеет вид краевых условий, когда задаются значения искомой функции и ее производных или при одном значении, или при разных значениях аргумента в области, где и рассматривается исследуемый процесс. Неправильно было бы думать, что дополнительные условия, присоединяемые к дифференциальному уравнению, могут иметь такой диапазон произвола, что о их характере не нужно заботиться. Ведь эти условия можно задать такими, что уравнение вообще решений иметь не будет! С другой стороны, иногда краевые условия носят такой характер, что задача имеет не одно, а несколько решений, что тоже нас устроить не может: какое из полученных решений описывает исследуемый процесс остается неизвестным. Поэтому каждое дифференциальное уравнение с присоединенными к нему дополнительными условиями, накладываемыми на искомое решение (краевые условия), составляет самостоятельную задачу, от которой требуют существования и единственности решения. Условия, обеспечивающие существование и единственность решения той или иной задачи, обычно даются так называемыми теоремами существования и единственности решения дифференциального уравнения. С некоторыми из них мы встречались в общем курсе высшей математики и в ч. 2 пособия. Вопрос существования и единственности решения краевой задачи в каждом конкретном случае решают по-своему. Но можно выделить такие краевые задачи, которые наиболее часто встречаются в приложениях и поэтому в первую очередь привлекают к себе внимание. В теории обыкновенных дифференциальных уравнений одной из таких задач является задача Коши. При весьма общих условиях, накладываемых на уравнение, легко доказать, что задача Коши имеет вполне определенное единственное решение. Наиболее часто встречающийся способ отыскания этого решения, например, для линейного уравнения n-го порядка состоит в том, что сначала находят общее решение этого уравнения, которое, как известно, зависит от n произвольных постоянных, а затем в силу начальных условий определяют сами произвольные постоянные, сводя эту задачу к решению системы линейных алгебраических уравнений.

Все сказанное выше в большей части относится к дифференциальным уравнениям в частных производных. Пожалуй самое яркое отличие здесь заключается в том, что задачи, приводящие к дифференциальным уравнениям в частных производных, почти нигде не встречаются в такой форме, чтобы искомым являлось общее решение. Мало того, если вы располагаете таким решением, то знание его, за редким исключением, почти ничего не дает для решения основной частной задачи: ведь общее решение уравнения в частных производных n-го порядка зависит от n произвольных функций и для нахождения частного решения придется рассматривать систему столь сложных функциональных соотношений, что отыскание произвольных функций окажется практически невозможным.

В связи с этим каждая задача математической физики ставится как задача решения некоторого уравнения в частных производных при определенных дополнительных условиях. Эти дополнительные условия в большинстве случаев диктуются физическими свойствами, характером системы, описываемой уравнением. Очевидно, дополнительные условия, о которых здесь идет речь, например, для волнового уравнения (1.34) будут отличаться, вообще говоря, от дополнительных условий для уравнения теплопроводности (1.40) и уравнения Лапласа (1.43), точно так же как дополнительные условия для последнего уравнения будут отличаться от дополнительных условий для уравнения теплопроводности. Как известно, волновое уравнение и уравнение теплопроводности описывают так называемые нестационарные явления, а уравнение Лапласа описывает установившиеся стационарные процессы. Это обстоятельство и отличает в корне дополнительные условия для первых двух уравнений от дополнительных условий для третьего. Действительно, если искомая функция, описывающая нестационарный процесс, зависит от времени, то естественно такую задачу назвать задачей с начальными условиями. Ведь для нахождения вполне определенной функции, входящей как неизвестное в уравнение, необходимо знать величины, характеризующие ее в некоторый момент времени, обычно называемый «начальным» моментом, а также все функции возмущений (внешние силы) во все последующие моменты времени после начального. Приведем пример из механики, о котором шла речь в § 1.1 (движение пружины). Если бы сила F, действующая на пружину (см. рис. 1.2), менялась со временем, то необходимо было бы знать смещения и скорости всех точек пружины в некоторый определенный момент времени для того, чтобы найти зависящие от времени деформации внутри пружины для всех последующих моментов времени. Кроме того, требовалось бы знать все силы, действующие на границе пружины. В стационарных задачах дело обстоит иначе. Так как возмущение (внешние силы) и искомая функция во времени не изменяются, то для анализа системы достаточно определения их величин и распределения внутри поля в некоторой области. А для этого нужно знать как ведет себя неизвестная функция на границе этой области. Поэтому задачи стационарного поля называют часто граничными или краевыми задачами.

Укажем на некоторые из возможных постановок задач для уравнений, которые мы вывели в § 1.4. Причем по аналогии с обыкновенными дифференциальными уравнениями эти задачи будем называть краевыми задачами.

1. 5. 1. Краевые задачи для волнового уравнения

Рассмотрим сначала простейшее волновое одномерное уравнение

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \tag{1.44}$$

описывающее колебания струны длиной l, т. е. $0 \leqslant x \leqslant l$. В левой части этого уравнения стоит величина $\frac{\partial^2 u}{\partial t^2}$ — вторая частная производная по времени от искомой функции $u(x, t)$, выражающая ускорение произвольной точки струны. Известно, что движение любой механической системы, в которой действующие силы, а следовательно, и ускорения выражаются через координаты участвующих в движении тел, вполне определяется, если заданы начальные положения и скорости всех точек системы. В связи с этим движение струны будет вполне определено, если задать в начальный момент времени $t=0$ положения

$$u(x, t)|_{t=0} = f(x)$$

и скорости

$$\left.\frac{\partial u(x, t)}{\partial t}\right|_{t=0} = \varphi(x) \tag{1.45}$$

всех точек струны. (Функции $f(x)$ и $\varphi(x)$ предполагаются известными.)

Условия (1.45) показывают, в каком состоянии находится струна в момент начала колебания. Поэтому эти условия часто называются начальными условиями.

Очевидно, движение струны существенно зависит от того, как она себя ведет на концах, т. е. в точках $x=0$ и $x=l$. Ведь на концах струны перестают действовать те формулы, которые выражали законы механики для внутренних точек. В связи с этим придется еще задать дополнительные условия, заранее обеспечивающие режим струны в точках $x=0$ и $x=l$.

Если предположить, что в положении равновесия струна закреплена на обоих концах, то эти дополнительные условия выразятся равенствами

$$u(x, t)|_{x=0} = 0; \quad u(x, t)|_{x=l} = 0. \tag{1.46}$$

В связи с тем, что условия (1.46) относятся к границе области, где должна быть определена функция $u(x, t)$, они носят название граничных или краевых условий.

Итак, задачу, к которой приводит изучение свободных колебаний струны, закрепленной на обоих концах, можно окончательно сформулировать следующим образом.

Найти решение $u(x, t)$ однородного дифференциального уравнения в частных производных второго порядка с постоянными коэффициентами

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2},$$

удовлетворяющее начальным условиям

$$u(x, t)\Big|_{t=0} = f(x); \quad \left.\frac{\partial u(x, t)}{\partial t}\right|_{t=0} = \varphi(x)$$

и краевым условиям

$$u(x, t)|_{x=0} = 0; \quad u(x, t)|_{x=l} = 0.$$

Очевидно, можно ставить и другие задачи для уравнения (1.44), когда граничные условия имеют вид, отличный от (1.46). Например, можно задать законы колебания концов струны: в момент $t \geqslant 0$ смещение левого конца струны равно $\psi(t)$, а смещение правого конца — $\theta(t)$. Тогда краевая задача будет ставиться так.

Найти решение $u(x, t)$ уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2},$$

удовлетворяющее начальным условиям

$$u(x, t)\Big|_{t=0} = f(x); \quad \frac{\partial u(x, t)}{\partial t}\Big|_{t=0} = \varphi(x)$$

и краевым условиям

$$u(x, t)\,|_{x=0} = \psi(t); \quad u(x, t)_{x=l} = \theta(t). \tag{1.47}$$

Возможно задание также и других условий на концах.

Если струна закреплена на обоих концах, т. е. искомое решение удовлетворяет краевым условиям (1.46), то соответствующую задачу называют краевой задачей с однородными краевыми условиями. Если же краевые условия имеют вид (1.47) ($\psi(t) \not\equiv 0$; $\theta(t) \not\equiv 0$), то соответствующую задачу называют краевой задачей с неоднородными краевыми условиями.

Сразу отметим, что уравнение (1.44) имеет единственное решение, удовлетворяющее краевым условиям (1.47) и начальным условиям (1.45). Это означает, что уравнение (1.44) совместно с дополнительными условиями (1.45), (1.47) содержит всю информацию, необходимую для исследования колебания струны (решение единственно), и не содержит избыточной, противоречивой информации (решение существует).

Может показаться, что условия (1.45) и (1.47) независимы друг от друга. На самом деле это не так. Например, если потребовать, чтобы искомая функция $u(x, t)$ была непрерывной не только внутри, но и на границе области своего определения, необходимо, чтобы

$$f(0) = \psi(0); \quad f(l) = \theta(0), \tag{1.48}$$

ибо в противном случае $u(x, 0) \neq u(0, t)$ и $u(x, 0) \neq u(l, t)$. Соотношения (1.48) называются условиями согласования. В случае однородных краевых условий (1.46) они имеют особенно простой вид

$$f(0) = f(l) = 0.$$

Если потребовать, чтобы на границе области были непрерывны также некоторые из производных функции $u(x, t)$, то могут, естественно, возникнуть новые условия согласования, на чем мы останавливаться не будем.

Аналогично разобранным ставятся краевые задачи и для трехмерного (1.34) волнового уравнения. Рассмотрим одну из таких задач.

Найти решение $p(x, y, z, t)$ уравнения

$$\frac{\partial^2 p}{\partial t^2} = a^2\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial y^2} + \frac{\partial^2 p}{\partial z^2}\right),$$

удовлетворяющее начальным условиям

$$p(x, y, z, t)\Big|_{t=0} = F(x, y, z); \quad \frac{\partial p(x, y, z, t)}{\partial t}\Big|_{t=0} = \Phi(x, y, z)$$

и краевому условию

$$p(x, y, z, t)|_S = \psi(x, y, z, t) = \psi(P, t), \tag{1.49}$$

где F, Φ и ψ — заданные функции, причем ψ зависит от точки P на поверхности S и времени t.

Краевое условие (1.49) выражает физический закон для частиц, находящихся на границе S рассматриваемого объема V. Этот закон может быть задан и в другой форме, например, в виде задания производной $\frac{\partial p}{\partial n}$ на границе объема, или линейной комбинации искомой функции p и производной $\frac{\partial p}{\partial n}$ на этой границе. Во всяком случае, при весьма общих ограничениях можно доказать, что сформулированная краевая задача имеет вполне определенное единственное решение.

Может представиться и такой случай, когда исследователя интересует явление в течение такого промежутка времени, когда влияние границ еще несущественно или когда струна (тело) имеет неограниченные размеры и влиянием концов (границ) можно пренебречь. Тогда вместо полной краевой задачи (ее иногда называют смешанной) рассматривается краевая задача только с начальными условиями для неограниченной области. Рассмотрим пример.

Найти решение $u(x, t)$ уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < +\infty, \quad t > 0,$$

удовлетворяющее начальным условиям

$$u(x, t)\Big|_{t=0} = f(x); \quad \frac{\partial u(x, t)}{\partial t}\Big|_{t=0} = \varphi(x),$$

где $f(x)$ и $\varphi(x)$ — заданные функции при $-\infty < x < +\infty$.

Такую краевую задачу обычно называют задачей Коши. Как и в общем случае она имеет вполне определенное единственное решение, в чем мы сможем убедиться несколько позже.

1. 5. 2. Краевые задачи для уравнения теплопроводности

Рассмотрим однородное уравнение теплопроводности

$$\frac{\partial T}{\partial t} = a^2\left(\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2}\right), \tag{1.50}$$

где $a^2 = \dfrac{k}{\rho c_v}$.

Как известно (см. § 1.4), уравнение (1.50) описывает распределение температуры среды во времени.

Точно так же, как в предыдущем пункте, для выделения единственного решения уравнения (1.50) необходимо к этому уравнению присоединить начальные и краевые условия.

В левой части уравнения (1.50) стоит величина $\dfrac{\partial T}{\partial t}$ — первая производная по времени от искомой функции $T(x, y, z, t)$ — температуры среды. В сравнении с волновым уравнением это обстоятельство скажется на характере начальных условий для рассматриваемой задачи. Физические соображения подтверждают эту мысль. Действительно, чтобы найти температуру внутри некоторого объема в любой момент времени, совершенно необходимо знать распределение температуры внутри этого объема в начальный момент времени $t = 0$. Так появляется начальное условие для уравнения (1.50)

$$T(x, y, z, t)|_{t=0} = f(x, y, z), \tag{1.51}$$

где f — заданная функция.

Безусловно, на распределении температуры внутри некоторого объема будет сказываться и тот тепловой режим, который имеется на границе S объема. Так появляются краевые (граничные) условия для уравнения (1.50). Они могут задаваться различными способами.

а) В каждой точке поверхности S задается температура

$$T(x, y, z, t)|_S = \varphi(P, t), \tag{1.52}$$

где $\varphi(P, t)$ — известная функция точки P поверхности S и времени $t \geqslant 0$.

Если предположить, что температура внешней среды равна $\varphi(x, y, z, t)$, то условие (1.52) означает, что на границе S объем имеет температуру внешней среды.

б) На поверхности S задается тепловой поток (1.38)

$$q = -k\frac{\partial T}{\partial n}.$$

Это означает, что

$$\left.\frac{\partial T}{\partial n}\right|_S = \psi(P, t), \tag{1.53}$$

где $\psi(P, t)$ — известная функция точки P поверхности S и времени $t \geqslant 0$ $\left(\psi(P, t) = -\frac{q(P, t)}{k}\right)$. Может оказаться, что $\psi(P, t) \equiv 0$. Это означает, что переход тепла через границу S объема исключен, т. е. граница объема теплоизолирована.

в) На поверхности S объема происходит теплообмен с окружающей средой, температура которой $\tilde{T}(x, y, z, t)$ известна. Найдем закон этого теплообмена.

Разность $T(P, t) - \tilde{T}(P, t)$ между температурой тела в точке границы и температурой окружающей среды в этой точке называется перепадом температур в точке P границы. Существует физический закон (закон Ньютона), устанавливающий, что поток тепла изнутри объема через любую часть поверхности пропорционален перепаду температур на этой части границы:

$$q = h(T - \tilde{T}),$$

где h — коэффициент теплообмена, зависящий от физических свойств среды и характера поверхности S. Для определенности этот коэффициент будем считать постоянным. С учетом (1.38)

$$h(T - \tilde{T}) = -k\frac{\partial T}{\partial n}$$

или

$$\frac{\partial T}{\partial n} + H(T - \tilde{T})|_S = 0, \tag{1.54}$$

где $H = \frac{h}{k}$.

Итак, если на поверхности S происходит теплообмен с окружающей средой, температура которой $\tilde{T}(x, y, z, t)$, то температура $T(x, y, z, t)$ изучаемого объема должна удовлетворять краевому условию (1.54).

Теперь мы можем сформулировать краевую задачу для уравнения теплопроводности.

Найти решение $T(x, y, z, t)$ уравнения (1.50), удовлетворяющее начальному условию

$$T(x, y, z, t)|_{x=0} = f(x, y, z)$$

и краевому условию

$$T(x, y, z, t)|_S = \varphi(P, t)$$

или одному из краевых условий (1.53) или (1.54).

Заметим, что каждая из трех упомянутых здесь краевых задач для уравнения теплопроводности при некоторых ограничениях имеет вполне определенное единственное решение.

Краевые условия (1.52) — (1.54), очевидно, отпадают, если рассматривать краевую задачу для неограниченного объема (тела). Сформулированная выше задача без таких условий носит название, как и в случае волнового уравнения в подобной ситуации, задачи Коши. Только теперь начальных условий будет уже не два, а одно (1.51).

1. 5. 3. Краевые задачи для уравнения Лапласа

Рассмотрение краевых задач для уравнения Лапласа начнем со случая стационарного распределения температуры внутри некоторого объема (тела) V.

Пусть в некотором объеме V процесс теплопередачи установился, т. е. распределение температуры не меняется с течением времени. Тогда функция температуры T от времени не зависит и, как известно, удовлетворяет уравнению Лапласа

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = 0.$$

Очевидно, говорить здесь о начальных условиях не имеет смысла: ведь распределение температуры в начальный и во все последующие моменты времени одно и то же и совпадает с искомой функцией $T(x, y, z)$.

Относительно граничных (краевых) условий можно сказать слово в слово то же, что и в предыдущем пункте, когда речь шла о нестационарном распределении температуры. Нам остается перечислить возможные здесь постановки краевых задач и привести их название.

Задача Дирихле. Найти решение $T(x, y, z)$ уравнения

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = 0,$$

удовлетворяющее краевому условию

$$T(x, y, z)|_S = f(P),$$

где f — заданная функция точки P поверхности S, ограничивающей объем V.

Задача Дирихле при слабых ограничениях имеет одно и только одно решение. Одно из этих ограничений бросается в глаза. Действительно, если речь идет о распределении температуры внутри некоторого объема (тела) V, то необходимо потребовать, чтобы функция $T(x, y, z)$ была непрерывной в области V вплоть до ее границы S, т. е. была непрерывной в замкнутой области. Если отказаться от этого требования, то любую функцию, равную постоянной C внутри V и функции $f(x, y, z)$ на S, можно было бы рассматривать как решение нашей задачи. В этом случае единственность решения задачи Дирихле была бы потеряна.

Если обратиться к примеру установившегося движения несжимаемой жидкости, то для потенциала $u(x, y, z)$ скоростей $\vec{v}$ частиц этой жидкости мы придем к уравнению Лапласа:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0,$$

для которого ставить задачу Дирихле уже не имеет смысла из физических соображений: определить потенциал скорости на границе S области V не представляется возможным. Но мы можем измерить саму скорость $\vec{v}$ на границе S области V или проекцию скорости $\vec{v}$ на направление внешней нормали $\vec{n}$ к поверхности S. Останавливаясь на втором случае, заметим, что так как $\vec{v} = -\operatorname{grad} u$, то $\text{пр}_{\vec{n}}\, \vec{v} = -\text{пр}_{\vec{n}} \operatorname{grad} u = -\frac{\partial u}{\partial n}$ (см. ч. 2, стр. 19). А если будет известна проекция скорости $\vec{v}$ на направление внешней нормали $\vec{n}$ к поверхности S, то будет известна и производная от потенциала u по направлению нормали в любой точке поверхности, т. е.

$$\left.\frac{\partial u}{\partial n}\right|_S = \varphi(x, y, z),$$

где $\varphi(x, y, z)$ — известная функция, определенная на поверхности S. Итак, мы приходим к следующей краевой задаче для уравнения Лапласа.

Задача Неймана. Найти решение $u(x, y, z)$ уравнения

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0,$$

удовлетворяющее краевому условию

$$\left.\frac{\partial u(x, y, z)}{\partial n}\right|_S = \varphi(P),$$

где φ — заданная функция точки P поверхности S, ограничивающей объем V.

Задача Неймана имеет чрезвычайно важную особенность. Если предположить, что функция $\varphi(x, y, z)$ совершенно произвольна на границе S области, то эта задача решения может вовсе не иметь. В дальнейшем мы выясним, что необходимое условие существования решения задачи Неймана состоит в выполнении равенства

$$\iint_S \varphi(x, y, z)\, dS = 0. \tag{1.55}$$

Таким образом, совершенно теряет смысл постановка задачи Неймана, если условие (1.55) не выполняется.

Кроме задач Дирихле и Неймана для уравнения Лапласа можно ставить и более общую краевую задачу, когда отыскивается такое его решение, которое удовлетворяет краевому условию типа (1.54).

Как и в предыдущих случаях, совсем необязательно отыскивать решение уравнения Лапласа лишь в конечной области. Часто бывает необходимым решить уравнение Лапласа для области, неограниченной. Это бывает, например, тогда, когда размеры рассматриваемой области очень велики по сравнению с масштабом изучаемого явления.

В случае бесконечных областей важно знать, как ведет себя искомое решение в далеких точках изучаемой области, т. е. на бесконечности. Дело в том, что во многих случаях задача остается неопределенной (нарушается единственность), если не задать условий, которым должна удовлетворять искомая функция на бесконечности. Обычно эти условия заключаются в том, чтобы решение обращалось на бесконечности в нуль.

Заметим, что если отыскивается решение задачи Дирихле или Неймана в области V внутренней (или внешней) по отношению к поверхности S, то соответствующую задачу называют внутренней (или внешней) краевой задачей.

В заключение параграфа сделаем ряд общих замечаний о постановке задач математической физики.

Разобранные выше примеры постановки краевых задач для уравнений математической физики познакомили нас с основными типами таких задач, заставили с большим вниманием отнестись к проблеме интегрирования уравнений в частных производных. Каждый раз подчеркивая физическое содержание задач, мы невольно оставляем в стороне чисто математические вопросы, которые здесь имеют немаловажное значение. Чтобы вас сразу ввести в курс возникающих здесь серьезных проблем, поступим следующим образом. Предположим, что задана краевая задача: выписано уравнение в частных производных и краевые условия. Предлагается решить эту задачу, т. е. найти функцию, удовлетворяющую уравнению и краевым условиям (в краевые условия могут, конечно, входить и начальные условия). Трудно предположить, что решение сформулированной задачи вы начнете с попыток нахождения искомой функции, с подбора метода решения задачи. Ведь может оказаться, что решение не существует. Поэтому нужно постараться в первую очередь убедиться в том, что решение задачи на самом деле существует. Обычно искомую функцию подчиняют некоторым ограничениям общего характера. Эти ограничения, в свою очередь, вынуждают накладывать некоторые ограничения и на заданные функции, входящие в правые части дифференциального уравнения, и краевые условия. Так очерчивается круг условий, при которых и рассматривается наша задача. Теперь остается доказать, что решение заданной краевой задачи существует. Если удается решить этот вопрос положительно, встает следующая проблема — проблема единственности решения. Может оказаться, что при рассматриваемых условиях задача имеет не одно, а два или несколько решений, что нас по известным причинам устроить не может. Тогда такую задачу нельзя считать поставленной правильно, как говорят, поставленной корректно. Чтобы она стала корректной, надо пересмотреть исходные данные, изменить условия постановки задачи. Обычно стараются ставить краевые условия так, чтобы краевая задача имела одно и только одно решение. Весь этот комплекс проблем решается в каждом конкретном случае теоремой существования и единственности решения краевой задачи. Предположим, нам удалось убедиться в том, что рассматриваемая краевая задача имеет единственное решение. Но и этого чаще всего не достаточно, чтобы сделать положительный вывод о «доброкачественности» полученного решения. Дело в том, что в основе определения физических величин в конечном счете лежит процесс измерения, который всегда связан с некоторой погрешностью. Очевидно, с погрешностью определяются и данные задачи, в том числе и краевые условия. Поэтому естественен вопрос: как погрешность в данных краевой задачи отразится на ее решении? Очевидно, если незначительные изменения начальных данных приведут к незначительному изменению решения краевой задачи, то полученное решение можно считать правильно отражающим физический процесс. В противном случае найденное решение нельзя считать удовлетворительным.

Если малые изменения данных задачи влекут за собой малые изменения решения, то говорят, что решение краевой задачи устойчиво (или непрерывно зависит от данных задачи). Требование устойчивости сопровождает почти каждую краевую задачу.

Итак, краевая задача должна быть поставлена так, чтобы:

1) решение существовало,

2) решение было единственным,

3) решение было устойчивым.

Если все эти требования выполняются, то краевую задачу называют поставленной корректно.

Исследование корректности постановки задач математической физики представляет важнейшую и притом весьма трудную задачу теории уравнений в частных производных. Каждый тип задач имеет свои условия, обеспечивающие корректность постановки той или иной краевой задачи. Например, для уравнения Лапласа можно выделить группу условий, гарантирующих корректность постановки соответствующей краевой задачи. Искомая функция, дающая решение краевой задачи для уравнения в частных производных второго порядка, должна удовлетворять следующим условиям:

1) быть непрерывной в области, в которой ставится задача, вплоть до границы области;

2) внутри области иметь непрерывные вторые производные и удовлетворять заданному уравнению;

3) удовлетворять заданному краевому условию;

4) если область трехмерна и бесконечна, то при перемещении к бесконечно удаленной точке вдоль любого луча, принадлежащего области, стремиться к нулю.

В дальнейшем, если не будет оговорено особо, мы всюду будем предполагать, что рассматриваемые нами краевые задачи поставлены корректно.

§ 1. 6. ПРИВЕДЕНИЕ УРАВНЕНИЯ ВТОРОГО ПОРЯДКА К КАНОНИЧЕСКОМУ ВИДУ

В курсе аналитической геометрии изучалась задача о приведении общего уравнения кривой второго порядка к каноническому виду. Схема решения этой задачи состояла в следующем. Предложенное уравнение кривой второго порядка

$$Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F = 0$$

путем линейного преобразования координат

$$\bar{x} = \alpha_1 x + \beta_1 y; \quad \bar{y} = \alpha_2 x + \beta_2 y$$

с определителем

$$\Delta = \begin{vmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{vmatrix} \neq 0$$

(преобразование невырожденное), приводилось к виду

$$A'\bar{x}^2 + 2B'\bar{x}\bar{y} + C'\bar{y}^2 + D'\bar{x} + E'\bar{y} + F = 0.$$

Элементы преобразования α_i, β_i, $i = 1,2$, подбирались таким образом, чтобы коэффициент при произведении координат $\bar{x}$, $\bar{y}$ обратился в нуль, т. е. $B' = 0$. Тогда в новой системе координат $\bar{x}O\bar{y}$ уравнение кривой приобретало следующий вид

$$A'\bar{x}^2 + C'\bar{y}^2 + D'\bar{x} + E'\bar{y} + F = 0,$$

который назывался каноническим видом кривой второго порядка. Причем, если $\delta = B^2 - AC < 0$, то заданная кривая называлась кривой эллиптического типа; если $\delta = B^2 - AC > 0$, то кривая называлась кривой гиперболического типа и, наконец, если $\delta = B^2 - AC = 0$, то кривая называлась кривой параболического типа. Приведение общего уравнения кривой второго порядка к каноническому виду позволяло без больших помех получить ту кривую, которую и описывало заданное уравнение: эллипс, гиперболу или параболу (не исключены, конечно, случаи вырождения).

Представьте теперь себе, что вы имеете дело не с уравнением кривой второго порядка, а с линейным уравнением в частных производных второго порядка:

$$a_{11}\frac{\partial^2 u}{\partial x^2} + 2a_{12}\frac{\partial^2 u}{\partial x\,\partial y} + a_{22}\frac{\partial^2 u}{\partial y^2} + b_1\frac{\partial u}{\partial x} + b_2\frac{\partial u}{\partial y} + cu + f = 0, \tag{1.56}$$

где a_{11}, a_{12}, a_{22}, b_1, b_2, c и f — некоторые функции переменных x и y. Спрашивается, нельзя ли здесь поступить точно так же как в предыдущем случае? Нельзя ли подобрать некоторое преобразование координат так, чтобы при его помощи максимально упростить уравнение (1.56)? Положительный ответ на поставленный вопрос открывал бы хорошие перспективы в деле интегрирования линейного уравнения второго порядка в частных производных. Ведь может оказаться, что преобразованное уравнение будет иметь такой вид, что мы сможем выписать его решение, проинтегрировать. После этого, возвращаясь к старым переменным, найдем решение исходного уравнения. Попробуем реализовать описанный выше план для упрощения уравнения (1.56).

Произведем преобразование в уравнении (1.56) независимых переменных x и y по формулам

$$\xi = \varphi(x, y); \quad \eta = \psi(x, y). \tag{1.57}$$

Здесь ξ и η — новые независимые переменные. Считая функции φ и ψ пока неизвестными, будем требовать, чтобы они были непрерывны вместе с их частными производными первого и второго порядков. Будем предполагать также, что функции (1.57) устанавливают взаимно однозначное соответствие между точками (ξ, η) и (x, y) соответствующих областей. Аналитически это условие выражается тем, что якобиан преобразования

$$\frac{D(\varphi, \psi)}{D(x, y)} = \begin{vmatrix} \dfrac{\partial\varphi}{\partial x} & \dfrac{\partial\varphi}{\partial y} \\ \dfrac{\partial\psi}{\partial x} & \dfrac{\partial\psi}{\partial y} \end{vmatrix} = \frac{\partial\varphi}{\partial x}\frac{\partial\psi}{\partial y} - \frac{\partial\varphi}{\partial y}\frac{\partial\psi}{\partial x}$$

нигде в области рассмотрения не обращается в нуль, т. е.

$$\frac{D(\varphi, \psi)}{D(x, y)} = \frac{\partial\varphi}{\partial x}\frac{\partial\psi}{\partial y} - \frac{\partial\varphi}{\partial y}\frac{\partial\psi}{\partial x} \neq 0.$$

Чтобы выполнить требуемую замену переменных, необходимо выразить частные производные от функции u по x и y, входящие в уравнение (1.56), через производные от u по ξ и η. Используя правила дифференцирования сложной функции и помня о том, что после замены переменных u зависит уже от ξ и η, которые в свою очередь зависят от x и y, получим

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x}; \quad \frac{\partial u}{\partial y} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial y} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial y}. \tag{1.58}$$

Совершенно аналогично вычислим производные второго порядка

$$\left.\begin{aligned}
\frac{\partial^2 u}{\partial x^2} &= \frac{\partial^2 u}{\partial \xi^2}\left(\frac{\partial \xi}{\partial x}\right)^2 + 2\frac{\partial^2 u}{\partial \xi\,\partial \eta}\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial x} + \frac{\partial^2 u}{\partial \eta^2}\left(\frac{\partial \eta}{\partial x}\right)^2 + \\
&\qquad + \frac{\partial u}{\partial \xi}\frac{\partial^2 \xi}{\partial x^2} + \frac{\partial u}{\partial \eta}\frac{\partial^2 \eta}{\partial x^2}; \\
\frac{\partial^2 u}{\partial x\,\partial y} &= \frac{\partial^2 u}{\partial \xi^2}\frac{\partial \xi}{\partial x}\frac{\partial \xi}{\partial y} + \frac{\partial^2 u}{\partial \xi\,\partial \eta}\left(\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial y} + \frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial x}\right) + \\
&\qquad + \frac{\partial^2 u}{\partial \eta^2}\frac{\partial \eta}{\partial x}\frac{\partial \eta}{\partial y} + \frac{\partial u}{\partial \xi}\frac{\partial^2 \xi}{\partial x\,\partial y} + \frac{\partial u}{\partial \eta}\frac{\partial^2 \eta}{\partial x\,\partial y}; \\
\frac{\partial^2 u}{\partial y^2} &= \frac{\partial^2 u}{\partial \xi^2}\left(\frac{\partial \xi}{\partial y}\right)^2 + 2\frac{\partial^2 u}{\partial \xi\,\partial \eta}\frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial y} + \frac{\partial^2 u}{\partial \eta^2}\left(\frac{\partial \eta}{\partial y}\right)^2 + \\
&\qquad + \frac{\partial u}{\partial \xi}\frac{\partial^2 \xi}{\partial y^2} + \frac{\partial u}{\partial \eta}\frac{\partial^2 \eta}{\partial y^2}.
\end{aligned}\right\} \tag{1.59}$$

Обратите внимание на то, что правые части равенств (1.58) и (1.59) линейны относительно частных производных по новым переменным $\frac{\partial u}{\partial \xi}$, $\frac{\partial u}{\partial \eta}$, $\frac{\partial^2 u}{\partial \xi^2}$, $\frac{\partial^2 u}{\partial \eta^2}$ $\frac{\partial^2 u}{\partial \xi \partial \eta}$. Если сейчас подставить найденные соотношения в уравнение (1.56), то следует ожидать, что опять получится линейное уравнение второго порядка, где неизвестной уже будет функция u с неизвестными переменными ξ и η.

После подстановки получим

$$\bar{a}_{11}\frac{\partial^2 u}{\partial \xi^2}+2\bar{a}_{12}\frac{\partial^2 u}{\partial \xi \partial \eta}+\bar{a}_{22}\frac{\partial^2 u}{\partial \eta^2}+\bar{F}\left(\xi,\ \eta,\ u,\ \frac{\partial u}{\partial \xi};\ \frac{\partial u}{\partial \eta}\right)=0, \quad (1.60)$$

где

$$\bar{a}_{11}=a_{11}\left(\frac{\partial \xi}{\partial x}\right)^2+2a_{12}\frac{\partial \xi}{\partial x}\frac{\partial \xi}{\partial y}+a_{22}\left(\frac{\partial \xi}{\partial y}\right)^2;$$

$$\bar{a}_{12}=a_{11}\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial x}+a_{12}\left(\frac{\partial \xi}{\partial x}\frac{\partial \eta}{\partial y}+\frac{\partial \eta}{\partial x}\frac{\partial \xi}{\partial y}\right)+a_{22}\frac{\partial \xi}{\partial y}\frac{\partial \eta}{\partial y};$$

$$\bar{a}_{22}=a_{11}\left(\frac{\partial \eta}{\partial x}\right)^2+2a_{12}\frac{\partial \eta}{\partial x}\frac{\partial \eta}{\partial y}+a_{22}\left(\frac{\partial \eta}{\partial y}\right)^2,$$

а $\bar{F}$ — линейная функция относительно u, $\frac{\partial u}{\partial \xi}$, $\frac{\partial u}{\partial \eta}$.

Казалось бы, что мы ничего не достигли. Действительно, опять получено уравнение того же вида, что и исходное (1.56). Но на самом деле это не так. Коэффициенты $\bar{a}_{11}$, $\bar{a}_{12}$, $\bar{a}_{22}$ преобразованного уравнения (1.60) выражаются через функции $\xi=\varphi(x,\ y)$ и $\eta=\psi(x,\ y)$ и может оказаться так, что хотя бы часть из них обратится в нуль. Если это случится, то наша цель будет достигнута: уравнение (1.56) путем преобразования (1.57) будет приведено к более простому виду. Вот и наметился план дальнейших действий. Попробуем выбрать переменные ξ и η так, чтобы коэффициент $\bar{a}_{11}$ был равен нулю. Тогда получим

$$a_{11}\left(\frac{\partial \xi}{\partial x}\right)^2+2a_{12}\frac{\partial \xi}{\partial x}\frac{\partial \xi}{\partial y}+a_{22}\left(\frac{\partial \xi}{\partial y}\right)^2=0.$$

Другими словами, если $z=\varphi(x,\ y)$ — какое-нибудь частное решение уравнения в частных производных первого порядка

$$a_{11}\left(\frac{\partial z}{\partial x}\right)^2+2a_{12}\frac{\partial z}{\partial x}\frac{\partial z}{\partial y}+a_{22}\left(\frac{\partial z}{\partial y}\right)^2=0, \quad (1.61)$$

то положив $\xi=z=\varphi(x,\ y)$, мы обратим коэффициент $\bar{a}_{11}$ в нуль.

Таким образом, наша задача подбора переменных ξ и η свелась к задаче интегрирования уравнения (1.61). Нижеследующая лемма поможет обойти его непосредственное интегрирование.

Лемма. *Если* $z=\varphi(x, y)$ *является частным решением уравнения (1.61), то соотношение* $\varphi(x, y)=C$ *представляет собой общий интеграл обыкновенного дифференциального уравнения*

$$a_{11}dy^2-2a_{12}dxdy+a_{22}dx^2=0 \tag{1.62}$$

и наоборот.

Проверим, что функция $\varphi(x, y)=C$ есть интеграл уравнения (1.62). Уравнение (1.62) может быть переписано в следующем виде

$$a_{11}\left(\frac{dy}{dx}\right)^2-2a_{12}\frac{dy}{dx}+a_{22}=0.$$

Надо показать, что $\varphi(x, y)=C$ удовлетворяет этому уравнению. По правилу дифференцирования неявно заданной функции получим $\frac{\partial\varphi}{\partial y}\frac{\partial y}{\partial x}+\frac{\partial\varphi}{\partial x}=0$. Отсюда $\frac{dy}{dx}=-\frac{\frac{\partial\varphi}{\partial x}}{\frac{\partial\varphi}{\partial y}}$. Подставим полученное в левую часть исследуемого уравнения. Тогда оно приобретет следующий вид

$$a_{11}\left(-\frac{\frac{\partial\varphi}{\partial x}}{\frac{\partial\varphi}{\partial y}}\right)^2-2a_{12}\left(-\frac{\frac{\partial\varphi}{\partial x}}{\frac{\partial\varphi}{\partial y}}\right)+a_{22}$$

или

$$a_{11}\left(\frac{\partial\varphi}{\partial x}\right)^2+2a_{12}\frac{\partial\varphi}{\partial x}\frac{\partial\varphi}{\partial y}+a_{22}\left(\frac{\partial\varphi}{\partial y}\right)^2=0,$$

так как $\varphi(x, y)$ — решение уравнения (1.61). Итак, функция $\varphi(x, y)=C$ удовлетворяет уравнению (1.62).

Полное доказательство сформулированной леммы мы опускаем [35, стр. 13].

Итак, уравнение (1.62) приобретает ключевое значение для решения нашей задачи. Назовем его *характеристическим уравнением.*

Очевидно, характеристическое уравнение (1.62) распадается на два уравнения, которые получим, разрешая его как квадратное относительно $\frac{dy}{dx}$,

$$\frac{dy}{dx}=\frac{a_{12}+\sqrt{a_{12}^2-a_{11}a_{22}}}{a_{11}} \tag{1.63}$$

и

$$\frac{dy}{dx} = \frac{a_{12} - \sqrt{a_{12}^2 - a_{11}a_{22}}}{a_{11}}. \tag{1.64}$$

Если $\varphi(x, y) = C$ есть общий интеграл уравнения (1.63), то, полагая $\xi = \varphi(x, y)$, мы обращаем в нуль коэффициент при $\frac{\partial^2 u}{\partial \xi^2}$. Если $\psi(x, y) = C$ является общим интегралом уравнения (1.64), независимым от $\varphi(x, y)$, то, полагая $\eta = \psi(x, y)$, мы обратим в нуль также и коэффициент при $\frac{\partial^2 u}{\partial \eta^2}$.

Интегральные кривые характеристического уравнения, т. е. все кривые, входящие в семейства $\varphi(x, y) = C$ и $\psi(x, y) = C$, называются *характеристиками* заданного дифференциального уравнения (1.56). Если $a_{12}^2 - a_{11}a_{22} > 0$, то получим характеристики действительные и различные. В случае $a_{12}^2 - a_{11}a_{22} = 0$ получаются действительные характеристики, совпадающие между собой, а в случае $a_{12}^2 - a_{11}a_{22} < 0$ — различные и комплексные.

Непосредственной проверкой можно установить, что имеет место тождество

$$\bar{a}_{12}^2 - \bar{a}_{11}\bar{a}_{22} = (a_{12}^2 - a_{11}a_{22}) \left[\frac{D(\varphi, \psi)}{D(x, y)}\right]^2,$$

из которого следует, что знак дискриминанта $a_{12}^2 - a_{11}a_{22}$ не меняется при преобразовании переменных. Значит, все множество линейных уравнений в частных производных второго порядка можно разбить на следующие типы, положив в основу классификации (как и в случае общего уравнения кривой второго порядка) знак дискриминанта $a_{12}^2 - a_{11}a_{22}$.

Определение. *Уравнение (1.56) называется в точке M области G уравнением*

гиперболического типа, если в точке M $a_{12}^2 - a_{11}a_{22} > 0$;

эллиптического типа, если в точке M $a_{12}^2 - a_{11}a_{22} < 0$.

параболического типа, если в точке M $a_{12}^2 - a_{11}a_{22} = 0$.

Рассмотрим область G, во всех точках которой уравнение (1.56) имеет один и тот же тип, и разберем каждый из трех случаев в отдельности, возвращаясь к нашей основной задаче — задаче упрощения этого уравнения, приведения его к канонической форме. Из предыдущего следует, что тип уравнения не меняется при невырожденном преобразовании независимых переменных и для каждого типа уравнений существует своя каноническая форма.

1. 6. 1. Уравнение гиперболического типа

Правые части уравнений (1.63) и (1.64) действительны и различны. Общие интегралы их $\varphi(x, y) = C$ и $\psi(x, y) = C$ определяют действительные семейства характеристик. Положим

$$\xi = \varphi(x, y); \quad \eta = \psi(x, y).$$

Тогда уравнение (1.60) примет вид

$$2\bar{a}_{12}\frac{\partial^2 u}{\partial\xi\,\partial\eta} + \bar{F} = 0.$$

После деления на $2\bar{a}_{12}$ окончательно получим

$$\frac{\partial^2 u}{\partial\xi\,\partial\eta} = \Phi\left(\xi, \eta, u, \frac{\partial u}{\partial\xi}, \frac{\partial u}{\partial\eta}\right), \tag{1.65}$$

где $\Phi = -\dfrac{\bar{F}}{2\bar{a}_{12}}$.

В связи с тем, что полученное уравнение (1.65) имеет простейшую форму, его называют *канонической формой* уравнения (1.56).

Иногда пользуются другой канонической формой уравнения гиперболического типа. Получим ее.

Сделаем в уравнении (1.65) замену переменных по закону

$$\xi = t + \tau; \quad \eta = t - \tau,$$

где t и τ — новые переменные.

В результате этого преобразования уравнение (1.65) примет вид

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial\tau^2} = \bar{\Phi}, \tag{1.66}$$

где $\bar{\Phi} = 4\Phi$. (Проверьте!)

Пример 1. Привести к каноническому виду уравнение

$$x^2\frac{\partial^2 u}{\partial x^2} - y^2\frac{\partial^2 u}{\partial y^2} = 0, \quad x > 0; \quad y > 0, \tag{1.67}$$

и найти его решение.

Решение. Определим сначала тип предложенного уравнения. Здесь $a_{11} = x^2$; $a_{12} = 0$; $a_{22} = -y^2$. Тогда $a_{12}^2 - a_{11}a_{12} = x^2y^2 > 0$. Значит уравнение гиперболического типа.

Составляем характеристическое уравнение

$$x^2dy^2 - y^2dx^2 = 0.$$

Оно распадается на два уравнения с разделяющимися переменными

$$xdy - ydx = 0; \quad xdy + ydx = 0.$$

Интегрируя эти уравнения, получим

$$xy = C_1; \quad \frac{y}{x} = C_2.$$

Таким образом, прямые $y = C_2 x$ и гиперболы $y = \frac{C_1}{x}$ являются характеристиками заданного уравнения.

Согласно общей теории нужно ввести новые переменные по формулам

$$\xi = xy; \quad \eta = \frac{y}{x}.$$

Используя формулы (1.59), найдем $\frac{\partial^2 u}{\partial x^2}$ и $\frac{\partial^2 u}{\partial y^2}$:

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial \xi^2}\left(\frac{\partial \xi}{\partial x}\right)^2 + 2\frac{\partial^2 u}{\partial \xi\,\partial \eta}\,\frac{\partial \xi}{\partial x}\,\frac{\partial \eta}{\partial x} + \frac{\partial^2 u}{\partial \eta^2}\left(\frac{\partial \eta}{\partial x}\right)^2 + \frac{\partial u}{\partial \xi}\,\frac{\partial^2 \xi}{\partial x^2} +$$
$$+ \frac{\partial u}{\partial \eta}\,\frac{\partial^2 \eta}{\partial x^2} = y^2\frac{\partial^2 u}{\partial \xi^2} - 2\frac{y^2}{x^2}\,\frac{\partial^2 u}{\partial \xi\,\partial \eta} + \frac{y^2}{x^4}\,\frac{\partial^2 u}{\partial \eta^2} + 2\frac{y}{x^3}\,\frac{\partial u}{\partial \eta};$$
$$\frac{\partial^2 u}{\partial y^2} = x^2\frac{\partial^2 u}{\partial \xi^2} + 2\frac{\partial^2 u}{\partial \xi\,\partial \eta} + \frac{1}{x^2}\,\frac{\partial^2 u}{\partial \eta^2}.$$

Подставляя эти значения вторых производных в уравнение (1.67), окончательно получим

$$\frac{\partial^2 u}{\partial \xi\,\partial \eta} - \frac{1}{2\xi}\,\frac{\partial u}{\partial \eta} = 0, \quad \xi > 0, \quad \eta > 0.$$

Последнее уравнение легко интегрируется (см. пример 6 § 1.2). Его общее решение имеет вид

$$u = \sqrt{\xi}\,\theta(\eta) + \lambda(\xi),$$

где θ и λ — произвольные функции. Возвращаясь к старым переменным x и y, получим

$$u = \sqrt{xy}\,\theta\left(\frac{y}{x}\right) + \lambda(x\,y).$$

1. 6. 2. Уравнение параболического типа

В этом случае уравнения (1.63) и (1.64) совпадают и мы получаем один общий интеграл уравнения (1.63) $\varphi(x, y) = C$, определяющий одно семейство характеристик.

В этом случае можно положить

$$\xi = \varphi(x, y); \quad \eta = \psi(x, y),$$

где $\psi(x, y)$ — любая функция, независимая от φ, лишь бы она была дифференцируема нужное число раз.

Очевидно, при выбранной замене переменных коэффициент $\bar{a}_{11}$ обращается в нуль, т. е.

$$\bar{a}_{11} = a_{11}\left(\frac{\partial \xi}{\partial x}\right)^2 + 2a_{12}\frac{\partial \xi}{\partial x}\,\frac{\partial \xi}{\partial y} + a_{22}\left(\frac{\partial \xi}{\partial y}\right)^2 = 0.$$

С учетом того, что $a_{12}^2 = a_{11}a_{22}$ или $a_{12} = \sqrt{a_{11}}\sqrt{a_{22}}$, последнее равенство можно переписать в следующем виде

$$a_{11}\left(\frac{\partial\xi}{\partial x}\right)^2 + 2a_{12}\frac{\partial\xi}{\partial x}\frac{\partial\xi}{\partial y} + a_{22}\left(\frac{\partial\xi}{\partial y}\right)^2 = \left(\sqrt{a_{11}}\frac{\partial\xi}{\partial x} + \sqrt{a_{22}}\frac{\partial\xi}{\partial y}\right)^2 = 0.$$

Это и поможет показать, что коэффициент $\bar{a}_{12}$ тоже обращается в нуль. Действительно,

$$\bar{a}_{12} = a_{11}\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial x} + a_{12}\left(\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial y} + \frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial x}\right) + a_{22}\frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial y} =$$
$$= a_{11}\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial x} + \sqrt{a_{11}}\sqrt{a_{22}}\frac{\partial\xi}{\partial x}\frac{\partial\eta}{\partial y} + \sqrt{a_{11}}\sqrt{a_{22}}\frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial x} +$$
$$+ a_{22}\frac{\partial\xi}{\partial y}\frac{\partial\eta}{\partial y} = \sqrt{a_{11}}\frac{\partial\xi}{\partial x}\left(\sqrt{a_{11}}\frac{\partial\eta}{\partial x} + \sqrt{a_{22}}\frac{\partial\eta}{\partial y}\right) +$$
$$+ \sqrt{a_{22}}\frac{\partial\xi}{\partial y}\left(\sqrt{a_{11}}\frac{\partial\eta}{\partial x} + \sqrt{a_{22}}\frac{\partial\eta}{\partial y}\right) =$$
$$= \left(\sqrt{a_{11}}\frac{\partial\eta}{\partial x} + \sqrt{a_{22}}\frac{\partial\eta}{\partial y}\right)\left(\sqrt{a_{11}}\frac{\partial\xi}{\partial x} + \sqrt{a_{22}}\frac{\partial\xi}{\partial y}\right) = 0.$$

Итак, уравнение (1.60) принимает вид

$$\bar{a}_{22}\frac{\partial^2 u}{\partial\eta^2} + \bar{F} = 0.$$

После деления на $\bar{a}_{22}$ окончательно получим

$$\frac{\partial^2 u}{\partial\eta^2} = \bar{\Phi}\left(\xi,\ \eta,\ u,\ \frac{\partial u}{\partial\xi},\ \frac{\partial u}{\partial\eta}\right), \tag{1.68}$$

где

$$\bar{\Phi} = -\frac{\bar{F}}{\bar{a}_{22}}.$$

Уравнение (1.68) называют канонической формой уравнения параболического типа.

Интересно отметить, что если правая часть уравнения (1.68) не содержит $\frac{\partial u}{\partial\xi}$, то оно становится обыкновенным дифференциальным уравнением, где роль параметра играет ξ.

Пример 2. Привести к каноническому виду уравнение

$$x^2\frac{\partial^2 u}{\partial x^2} + 2xy\frac{\partial^2 u}{\partial x\,\partial y} + y^2\frac{\partial^2 u}{\partial y^2} = 0, \quad x > 0 \tag{1.69}$$

и найти его решение.

Решение. Определим тип заданного уравнения. Здесь

$$a_{11} = x^2; \quad a_{12} = xy; \quad a_{22} = y^2;$$
$$a_{12}^2 - a_{11}a_{22} = x^2y^2 - x^2y^2 = 0.$$

Итак, уравнение (1.69) параболического типа.

Согласно общей теории составим характеристическое уравнение

$$x^2dy^2 - 2\,xy\,dx\,dy + y^2dx^2 = 0.$$

Легко увидеть, что здесь

$$(xdy - ydx)^2 = 0,$$

т. е. $xdy - ydx = 0$.

Разделяя переменные и интегрируя, получим одно семейство характеристик $\frac{y}{x} = C$ — семейство гипербол.

Новые переменные вводим по формулам

$$\xi = \frac{y}{x}; \quad \eta = y.$$

Выбор второй функции сделан с учетом того, чтобы она была наиболее простой, и с другой стороны, чтобы якобиан преобразования был отличен от нуля.

Находим теперь $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial x\,\partial y}$ и $\frac{\partial^2 u}{\partial y^2}$:

$$\frac{\partial^2 u}{\partial x^2} = \frac{y^2}{x^4}\frac{\partial^2 u}{\partial \xi^2} + \frac{2y}{x^3}\frac{\partial u}{\partial \xi};$$

$$\frac{\partial^2 u}{\partial x\,\partial y} = -\frac{y}{x^3}\frac{\partial^2 u}{\partial \xi^2} - \frac{y}{x^2}\frac{\partial^2 u}{\partial \xi\,\partial \eta} - \frac{1}{x^2}\frac{\partial u}{\partial \xi};$$

$$\frac{\partial^2 u}{\partial y^2} = \frac{1}{x^2}\frac{\partial^2 u}{\partial \xi^2} + \frac{2}{x}\frac{\partial^2 u}{\partial \xi\,\partial \eta} + \frac{\partial^2 u}{\partial \eta^2}.$$

Подставляя полученные выражения в уравнение (1.69), окончательно запишем это уравнение в канонической форме

$$\frac{\partial^2 u}{\partial \eta^2} = 0.$$

Полученное уравнение легко интегрируется

$$\frac{\partial u}{\partial \eta} = \theta(\xi); \quad u = \int \theta(\xi)\,d\eta + \lambda(\xi)$$

или

$$u = \theta(\xi)\,\eta + \lambda(\xi),$$

где θ и λ — произвольные функции.

Возвращаясь к старым переменным x и y, получим общее решение исходного уравнения

$$u = \theta\left(\frac{y}{x}\right)y + \lambda\left(\frac{y}{x}\right).$$

1. 6. 3. Уравнение эллиптического типа

Правые части уравнений (1.63) и (1.64) — комплексно сопряжены. Пусть $\varphi(x, y)$ — комплексный интеграл уравнения (1.63). Тогда $\varphi^*(x, y) = C$ интеграл уравнения (1.64), где φ^* — функция комплексно сопряженная с функцией φ.

Если сейчас перейти к комплексным переменным $\xi = \varphi(x, y)$, $\eta = \varphi^*(x, y)$, то согласно общей теории уравнение (1.56) приведется к виду

$$\frac{\partial^2 u}{\partial\xi\,\partial\eta} = \Phi\left(\xi,\ \eta,\ u,\ \frac{\partial u}{\partial\xi},\ \frac{\partial u}{\partial\eta}\right),$$

т. е. точно к такому же виду, как и гиперболическое уравнение. Желая остаться в действительной области, т. е. не иметь дела с комплексными переменными, сделаем еще одну замену переменных

$$\alpha = \frac{\varphi + \varphi^*}{2};\quad \beta = \frac{\varphi - \varphi^*}{2j},$$

где α и β — новые переменные.

Тогда

$$\xi = \alpha + j\beta;\quad \eta = \alpha - j\beta.$$

Легко показать, что при такой замене переменных $\bar{a}_{11} = \bar{a}_{22}$ и $\bar{a}_{12} = 0$. Таким образом уравнение (1.56) приведется к виду

$$\frac{\partial^2 u}{\partial\alpha^2} + \frac{\partial^2 u}{\partial\beta^2} = \Phi\left(\alpha,\ \beta,\ u,\ \frac{\partial u}{\partial\alpha},\ \frac{\partial u}{\partial\beta}\right),\quad \Phi = -\frac{\overline{F}}{\bar{a}_{22}},$$

который называется каноническим видом уравнения эллиптического типа.

Пример 3. Привести к каноническому виду уравнение

$$\frac{\partial^2 u}{\partial x^2} - 4\frac{\partial^2 u}{\partial x\,\partial y} + 5\frac{\partial^2 u}{\partial y^2} = 0. \tag{1.70}$$

Решение. Определяем тип заданного уравнения. Здесь $a_{11} = 1$; $a_{12} = -2$; $a_{22} = 5$. Отсюда $a_{12}^2 - a_{11}a_{22} = 4 - 5 = -1 < 0$. Уравнение (1.70) принадлежит эллиптическому типу.

Характеристическое уравнение будет иметь вид

$$dy^2 + 4\,dx\,dy + 5\,dx^2 = 0.$$

Выпишем уравнения первого порядка, на которые оно распадается, соответствующие (1.63) и (1.64),

$$y' = -2 + j;\quad y' = -2 - j.$$

Интегрируя полученные уравнения, найдем два общих интеграла характеристического уравнения

$$y = (-2 + j)\,x + C_1 \quad\text{и}\quad y = (-2 - j)\,x + C_2.$$

Здесь

$$\varphi = y - (-2 + j)\,x;\quad \varphi^* = y + (2 + j)\,x;$$

$$\alpha = \frac{\varphi + \varphi^*}{2} = y + 2x;\quad \beta = \frac{\varphi - \varphi^*}{2j} = -x.$$

Согласно общей теории нужно ввести новые переменные α и β по формулам

$$\xi = \alpha + j\beta = y + 2x - jx; \quad \eta = \alpha - j\beta = y + 2x + jx.$$

Вычисляя $\frac{\partial^2 u}{\partial x^2}$, $\frac{\partial^2 u}{\partial x\, \partial y}$ и $\frac{\partial^2 u}{\partial y^2}$ по формулам (1.59), после всех преобразований получим

$$\frac{\partial^2 u}{\partial \xi^2} + \frac{\partial^2 u}{\partial \eta^2} = 0,$$

т. е. получим уравнение (1.70) эллиптического типа в каноническом виде. Обращаем ваше внимание на то, что каноническое уравнение представляет собой двумерное уравнение Лапласа.

В заключение параграфа сделаем несколько замечаний.

Очевидно, представляет интерес рассмотрение случая, когда уравнение (1.56) имеет постоянные коэффициенты. Заметим, что такому уравнению соответствует характеристическое уравнение с постоянными коэффициентами, а поэтому характеристиками такого уравнения будут прямые

$$y = \frac{a_{12} + \sqrt{a_{12}^2 - a_{11}a_{22}}}{a_{11}}\, x + C_1$$

и

$$y = \frac{a_{12} - \sqrt{a_{12}^2 - a_{11}a_{22}}}{a_{11}}\, x + C_2.$$

Не останавливаясь на деталях, отметим, что после соответствующей замены переменных исходное уравнение приводится к одной из простейших форм

$$\frac{\partial^2 u}{\partial \xi\, \partial \eta} + b_1 \frac{\partial u}{\partial \xi} + b_2 \frac{\partial u}{\partial \eta} + cu + f = 0$$

или

$$\frac{\partial^2 u}{\partial \xi^2} - \frac{\partial^2 u}{\partial \eta^2} + b_1 \frac{\partial u}{\partial \xi} + b_2 \frac{\partial u}{\partial \eta} + cu + f = 0,$$

когда $a_{12}^2 - a_{11}a_{22} > 0$ — гиперболический тип;

$$\frac{\partial^2 u}{\partial \xi\, \partial \eta} + b_1 \frac{\partial u}{\partial \xi} + b_2 \frac{\partial u}{\partial \eta} + cu + f = 0,$$

когда $a_{12}^2 - a_{11}a_{22} = 0$ — параболический тип;

$$\frac{\partial^2 u}{\partial \xi^2} + \frac{\partial^2 u}{\partial \eta^2} + b_1 \frac{\partial u}{\partial \xi} + b_2 \frac{\partial u}{\partial \eta} + cu + f = 0,$$

когда $a_{12}^2 - a_{11}a_{22} < 0$ — эллиптический тип.

Эти уравнения можно еще более упростить, если ввести новую неизвестную функцию $v(\xi, \eta)$ по формуле

$$u(\xi, \eta) = e^{\lambda\xi+\mu\eta} v(\xi, \eta).$$

Дело в том, что числа λ и μ можно подобрать так, чтобы исчезли члены с производными первого поряда в уравнениях гиперболического и эллиптического типов, а в уравнении параболического типа исчез один из членов с производной первого порядка и член с неизвестной функцией.

Производя необходимые здесь очевидные вычисления, окончательно получим

$$\frac{\partial^2 v}{\partial\xi\,\partial\eta} + \gamma v + f_1 = 0$$

или

$$\frac{\partial^2 v}{\partial\xi^2} - \frac{\partial^2 v}{\partial\eta^2} + \gamma v + f_1 = 0 \text{ (гиперболический тип);}$$

$$\frac{\partial^2 v}{\partial\xi^2} + b_2 \frac{\partial v}{\partial\eta} + f_1 = 0 \text{ (параболический тип);}$$

$$\frac{\partial^2 v}{\partial\xi^2} + \frac{\partial^2 v}{\partial\eta^2} + \gamma v + f_1 = 0 \text{ (эллиптический тип).}$$

Одномерное волновое уравнение (1.23) принадлежит, очевидно, к гиперболическому типу. Оно легко может быть приведено к каноническому виду методами, изложенными выше. Одномерное уравнение теплопроводности

$$\rho c_v \frac{\partial T}{\partial t} = k \frac{\partial^2 T}{\partial x^2} + \varphi(x, t)$$

относится к параболическому типу и имеет канонический вид. Двумерное уравнение Лапласа

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0$$

представляет собой уравнение эллиптического типа в канонической форме.

Если число независимых переменных больше двух, то и в этом случае допустима классификация уравнений второго порядка, аналогичная той, которая приведена выше.

Классификация уравнений в частных производных второго порядка имеет фундаментальное значение для теории этих уравнений по ряду причин. Во-первых, решения уравнений одного и того же типа имеют много общих свойств в связи с чем и методы решения этих уравнений будут иметь специфический характер, учитывающий тип уравнения. Во-вторых, уравнения одного и того же типа описывают определенный круг физических и технических задач, имеющих много общих черт. Так, например, с дифференциальными уравнениями в частных производных гиперболического типа мы встречаемся при изучении волновых физических полей, а также в тех областях техники, где играет роль волновое движение. Пожалуй, наиболее важными из них можно назвать следующие.

1. Колебание струн и мембран.
2. Крутильные колебания стержней.
3. Распространение волн в жидкостях (акустика, звуколокационные устройства и т. п.).
4. Распространение волн в земле (например, сейсмология).
5. Передача электрических сигналов по линиям связи.
6. Распространение электромагнитных волн в пространстве и волноводах.
7. Квантовая механика.

Заметим, что из трех типов уравнений в частных производных второго порядка уравнения гиперболического типа труднее всего поддаются решению.

Если обратиться к уравнениям параболического типа, то можно отметить, что они встречаются при изучении переходных процессов, которые осуществляются в системах, содержащих как распределенные элементы рассеяния энергии, так и один из двух видов накопителей энергии. Особое значение параболические уравнения имеют в области проблем термодинамики. Ведь тепловые поля всегда содержат тепловые сопротивления и теплоемкости. Здесь температура является потенциальной функцией, тепловой поток — рассеивателем энергии. Из важных областей, где приходится иметь дело с уравнениями параболического типа, укажем фильтрацию жидкости в пористой среде, диффузию мелких частиц в воз духе или в какой-либо жидкости.

Уравнения эллиптического типа описывают физические поля, в которых достигнут установившийся или статический режим. Так уравнение Лапласа описывает поля, не имеющие внутренних источников, а уравнение Пуассона — поля с распределенными внутренними источниками. Эти уравнения встречаются почти во всех областях прикладной физики, поэтому они считаются наиболее важными из трех типов дифференциальных уравнений в частных производных.

Глава 2. ПРОСТЕЙШИЕ МЕТОДЫ ИНТЕГРИРОВАНИЯ УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ ВТОРОГО ПОРЯДКА

§ 2.1. МЕТОД ХАРАКТЕРИСТИК

Изучение методов интегрирования уравнений в частных производных второго порядка мы начинаем с метода характеристик не случайно. В предыдущем параграфе, по существу, уже изложены его основы. На примерах, продемонстрированных там, убедительно были показаны максимальные возможности примененного метода: уравнение, преобразованное к характеристикам, становилось обыкновенным дифференциальным уравнением, допускающим его интегрирование в квадратурах. Получающееся при этом общее решение исходного уравнения тогда нас вполне удовлетворяло. Теперь же мы поставим задачу несколько шире. Потребуем, чтобы уравнение, отнесенное к характеристикам, имело такое общее решение, которое позволяет легко получать решения различных краевых задач. Таким свойством обладает уравнение в частных производных второго порядка не каждого типа. Если рассмотреть, например, двумерное уравнение Лапласа

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0,$$

то его общее решение, как известно (см. пример 5 § 1.2), может быть представлено в виде $u(x, y)=\varphi(x+jy)+\psi(x-jy)$: ведь в этом случае характеристики уравнения мнимые ($x+jy=C_1$ и $x-jy=C_2$) и поэтому не могут иметь непосредственного значения для решения краевых задач. В случае уравнений гиперболического типа метод характеристик становится удобным и эффективным средством решения и исследования ряда краевых задач. Подробное знакомство с методом характеристик мы начнем с примера решения задачи Коши для однородного волнового уравнения, уравнения колебания струны. Впоследствии этим же методом будет решена краевая задача для закрепленной струны и рассмотрены некоторые вопросы применения метода характеристик к изучению электрических колебаний в проводах.

2. 1. 1. Колебания бесконечной струны. Формула Даламбера

Предположим, что нас интересуют свободные колебания настолько длинной струны, что ее практически можно считать бесконечной. Как известно (§ 1.4), эта задача сводится к решению уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < +\infty; \; t > 0, \tag{2.1}$$

при начальных условиях

$$u(x,\ t)\Big|_{t=0} = f(x); \quad \frac{\partial u(x,\ t)}{\partial t}\Big|_{t=0} = \varphi(x), \tag{2.2}$$

где $f(x)$ и $\varphi(x)$ — заданные функции: $f(x)$ определяет начальное отклонение струны, а $\varphi(x)$ — ее начальную скорость.

Другими словами, нам необходимо решить задачу Коши для одномерного волнового уравнения.

Получим сначала общее решение уравнения (2.1), а затем, воспользовавшись начальными условиями (2.2), исключим две произвольные функции, входящие в это решение.

Приведем уравнение (2.1) к характеристикам. Для этого составим характеристическое уравнение

$$a_{11}dt^2 - 2a_{12}\,dxdt + a_{22}dx^2 = 0.$$

В рассматриваемом случае $a_{11} = a^2$; $a_{12} = 0$; $a_{22} = -1$. Характеристическое уравнение будет иметь вид

$$a^2dt^2 - dx^2 = 0.$$

Оно распадается на два уравнения

$$adt - dx = 0 \text{ и } adt + dx = 0.$$

Интегрируя их, получим два семейства характеристик

$$x - at = C_1 \text{ и } x + at = C_2.$$

Для упрощения исходного уравнения введем новые независимые переменные

$$\xi = x - at; \quad \eta = x + at.$$

Известно (§ 1.6), что преобразованное уравнение будет иметь канонический вид

$$\frac{\partial^2 u}{\partial \xi\, \partial \eta} = 0.$$

Для решения этого уравнения поступаем точно так же, как и при решении примера 4 § 1.2. В результате получим

$$u(x,\ t) = \theta_1(\xi) + \theta_2(\eta),$$

где θ_1 и θ_2 — произвольные функции.

Возвращаясь к старым переменным, окончательно находим

$$u(x,\ t) = \theta_1(x - at) + \theta_2(x + at). \tag{2.3}$$

Очевидно, чтобы функции $u(x, t)$, определяемые формулой (2.3), были решениями заданного уравнения, необходимо наложить на θ_1 и θ_2 некоторые ограничения. Они должны быть дважды дифференцируемы по обеим переменным. Тогда легко убедиться простой подстановкой, что и $\theta_1(x-at)$, и $\theta_2(x+at)$, а также их сумма будут решениями уравнения (2.1).

Решение (2.3) называется *решением Даламбера уравнения колебания струны.*

Определенный интерес представляет физическое истолкование решения (2.3), описывающего процесс колебания бесконечной струны.

Как показывает формула (2.3) любое решение уравнения (2.1) представимо в виде суперпозиции двух решений $u_1 = \theta_1(x-at)$ и $u_2 = \theta_2(x+at)$. Разберем каждое из этих решений отдельно. Остановимся сначала на физической интерпретации решения $u_1 = \theta_1(x-at)$, представляющего величину отклонения точек струны от положения равновесия. Зафиксируем точку $x = x_0$ и предположим, что из нее в положительном направлении оси Ox в начальный момент $t=0$ начинает двигаться наблюдатель со скоростью a. Очевидно, через t_1 *сек* этот наблюдатель будет в точке с абсциссой $x_1 = x_0 + at_1$. Величина отклонения, которое наблюдатель будет видеть в точке x_1 в момент t_1, легко определяется по формуле $u_1 = \theta_1(x_1 - at_1)$. Но так как $x_1 - at_1 = x_0$, то эта величина совпадает с величиной первоначального отклонения $\theta_1(x_0)$, т. е.

$$\theta_1(x_1 - at_1) = \theta_1(x_0).$$

Последнее равенство означает, что подвижный наблюдатель на протяжении всего времени будет фиксировать одну и ту же величину отклонения, которая равна $\theta_1(x_0)$. Если в начальный момент времени струна имеет определенную форму (профиль), то последняя будет двигаться по струне со скоростью a в положительном направлении оси Ox как жесткая система. Описанное явление, определяемое функцией $u_1 = \theta_1(x-at)$, называется распространением прямой волны. Решению $u_2 = \theta_2(x+at)$ можно дать аналогичную интерпретацию, только теперь волна двигается влево со скоростью a, а потому носит название обратной волны. Итак, общее решение уравнения (2.1) представляет собой наложение прямой и обратной волны (рис. 2.1).

Форму струны в момент t можно получить следующим образом. Вычертить графики функций $\theta_1(x)$ и $\theta_2(x)$ (в момент времени $t=0$), а затем раздвигать эти графики со скоростью a первый вправо, а второй влево по оси Ox. Через t *сек*, остановив движение, выполнить сложение ординат графиков. Наложение прямой и обратной волны может давать такую форму струны, которая очень далека от первоначальной формы каждого из слагаемых (рис. 2.2). Процесс распространения прямой

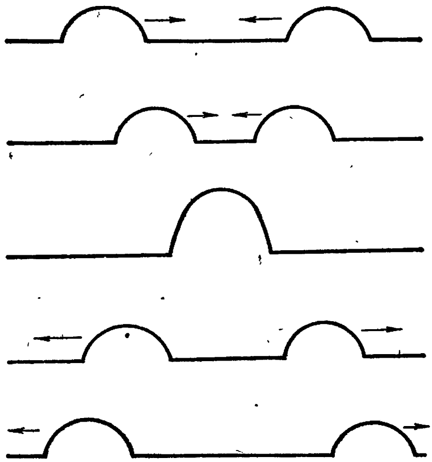

Рис. 2.1

и обратной волны имеет следующий геометрический смысл. Обратимся к характеристикам исследуемого уравнения:

$$x - at = C_1 \text{ и } x + at = C_2.$$

Очевидно, для всех точек характеристики $x - at = C_1$ величина отклонения для прямой волны будет одна и та же, т. е. $\theta_1(x - at) = \theta_1(C_1) = \text{const}$. Таким образом, роль линий уровня на плоскости xOt для функции двух переменных $\theta_1(x - at)$ будут играть характеристики заданного уравнения. Совершенно аналогично устанавливается, что линиями уровня обратной волны $\theta_2(x + at)$ будут служить линии второго семейства характеристик данного уравнения $x + at = C_2$. В связи с этим говорят, что прямая и обратная волны распространяются по характеристикам. Это замечательное свойство характеристик может быть использовано для нахождения значений функций $\theta_1(x - at)$ и $\theta_2(x + at)$ в некоторой фиксированной точке (x_0, t_0) плоскости xOt. Оказывается, значения этих функций в точке (x_0, t_0) вполне определяются их значениями в начальный момент времени $t = 0$. Покажем это. Проведем через точку (x_0, t_0) две характеристики $x - at = x_0 - at_0$ и $x + at = x_0 + at_0$ (рис. 2.3). Они пересекут ось Ox в двух точках x_1 и x_2. Но в связи с тем, что на характеристиках функции $\theta_1(x-at)$ и $\theta_2(x+at)$ принимают одни и те же значения $\theta_1(C_1)$ и $\theta_2(C_2)$, то и в точке (x_0, t_0) они примут такое же значение, что и в точках, соответственно, $(x_1, 0)$ и $(x_2, 0)$, т. е.

$$\theta_1(x_0-at_0)=\theta_1(x_1-a\cdot 0)=\theta_1(x_1);$$

$$\theta_2(x_0+at_0)=\theta_2(x_2+a\cdot 0)=\theta_2(x_2),$$

что и требовалось показать.

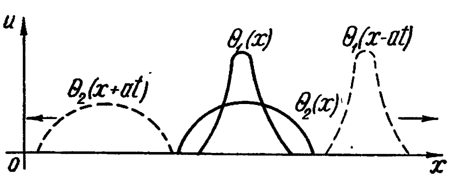

Рис. 2.2

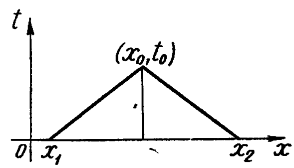

Рис. 2.3

Таким образом, чтобы определить величину отклонения $u(x, t)$ струны в точке (x_0, t_0) в момент времени t_0 можно провести через точку (x_0, t_0) плоскости xOt две характеристики $x-at=C_1$ и $x+at=C_2$. Тогда

$$u(x_0, t_0)=\theta_1(x_1)+\theta_2(x_2),$$

где x_1 и x_2 — точки пересечения оси Ox соответственно первой и второй характеристиками.

Проведенные здесь рассуждения еще раз подтверждают волновой характер решения уравнения колебания струны (2.1). В связи с этим уравнение (2.1) получило название волнового уравнения.

Вернемся теперь к решению задачи Коши. Другими словами, из множества решений (2.3) волнового уравнения (2.1) отыщем такое, которое удовлетворяет начальным условиям (2.2). В рассматриваемом случае это легко удается сделать. В силу первого из начальных условий находим

$$u(x, 0)=f(x)=\theta_1(x)+\theta_2(x).$$

Чтобы применить второе условие, придется продифференцировать равенство (2.3) по переменной t

$$u'_t(x, t)=(-a)\theta'_1(x-at)+a\theta'_2(x+at).$$

Подставляя в это равенство $t=0$, с учетом (2.2) получим

$$\varphi(x)=-a\theta'_1(x)+a\theta'_2(x).$$

Имеем два соотношения, связывающие функции θ_1 и θ_2, которые надо из полученных равенств исключить. Это удобнее будет сделать, если второе из этих соотношений будет содержать не производные функций θ_1 и θ_2, а просто сами функции. Проинтегрируем последнее равенство в пределах от 0 до x:

$$\int_0^x \varphi(x)\,dx = -a[\theta_1(x) - \theta_1(0)] + a[\theta_2(x) - \theta_2(0)]$$

или

$$-a\theta_1(x) + a\theta_2(x) = \int_0^x \varphi(x)\,dx + aC,$$

где $C = \theta_2(0) - \theta_1(0)$.

Итак, окончательно имеем следующую систему, из которой нам предстоит определить $\theta_1(x)$ и $\theta_2(x)$:

$$\theta_1(x) + \theta_2(x) = f(x);$$

$$-\theta_1(x) + \theta_2(x) = \frac{1}{a}\int_0^x \varphi(z)\,dz + C.$$

Отсюда

$$\theta_1(x) = \frac{1}{2}f(x) - \frac{1}{2a}\int_0^x \varphi(z)\,dz - \frac{C}{2};$$

$$\theta_2(x) = \frac{1}{2}f(x) + \frac{1}{2a}\int_0^x \varphi(z)\,dz + \frac{C}{2}.$$

В связи с тем, что последние равенства будут, очевидно, справедливы при любом значении x, заменяем в первом из них x на $x - at$, во втором x на $x + at$:

$$\theta_1(x - at) = \frac{1}{2}f(x - at) - \frac{1}{2a}\int_0^{x-at} \varphi(z)\,dz - \frac{C}{2};$$

$$\theta_2(x + at) = \frac{1}{2}f(x + at) + \frac{1}{2a}\int_0^{x+at} \varphi(z)\,dz + \frac{C}{2}.$$

Чтобы получить искомое решение (2.3) нужно эти равенства сложить. Тогда окончательно получим

$$u(x,\ t) = \frac{f(x - at) + f(x + at)}{2} + \frac{1}{2a}\int_{x-at}^{x+at} \varphi(z)\,dz. \qquad (2.4)$$

До сих пор мы не накладывали никаких ограничений на функции $\varphi(x)$ и $f(x)$, фигурирующие в начальных условиях (2.2). Формула (2.4) будет давать решение задачи Коши в том случае, если функция $f(x)$ обладает производными первого и второго порядков, а функция $\varphi(x)$ — производной первого порядка. К этому выводу вы сами придете, когда захотите убедиться в правильности полученного вами решения путем подстановки его в уравнение (2.1).

Формулу (2.4) называют формулой Деламбера решения задачи Коши для волнового уравнения (2.1).

Рассмотрим два частных случая задачи Коши, когда одно из начальных условий отсутствует. С физической точки зрения это может означать следующее. В начальный момент времени $t=0$ струна оттягивается, ей придается форма $f(x)$ и затем она отпускается. В этом случае, очевидно, начальная скорость точек струны равна нулю, т. е. $\varphi(x)\equiv 0$. Может быть и такая ситуация. В начальный момент времени, например, ударом молотка точкам струны сообщается некоторая мгновенная скорость, а затем изучается движение струны в последующие моменты времени. Теперь уже $f(x)\equiv 0$. Каждый из этих частных случаев представляет определенный интерес, а поэтому остановимся на них несколько подробнее.

1) Начальная скорость точек струны равна нулю. Для определенности будем считать, что в начальный момент времени $t=0$ струна имеет форму, изображенную на рис. 2.4. Вне промежутка (α, β) $f(x)=0$. При этих условиях величина отклонения точек струны в любой момент времени выразится формулой

$$u(x, t)=\frac{f(x-at)}{2}+\frac{f(x+at)}{2}.$$

(В формуле Даламбера (2.4) нужно положить $\varphi(x)\equiv 0$.) Полученное решение является суммой двух волн: прямой $\frac{1}{2}f(x-$ $-at)$ и обратной $\frac{1}{2}f(x+at)$. На рис. 2.4 изображен профиль струны в различные моменты времени. Начальная форма обеих волн определяется функцией $\frac{1}{2}f(x)$. В связи с тем, что функция $f(x)$ известна, мы сможем вычислить значение $u(x, t)$ при любых x и t. При этом можно пользоваться приемом с построением характеристик в плоскости xOt, который был описан выше.

Процесс колебания любой точки струны очень наглядно прослеживается на плоскости xOt с привлечением для этой цели характеристик уравнения. Каждая точка $M_0(x, t)$ плоскости соответствует точке струны с абсциссой x в момент времени

$t \geqslant 0$. Через точки α и β на оси Ox проведем по две характеристики. Они разобьют верхнюю полуплоскость на 6 частей (рис. 2.5). Очевидно, отклонение струны будет отлично от нуля только в тех точках плоскости xOt, для которых хотя бы одна из характеристик, проходящих через эти точки, пересекает ось абсцисс на отрезке (α, β).

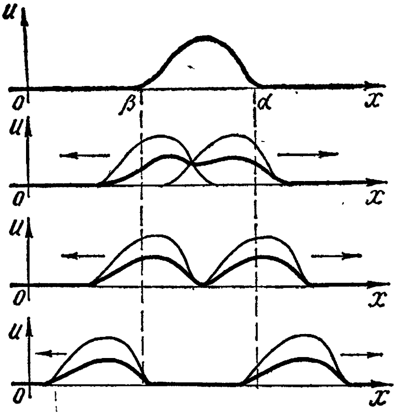

Рис. 2.4

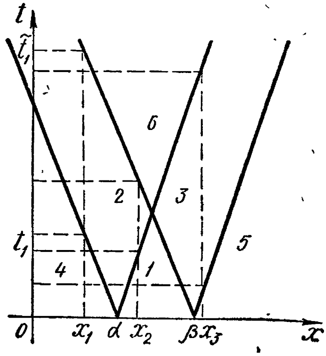

Рис. 2.5

Легко установить, что положению покоя соответствуют точки, принадлежащие зонам *4, 5* и *6*. Можно говорить о так называемых характеристических полосах, в которых отклонение точек струны отлично от нуля. Эти полосы соответствуют зонам *1, 2* и *3*. В зоне *3* действует только прямая волна, в зоне *2* — только обратная, а в зоне *1* — и та, и другая.

Интересно проследить, как меняется отклонение точки струны с абсциссой x_1 с течением времени t, начиная с начального.

Пусть сначала $x_1 < \alpha$. Тогда при $0 \leqslant t < t_1$ точка плоскости xOt находится в зоне *4* и $u(x, t) = 0$. Говорят, что в рассматриваемый промежуток времени колебание до точки еще не дошло. Волна, бегущая влево, дойдет до этой точки в момент времени $\tilde{t}_1$ и с этого момента точка струны начнет колебаться (до точки дошел передний фронт обратной волны). Как только волна пройдет через рассматриваемую точку, т. е. начиная с момента t_1, эта точка снова будет находиться в состоянии покоя (до точки дошел задний фронт обратной волны). Таким образом, точка x_1 участвует в колебательном процессе при $t_1 < t < \tilde{t}_1$.

Читатель без труда проследит происходящий процесс колебания точки струны в случае, когда $\alpha < x_2 < \beta$ и когда $x_3 > \beta$ (см. рис. 2.5).

2) Начальное отклонение точек струны равно нулю. По условию задачи $f(x) \equiv 0$, а поэтому решение (2.4) будет иметь вид

$$u(x, t) = \frac{1}{2a} \int_{x-at}^{x+at} \varphi(z)\, dz.$$

Если первообразную функции $\varphi(z)$ обозначить через $2a\Phi(z)$, то по формуле Лейбница — Ньютона получим

$$u(x, t) = \Phi(x + at) - \Phi(x - at).$$

Таким образом, по струне опять распространяются две волны: обратная и прямая. Желая выяснить роль начального импульса в процессе колебания струны, предположим, что начальная скорость струны равна $\varphi(x)$ на участке (α, β) и равна нулю вне этого промежутка. Попробуем проследить процесс колебания любой точки струны, воспользовавшись методом характеристик, примененным в предыдущем пункте (рис. 2.5). Как и ранее верхняя полуплоскость xOt разбивается характеристиками на 6 областей. Только теперь зоны *4* и *5* соответствуют точкам, для которых отклонение равно нулю, а вот в зоне *6* отклонение будет уже отличным от нуля. Покажем это. Как и ранее проследим, чему равно отклонение точки струны с абсциссой x_1 в различные моменты времени. Промежуток интегрирования $(x_1 - at,\ x_1 + at)$ при $t = 0$ вырождается в точку x_1, а затем с увеличением t расширяется в обе стороны со скоростью a. При $t < t_1$ он не имеет общих точек с промежутком (α, β), где $\varphi(x) \neq 0$, а поэтому для этих t $u(x, t) = 0$. Начиная с t_1, промежуток $(x_1 - at,\ x_1 + at)$ налегает на (α, β) и точка с абсциссой x_1 начинает колебаться. При $t > t_1$ промежуток $(x_1 - at,\ x_1 + at)$ полностью содержит промежуток (α, β). В этом случае интегрирование по промежутку $(x_1 - at,\ x_1 + at)$ сводится к интегрированию по промежутку (α, β), так как вне его $\varphi(x) = 0$. Отсюда при $t > t_1$

$$u(x, t) = \frac{1}{2a} \int_{\alpha}^{\beta} \varphi(z)\, dz = A. \tag{2.5}$$

Если момент $t = t_1$ называется моментом прохождения переднего фронта волны через точку x_1, то момент $t = \tilde{t}_1$ называется моментом прохождения заднего фронта волны через эту точку.

После того, как через точку x_1 пройдет задний фронт волны, она не будет возвращаться в прежнее положение равновесия, как это было ранее. Формула (2.5) говорит о том, что эта точка останется смещенной на постоянную величину A. Таким образом, с течением времени все точки струны окажутся смещенными на величину A и останутся без движения в этом положении. Волны, прошедшие через участок струны, соответствующий промежутку (α, β), оставляют как бы след своего прохождения. На рис. 2.6 представлена форма струны в некоторый момент $t > t_1$. Через участок III струны уже прошли прямая и обратная волны. В данный момент времени прямая волна проходит через участок IV, а обратная через участок II. Участки I и V находятся еще в состоянии покоя, так как до них волны еще не дошли.

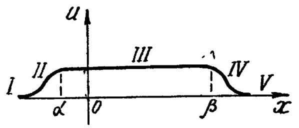

Рис. 2.6

2. 1. 2. КОЛЕБАНИЯ ПОЛУБЕСКОНЕЧНОЙ СТРУНЫ. ОТРАЖЕНИЕ ВОЛН

Неправильно было бы думать, что метод характеристик, который был применен в предыдущем пункте для решения задачи Коши волнового уравнения (2.1), не может быть использован для решения других краевых задач. Особый интерес для исследователя, конечно, представляет случай, когда струна имеет конечную длину. Так возникает смешанная краевая задача для уравнения (2.1), когда кроме начальных условий на концах струны задаются еще краевые условия. В этом пункте мы используем метод характеристик для решения краевой задачи, описывающей колебания полубесконечной струны $x \geqslant 0$ с закрепленным левым концом. Краевую задачу, описывающую колебание конечной струны с двумя закрепленными концами, мы рассмотрим в следующем параграфе. Это, конечно, не означает, что она не может быть решена изучаемым здесь методом характеристик.

Пусть требуется найти решение уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 \leqslant x < +\infty; \quad t > 0,$$

удовлетворяющее начальным условиям

$$u(x,\ t)\Big|_{t=0}=f(x);\quad \frac{\partial u(x,t)}{\partial t}\Big|_{t=0}=\varphi(x)$$

и краевому условию

$$u(x;\ t)\,|_{x=0}=u(0,\ t)=0. \tag{2.6}$$

Функции $f(x)$ и $\varphi(x)$ определены сейчас только на промежутке $[0, +\infty)$. Условия согласования будут иметь вид $f(0)=\varphi(0)$.

Для решения этой задачи мы хотим воспользоваться формулой Даламбера (2.4). Но, к сожалению, ее прямое использование в рассматриваемом случае затруднено. Это объясняется тем, что в формулу Даламбера входит разность $x-at$, которая может быть и отрицательной величиной, а для отрицательных значений аргумента функции $f(x)$ и $\varphi(x)$ не определены. Здесь мы попадаем почти в такую же ситуацию, как тогда, когда занимались представлением функции интегралом Фурье на промежутке $[0, +\infty)$ в ч. 1 пособия. Там мы доопределяли заданную функцию на промежутке $(-\infty, 0)$ четным или нечетным образом, что давало нам возможность решить поставленную задачу. Теперь мы поступим точно так же. Доопределим функции $f(x)$ и $\varphi(x)$ на промежутке $(-\infty, 0)$ (говорят, выполним продолжение этих функций на $(-\infty, 0)$). Казалось бы, что теперь нет никаких препятствий, чтобы воспользоваться формулой Даламбера для вновь полученных функций. Но нельзя забывать еще одну существенную деталь. Продолжения функций должны быть такими, чтобы кроме начальных условий выполнялось еще и краевое условие (2.6). Спрашивается, каким свойством должны обладать эти функции, чтобы формула Даламбера давала решение поставленной задачи?

Запишем формулу Даламбера для функций $F(x)$ и $\Phi(x)$, являющихся, соответственно, продолжениями функций $f(x)$ и $\varphi(x)$,

$$u(x,\ t)=\frac{F(x-at)+F(x+at)}{2}+\frac{1}{2a}\int\limits_{x-at}^{x+at}\Phi(z)\,dz. \tag{2.7}$$

В силу краевых условий (2.6) должны иметь

$$u(0,\ t)=\frac{F(-at)+F(at)}{2}+\frac{1}{2a}\int\limits_{-at}^{at}\Phi(z)\,dz\equiv 0.$$

Последнее тождество выполняется тогда, когда $F(at)=-F(-at)$ и $\int\limits_{-at}^{at}\Phi(z)\,dz=0$. Таким образом, функции $F(x)$ и $\Phi(x)$ должны быть нечетными.

Подведем итог нашим рассуждениям.

Для решения заданной краевой задачи мы продолжили функции $f(x)$ и $\varphi(x)$ нечетным образом на отрицательную часть оси Ox и обозначили их через $F(x)$ и $\Phi(x)$:

$$F(x)=\begin{cases} f(x) & \text{при } x\geqslant 0,\\ -f(-x) & \text{при } x<0;\end{cases} \qquad \Phi(x)=\begin{cases} \varphi(x) & \text{при } x\geqslant 0,\\ -\varphi(-x) & \text{при } x<0.\end{cases}$$

Тогда функция (2.7) — решение краевой задачи.

Рассмотрим процесс колебания струны вблизи закрепленного конца, связанный с важным вопросом об отражении волн.

Очевидно, решение рассматриваемой краевой задачи можно написать в известной форме

$$u(x,\ t)=\theta_1(x-at)+\theta_2(x+at). \tag{2.8}$$

Процесс колебания точек струны около закрепленного конца очень хорошо прослеживается на плоскости xOt с привлечением для этой цели характеристик уравнения.

Возьмем точку $(x_0,\ t_0)$ плоскости xOt. Желая найти величину отклонения $u(x_0,\ t_0)$, проведем через точку $(x_0,\ t_0)$ две характеристики $x-at=x_0-at_0$ и $x+at=x_0+at_0$, пересекающие ось Ox в точках $-x_1$ и x_2 (рис. 2.7). Как известно, отклонение $u(x_0,\ t_0)$ складывается из отклонения, обусловленного обратной волной, пришедшей из точки x_2, т. е. $\theta_2(x_2)$, и отклонения, обусловленного прямой волной, пришедшей из точки $-x_1$, т. е. $\theta_1(-x_1)$. Другими словами

$$u(x_0,\ t_0)=\theta_1(-x_1)+\theta_2(x_2).$$

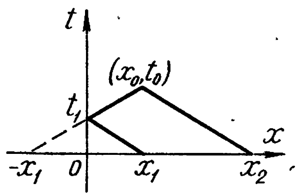

Рис. 2.7

Но в точке $-x_1$ в начальный момент времени $t=0$ никакого начального возбуждения не могло быть, поскольку начальные условия в рассматриваемой задаче заданы лишь для $x>0$. Выйти из создавшегося положения поможет соотношение, связывающее функции θ_1 и θ_2 в силу краевого условия (2.6). С учетом его из (2.8) находим

$$\theta_1(-at)+\theta_2(at)=0,$$

т. е. $\theta_1(-z)=-\theta_2(z)$. Таким образом, при $z=x_1$ $\theta_1(-x_1)=-\theta_2(x_1)$. Это означает, что вместо прямой волны, идущей из точки $-x_1$, можно рассматривать обратную волну, вышедшую в момент времени $t=0$ из симметричной относительно начала координат точки x_1. В момент $t=t_1$ эта волна достигнет точки $x=0$. Теперь уже ее нужно заменить прямой волной, вышедшей из точки $x=0$ в момент времени $t=t_1$ и несущей величину отклонения, равную $-\theta_1(-x_1)$, что следует из условия $\theta_2(x_1) = -\theta_1(-x_1)$. Итак, на конце струны $x = 0$ происходит явление отражения обратной волны с сохранением величины отклонения, но с изменением его знака на противоположный. После отражения по струне побегут вправо с одинаковой скоростью две волны, одна из которых в начальный момент $t = 0$ уже вышла из точки x_1, а вторая — отраженная волна, отстающая от первой по времени на t_1.

Предлагаем читателю самостоятельно убедиться в том, что при решении задачи, описывающей колебание струны длиной l с двумя закрепленными концами, функции $f(x)$ и $\varphi(x)$, определенные на промежутке $(0, l)$, необходимо продолжить на промежуток $(-l, 0)$ нечетным образом, а затем выполнить периодическое продолжение вновь полученных функций на всю числовую ось с периодом $2l$.

2. 1. 3. Электрические колебания в длинных линиях. Телеграфное уравнение *

Знакомство с теорией электрических цепей мы начали с рассмотрения колебательного контура, состоящего из последовательно соединенных сопротивления, емкости и катушки самоиндукции (см. ч. 1, стр. 107). Это был пример системы с сосредоточенными параметрами. Поведение такой системы полностью определяется способом взаимодействия перечисленных выше элементов с источником напряжения. При составлении уравнения, описывающего работу цепи, не принимались во внимание ни длина линий связи, ни порядок расположения элементов.

Но в практике приходится встречаться с задачами, когда электрическая цепь имеет такую протяженность, что упомянутые выше элементы необходимо считать распределенными вдоль всей ее длины. Так приходят к рассмотрению электрической цепи с распределенными параметрами.

Примером цепи с распределенными параметрами может служить длинная линия (телеграфная линия или линия передачи энергии), вдоль которой равномерно и непрерывно распределены сопротивление, индуктивность, емкость и проводимость (утечка изоляции).

Изучение колебательного процесса, происходящего в длинной линии, мы начнем с составления уравнения, управляющего работой этой системы. Для большей наглядности будем

* Перед чтением п. 2.1.3 рекомендуем повторить материал § 2.15 ч. 1 пособия.

предполагать, что линия двухпроводная, вдоль которой равномерно распределены рассчитанные на единицу длины сопротивление r, самоиндукция l, емкость c и утечка изоляции g. Проводники линии предполагаем расположенными вдоль оси Ox (рис. 2.8). В силу вышесказанного, сила тока i в сечении проводника и напряжение u, равное разности потенциалов между соответствующими точками A и A' обоих проводников, являются здесь неизвестными величинами, зависящими от двух переменных: абсциссы x и времени t. Задача состоит в том, чтобы найти уравнение, связывающее $i(x, t)$ и $u(x, t)$.

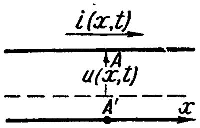

Рис. 2.8

Чтобы перейти от цепи с распределенными параметрами к цепи с сосредоточенными параметрами, рассмотрим в момент времени t элемент AA_1, малой длины Δx одного проводника

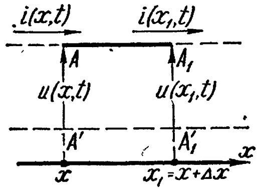

Рис. 2.9

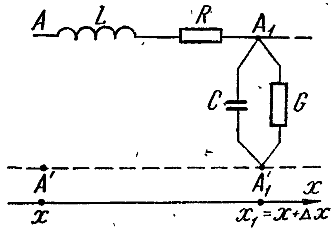

Рис. 2.10

между его сечениями в точках x и $x_1 = x + \Delta x$. Обозначим токи в этих сечениях через $i(x, t)$ и $i(x_1, t)$, напряжения соответственно — $u(x, t)$ и $u(x_1, t)$ (рис. 2.9). Тогда заданные параметры выделенного участка будут равны $r\Delta x$; $l\Delta x$; $c\Delta x$ и $g\Delta x$. Желая воспользоваться законами Кирхгофа, представим себе рассматриваемый участок в виде звена с сосредоточенными параметрами, где $C = c\Delta x$; $L = l\Delta x$; $R = r\Delta x$ и $G = g\Delta x$. Тогда можно силу тока, подходящего слева к точке A_1, ввиду малости Δx считать равной $i(x, t)$, а точку A_1 — узлом, где этот ток как бы разветвляется на ток $i(x_1, t)$, ток зарядки конденсатора i_c и ток утечки через изоляцию i_g (рис. 2.10). По первому закону Кирхгофа (см. ч. 1, стр. 109) сумма всех токов, протекающих через узел цепи из всех элементов, которые примыкают к этому узлу, равна нулю. В применении к рассматриваемому случаю, получим

$$i(x_1, t) + i_c + i_g - i(x, t) \approx 0. \qquad (2.9)$$

Как известно,

$$i_c = C\frac{du_c}{dt};\ i_g = Gu_g,$$

где u_c и u_g — падения напряжений на емкости и утечке изоляции, которые здесь можно считать равными падению напряжения в сечении x_1, т. е. $u_c = u_g \approx u(x_1, t)$. Тогда

$$i_c \approx c\Delta x \frac{\partial u(x_1, t)}{\partial t};\ i_g \approx g\Delta x\, u(x_1, t).$$

Подставляем полученные выражения в (2.9):

$$i(x_1, t) - i(x, t) + c\Delta x \frac{\partial u(x_1, t)}{\partial t} + g\Delta x\, u(x_1, t) \approx 0.$$

Деля обе части этого равенства на Δx и переходя к пределу при $\Delta x \to 0$ $(x_1 \to x)$, окончательно получим

$$\frac{\partial i(x, t)}{\partial x} + c\frac{\partial u(x, t)}{\partial t} + g\,u(x, t) = 0. \tag{2.10}$$

Чтобы получить второе соотношение, связывающее неизвестные функции $i(x, t)$ и $u(x, t)$, воспользуемся вторым законом Кирхгофа. Падение напряжения на элементе AA_1 равно $u(x, t) - u(x_1, t)$. Обозначив через u_r и u_l падения напряжений на сопротивлении и емкости, получим

$$u(x, t) - u(x_1, t) = u_r + u_l. \tag{2.11}$$

Как известно,

$$u_r = Ri_r;\ u_l = L\frac{di_l}{dt},$$

где i_r и i_l — токи, проходящие через сопротивление и индуктивность.

По предположению, сделанному выше, $i_r = i_l \approx i(x, t)$. Выражения для u_r и u_l могут быть переписаны следующим образом

$$u_r \approx r\Delta x\, i(x, t);\ u_l \approx l\Delta x \frac{\partial i(x, t)}{\partial t}$$

и равенство (2.11) приобретет вид

$$u(x_1, t) - u(x, t) + r\Delta x\, i(x, t) + l\Delta x \frac{\partial i(x, t)}{\partial t} \approx 0.$$

Деля обе части этого равенства на Δx и переходя к пределу при $\Delta x \to 0$ $(x_1 \to x)$, окончательно получим

$$\frac{\partial u(x, t)}{\partial x} + r\,i(x, t) + l\frac{\partial i(x, t)}{\partial t} = 0. \tag{2.12}$$

Уравнения (2.10) и (2.12) называются обычно дифференциальными уравнениями электрической линии, или системой телеграфных уравнений. Из этих уравнений можно исключить любую из неизвестных функций $i(x, t)$ или $u(x, t)$ и получить тем самым уравнение лишь для одной из них. Продифференцировав уравнение (2.10) по t и умножив его на l, получим

$$l\frac{\partial^2 i}{\partial x\,\partial t}+cl\frac{\partial^2 u}{\partial t^2}+gl\frac{\partial u}{\partial t}=0,$$

а продифференцировав уравнение (2.12) по x, найдем

$$\frac{\partial^2 u}{\partial x^2}+r\frac{\partial i}{\partial x}+l\frac{\partial^2 i}{\partial x\,\partial t}=0.$$

Вычтем из второго уравнения первое:

$$\frac{\partial^2 u}{\partial x^2}+r\frac{\partial i}{\partial x}-cl\frac{\partial^2 u}{\partial t^2}-gl\frac{\partial u}{\partial t}=0.$$

Заменяя здесь $\frac{\partial i}{\partial x}$ его значением из формулы (2.10), получим

$$\frac{\partial^2 u}{\partial t^2}-\frac{1}{lc}\frac{\partial^2 u}{\partial x^2}+\frac{rc+lg}{lc}\frac{\partial u}{\partial t}+\frac{rg}{lc}u=0. \qquad (2.13)$$

Уравнение (2.13) называется *телеграфным уравнением*. Аналогичным образом можно убедиться в том, что функция $i(x, t)$ удовлетворяет точно такому же уравнению

$$\frac{\partial^2 i}{\partial t^2}-\frac{1}{lc}\frac{\partial^2 i}{\partial x^2}+\frac{rc+lg}{lc}\frac{\partial i}{\partial t}+\frac{rg}{lc}i=0. \qquad (2.14)$$

Телеграфное уравнение есть линейное однородное уравнение в частных производных второго порядка гиперболического типа. Действительно, здесь $a_{11}=-\frac{1}{lc}$; $a_{12}=0$; $a_{22}=1$. Тогда

$$a_{12}^2-a_{11}a_{22}=\frac{1}{lc}>0.$$

Для исследования этого уравнения применим метод характеристик, изученный в предыдущих пунктах параграфа.

Считая линию простирающейся в обе стороны до бесконечности, составим задачу Коши для уравнения (2.13). Для этого нам необходимо задать начальные условия.

Предположим, что распределение напряжения и тока в момент времени $t=0$ известно

$$u(x, t)\big|_{t=0}=f(x);\quad i(x, t)\big|_{t=0}=\varphi(x). \qquad (2.15)$$

Телеграфное уравнение представляет собой уравнение второго порядка, поэтому должны быть заданы производные $\frac{\partial u(x, t)}{\partial t}$

и $\frac{\partial i(x, t)}{\partial t}$ в начальный момент времени, т. е. при $t = 0$. В рассматриваемом случае их можно отыскать из уравнений (2.10) и (2.12).

Из уравнения (2.10)

$$\frac{\partial u}{\partial t}\Big|_{t=0} = -\frac{g}{c}\, u(x, t)\Big|_{t=0} - \frac{1}{c}\,\frac{\partial i(x, t)}{\partial x}\Big|_{t=0}.$$

Условие (2.15) даст возможность определить $\frac{\partial i(x, t)}{\partial x}$ при $t = 0$. Действительно,

$$\frac{\partial i(x, t)}{\partial x}\Big|_{t=0} = \varphi'(x).$$

Итак,

$$\frac{\partial u(x, t)}{\partial t}\Big|_{t=0} = -\frac{g}{c} f(x) - \frac{1}{c}\varphi'(x).$$

Совершенно аналогично

$$\frac{\partial i(x, t)}{\partial t}\Big|_{t=0} = -\frac{r}{l}\varphi(x) - \frac{1}{l} f'(x).$$

Задача Коши для телеграфного уравнения может быть сформулирована следующим образом.

Найти решение уравнения

$$\frac{\partial^2 u}{\partial t^2} - \frac{1}{lc}\frac{\partial^2 u}{\partial x^2} + \frac{rc + lg}{lc}\frac{\partial u}{\partial t} + \frac{rg}{lc} u = 0,$$

$$-\infty < x < +\infty, \quad t > 0,$$

удовлетворяющее начальным условиям

$$u(x, t)\big|_{t=0} = f(x); \quad \frac{\partial u(x, t)}{\partial t}\Big|_{t=0} = -\frac{g}{c} f(x) - \frac{1}{c}\varphi'(x), \tag{2.16}$$

где $f(x)$ и $\varphi(x)$ — заданные функции.

Решение задачи Коши для телеграфного уравнения имеет определенные трудности. Мы рассмотрим лишь некоторые частные случаи этого уравнения, имеющие большое практическое значение.

1) Может оказаться, что сопротивление линии и проводимость изоляции настолько малы, что ими можно пренебречь. Аналитически это будет выражаться условием

$$r = g = 0.$$

Такую линию обычно называют линией без потерь.

Линия без потерь, как легко увидеть, будет описываться знакомым нам волновым уравнением

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial x^2},$$

где $a = \frac{1}{\sqrt{lc}}$, в которое перейдет телеграфное уравнение при $r = g = 0$.

Таким образом, мы попадаем в условие задачи, которая уже была решена ранее. Нам остается напомнить основной результат и записать формулу Даламбера для заданных начальных условий, которые будут теперь иметь вид

$$u(x, t)|_{t=0} = f(x); \quad \left.\frac{\partial u(x, t)}{\partial t}\right|_{t=0} = -\frac{1}{c}\varphi'(x).$$

С учетом этих начальных условий по формуле Даламбера получим

$$u(x, t) = \frac{f(x+at) + f(x-at)}{2} - \frac{1}{2ac}\int_{x-at}^{x+at}\varphi'(x)\,dx =$$

$$= \frac{f(x+at) + f(x-at)}{2} + \sqrt{\frac{l}{c}}\,\frac{\varphi(x-at) - \varphi(x+at)}{2}, \qquad (2.17)$$

где $\frac{1}{2ac} = \frac{\sqrt{lc}}{c} = \sqrt{\frac{l}{c}}$.

Можно сделать следующий вывод: решение телеграфного уравнения имеет волновой характер и представляет результат наложения волн произвольной формы, распространяющихся без искажения с определенной скоростью.

Читатель без особого труда получит решение задачи Коши относительно функции $i(x, t)$ для рассматриваемой линии без потерь.

2) Параметры изучаемой линии могут быть связаны между собой соотношением

$$\frac{r}{l} = \frac{g}{c}. \qquad (2.18)$$

Линию, обладающую этим свойством, называют линией без искажений. Это название оправдывается тем, что и в этом случае решение телеграфного уравнения будет иметь волновой характер, но только с той разницей, что амплитуда волны при ее распространении по линии уменьшается, т. е. происходит затухание волн. Покажем это.

Телеграфное уравнение (2.13) представляет собой уравнение в частных производных второго порядка с постоянными коэффициентами. В § 1.6 мы познакомились с приемом, который позволяет путем введения новой функции избавиться в изучаемом уравнении от члена, содержащего производную первого порядка $\frac{\partial u}{\partial t}$. Прием этот состоял в том, что вводилась неизвестная функция $v(x, t)$ по формуле

$$u(x, t) = e^{\lambda t} v(x, t).$$

Коэффициент λ подбирался таким образом, чтобы в преобразованном уравнении член с производной первого порядка от неизвестной функции исчез. Проделаем необходимые здесь вычисления

$$\frac{\partial^2 u}{\partial x^2} = e^{\lambda t}\frac{\partial^2 v}{\partial x^2}; \quad \frac{\partial^2 u}{\partial t^2} = \left(\frac{\partial^2 v}{\partial t^2} + 2\lambda\frac{\partial v}{\partial t} + \lambda^2 v\right)e^{\lambda t},$$

$$\frac{\partial u}{\partial t} = \left(\frac{\partial v}{\partial t} + \lambda v\right)e^{\lambda t}.$$

После подстановки найденных выражений в уравнение (2.13 и сокращения на $e^{\lambda t}$ получим

$$\frac{\partial^2 v}{\partial t^2} - \frac{1}{lc}\frac{\partial^2 v}{\partial x^2} + (2\lambda lc + rc + gl)\frac{\partial v}{\partial t} + [\lambda^2 lc + \lambda(rc + gl) + gr]\, v = 0.$$

Величина λ может быть определена из условия равенства нулю коэффициента при $\frac{\partial v}{\partial t}$

$$2\lambda lc + rc + gl = 0.$$

Отсюда

$$\lambda = -\frac{rc + gl}{2lc} = -\frac{1}{2}\left(\frac{r}{l} + \frac{g}{c}\right). \qquad (2.19)$$

Но из (2.18) $rc = gl$. Тогда для λ получим окончательно

$$\lambda = -\frac{r}{l}.$$

Интересно отметить, что при найденном значении λ с учетом (2.18) исчезает член, содержащий неизвестную функцию $v(x, t)$. Действительно,

$$\lambda^2 lc + \lambda(rc + gl) + gr = \frac{r^2}{l^2}lc - \frac{r}{l}\cdot 2rc + gr =$$

$$= \frac{r^2 c}{l} - \frac{2r^2 c}{l} + \frac{r^2 c}{l} = 0.$$

Итак, для функции $v(x, t)$ получаем уравнение

$$\frac{\partial^2 v}{\partial t^2} - \frac{1}{lc}\frac{\partial^2 v}{\partial x^2} = 0,$$

которое является волновым уравнением со скоростью распространения колебаний $a = \frac{1}{\sqrt{lc}}$.

Желая решить задачу Коши для этого уравнения, найдем начальные условия для функции $v(x, t)$, если начальные условия для $u(x, t)$ уже известны из равенства (2.16). Соотношение, связывающее функции $u(x, t)$ и $v(x, t)$, теперь имеет вид

$$u(x, t) = e^{-\frac{r}{l}t} v(x, t).$$

Отсюда получаем

$$u(x, t)\big|_{t=0} = v(x, t)\big|_{t=0} = f(x).$$

Проводя очевидные вычисления, убеждаемся, что

$$\left.\frac{\partial v(x, t)}{\partial t}\right|_{t=0} = \frac{r}{l} f(x) - \frac{g}{c} f(x) - \frac{1}{c}\varphi'(x) = -\frac{1}{c}\varphi'(x),$$

т. е. начальные условия для функции $v(x, t)$ совпадают с начальными условиями задачи Коши для линии без потерь. Поэтому решение рассматриваемой задачи будет отличаться от решения (2.17) только дополнительным множителем $e^{-\frac{r}{l}t}$

$$u(x, t) = e^{-\frac{r}{l}t}\left[\frac{f(x+at)+f(x-at)}{2} + \sqrt{\frac{l}{c}}\,\frac{\varphi(x-at)-\varphi(x+at)}{2}\right].$$

(Читатель без особого труда получит соответствующие формулы и решение задачи для функции $i(x, t)$.)

Итак, мы убедились, что решение телеграфного уравнения при условии (2.18) имеет волновой характер, причем волны напряжения распространяются без искажения с постоянной скоростью, если не считать наличие множителя $e^{-\frac{r}{l}t}$, благодаря которому они постепенно затухают, причем тем быстрее. чем больше показатель $\frac{r}{l}$, который называется коэффициентом затухания.

Полученный нами результат имеет исключительно важное практическое значение. Ведь за счет соответствующего подбора емкости и самоиндукции телеграфного кабеля удается добиться передачи сигналов почти в неискаженном виде (говорят, в пропорционально неискаженной форме). Другими словами, принятый по этой линии сигнал будет совпадать с точностью до множителя с переданным.

Разобранные случаи свидетельствуют о том, что при определенных условиях по длинной линии можно передавать сигналы любой формы. Наше исследование было бы неполным, если бы мы не остановились на характере телеграфного уравнения. И первый вопрос, который здесь возникает, следующий: можно ли по длинной линии с потерями передать сигнал произвольной формы?

Под длинной линией с потерями мы понимаем такую линию, которая описывается полным телеграфным уравнением (2.13). Из предыдущего известно, что путем введения новой функции $v(x, t)$ по формуле

$$u(x, t) = e^{\lambda t} v(x, t)$$

можно привести уравнение (2.13) к виду

$$\frac{\partial^2 v}{\partial t^2} = a^2 \frac{\partial^2 v}{\partial x^2} + b^2 v, \tag{2.20}$$

где коэффициенты a^2 и b^2 определяются с учетом того, что λ находится по формуле (2.19).

Предположим, что уравнение (2.20) имеет решение в виде прямой бегущей волны

$$v(x, t) = \theta(x - \alpha t), \tag{2.21}$$

где α — фазовая скорость распространения этой волны. Подстановка (2.21) в уравнение (2.20) дает

$$(a^2 - \alpha^2)\, \theta''(B) + b^2 \theta(B) = 0, \tag{2.22}$$

где через $B = x - \alpha t$ обозначена фаза волны.

В связи с тем, что по нашему предположению $\theta(B)$ — произвольная функция, равенство (2.22) будет иметь место только тогда, когда коэффициенты при $\theta''(B)$ и $\theta(B)$ обратятся в нуль

$$a^2 - \alpha^2 = 0; \quad b^2 = 0.$$

Но ведь в линии с потерями $b^2 \neq 0$. Значит, по длинной линии, описываемой уравнением (2. 20), волны произвольной формы распространяться не могут, т. е. сигналы произвольной формы по такой линии передавать нельзя.

Можно попытаться из уравнения (2.22) установить форму, которую могут иметь бегущие волны в рассматриваемом случае. Уравнение (2.22) представляет собой обыкновенное линейное дифференциальное уравнение с постоянными коэффициентами относительно $\theta(B)$. Поэтому его решения могут быть записаны в виде комплексных гармоник

$$C(\omega)\, e^{\pm j\omega B},$$

где $C(\omega)$ — комплексная амлитуда, а B — фаза.

Очевидно, волновое число ω удовлетворяет характеристическому уравнению

$$-(a^2 - \alpha^2)\, \omega^2 + b^2 = 0,$$

из которого легко получить выражение для скорости α бегущих волн:

$$\alpha = \pm\sqrt{a^2 - \frac{b^2}{\omega^2}}.$$

Замечаем, что возможные скорости бегущих волн не постоянны, как это было в случае линии без искажений ($b \neq 0$), а зависят от волнового числа ω.

Таким образом, разным волновым числам ω будут соответствовать бегущие волны

$$C(\omega)e^{\mp i\omega(x-\alpha t)}, \tag{2.23}$$

распространяющиеся с различными фазовыми скоростями α. Это явление носит название дисперсии волн.

Решение $v(x, t)$ можно считать суперпозицией ряда бегущих волн (2.23) с одним и тем же направлением распространения, но разными фазовыми скоростями. В связи с этим форма волны, которую имели в момент времени $t=0$, будет изменяться с течением времени.

§ 2. 2. МЕТОД РАЗДЕЛЕНИЯ ПЕРЕМЕННЫХ

Метод характеристик, изложенный в предыдущем параграфе, наряду с почти прозрачной простотой имеет целый ряд недостатков. Во-первых, он эффективно может быть применен только к уравнениям гиперболического типа. Что же касается уравнений параболического и эллиптического типов, то его применение или ограничено, или вообще невозможно. Во-вторых, использование метода характеристик для решения смешанной краевой задачи (задаются начальные и краевые условия) связано, как видели, с аналитическим продолжением заданных функций, что носит искусственный характер и далеко от физического существа дела.

Предлагаемый вашему вниманию метод разделения переменных (метод Фурье) во многом лишен этих недостатков. Приспособленный для решения, вообще говоря, смешанных краевых задач, он с успехом может быть применен к уравнениям и гиперболического, и параболического, и эллиптического типов.

В основе метода характеристик, как известно, лежит нахождения общего решения заданного уравнения. Метод Фурье предусматривает уже нахождение частных решений этого уравнения.

Сущность этого метода заключается в следующем. Искомая функция выражается через произведение двух или трех величин, каждая из которых является функцией одной из пространственных координат (или времени). Если такое произведение ввести в исходное дефференциальное уравнение, то вместо одного дифференциального уравнения в частных производных можно получить систему из двух или трех обыкновенных дифференциальных уравнений, каждое из которых решается отдельно. Так как исходное дифференциальное уравнение линейно относительно искомой функции и ее производных, то сумма произвольного числа найденных таким способом частных решений его есть опять решение (принцип суперпозиции). Большой набор частных решений позволит составить такое решение уравнения, которое будет удовлетворять заданным условиям. Существенным в описываемом методе является и то обстоятельство, что он применяется обычно в том случае, когда кривые (поверхности), на которых заданы граничные условия, являются координатными кривыми (поверхностями) в декартовых координатах или в какой-либо системе криволинейных координат (полярной, цилиндрической, сферической и т. д.). Именно в этом случае удается так подобрать входящие в частные решения параметры, чтобы все решение в целом удовлетворяло всем граничным и начальным условиям и представлялось бы при этом сходящимся рядом.

Применение метода Фурье продемонстрируем на ряде примеров.

2. 2. 1. Свободные колебания закрепленной струны

Пусть задана однородная струна длиной l, закрепленная в точках $x=0$ и $x=l$. Задача о свободных колебаниях такой струны может быть сформулирована следующим образом (см. § 1.5).

Найти решение $u(x, t)$ однородного уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \tag{2.24}$$

удовлетворяющее начальным условиям

$$u(x, t)\big|_{t=0} = f(x); \quad \left.\frac{\partial u(x, t)}{\partial t}\right|_{t=0} = \varphi(x) \tag{2.25}$$

и однородном краевым условиям

$$u(x, t)\big|_{x=0} = u(0, t) = 0; \quad u(x, t)\big|_{x=l} = u(l, t) = 0. \tag{2.26}$$

Решение этой задачи разобьем на несколько этапов.

1) Найдем сначала частные решения уравнения (2.24)

$$u_1(x, t),\ u_2(x, t),\ u_3(x, t),\ \ldots,\ u_n(x, t),\ \ldots,$$

не равные тождественно нулю (нетривиальные), в виде произведения двух функций, одна из которых зависит только от x, а другая — только от t

$$u(x, t) = X(x)T(t), \tag{2.27}$$

причем так, чтобы эти частные решения удовлетворяли граничным условиям (2.26).

2) Следуя принципу суперпозиции (см. § 1.3), составим ряд из бесконечной последовательности найденных частных решений

$$u(x, t) = C_1u_1(x, t) + C_2u_2(x, t) + C_3u_3(x, t) + \ldots + C_nu_n(x, t) + \ldots$$

При определенных условиях этот ряд будет решением уравнения (2.24), удовлетворяющим краевым условиям (2.26).

3) Подбираем коэффициенты C_i так, чтобы функция $u(x, t)$ удовлетворяла еще и начальным условиям (2.25).

Найденная таким образом функция $u(x, t)$ будет, очевидно, решением заданной краевой задачи.

Приступаем к реализации намеченного плана: отыскиваем решения уравнения (2.24) в виде (2.27). Продифференцируем выражение (2.27) дважды, сначала по x, а затем по t. Тогда получим

$$\frac{\partial^2 u}{\partial x^2} = X''(x)T(t); \quad \frac{\partial^2 u}{\partial t^2} = X(x)T''(t).$$

Подставим эти выражения в исходное уравнение (2.24):

$$X(x)T''(t) = a^2X''(x)T(t).$$

Разделяя переменные, получим

$$\frac{T''(t)}{a^2T(t)} = \frac{X''(x)}{X(x)}. \tag{2.28}$$

Обратите внимание на то, что левая часть равенства (2.28) зависит только от t, а правая — только от x. Значит, если в правой части этого равенства зафиксировать x, положив $x = x_0$, то левая часть при всех значениях t будет величиной постоянной. Точно такой же вывод можно сделать и относительно правой части равенства (2.28). Отсюда заключаем, что оба отношения $\frac{1}{a^2}\frac{T''(t)}{T(t)}$ и $\frac{X''(x)}{X(x)}$ являются величинами постоянными, т. е.

$$\frac{1}{a^2}\frac{T''(t)}{T(t)} = \frac{X''(x)}{X(x)} = \lambda,$$

где $\lambda = \text{const}$. Отсюда получим

$$T''(t) - a^2\lambda T(t) = 0; \tag{2.29}$$

$$X''(x) - \lambda X(x) = 0. \tag{2.30}$$

Таким образом, функции $T(t)$ и $X(x)$ должны удовлетворять обыкновенным дифференциальным уравнениям соответственно (2.29) и (2.30). Очевидно, краевые условия (2.26) накладывают на решения этих уравнений определенные условия. Найдем их. В силу (2.26) при любых t должны иметь

$$u(x, t)|_{x=0} = u(0, t) = X(0)\,T(t) = 0;$$
$$u(x, t)|_{x=l} = u(l, t) = X(l)\,T(t) = 0.$$

Отсюда $X(0) = 0$ и $X(l) = 0$, иначе решение (2.27) было бы тривиальным, что исключено условием.

Итак, задача свелась к тому, чтобы найти нетривиальные решения обыкновенного дифференциального уравнения (2.30), удовлетворяющие краевым условиям $X(0) = 0$; $X(l) = 0$.

Как известно (см. ч. 2, стр. 190), краевая задача

$$X''(x) - \lambda X(x) = 0; \tag{2.31}$$

$$X(0) = 0; \ X(l) = 0 \tag{2.32}$$

будет иметь нетривиальное решение, если определитель

$$D(0, l) = \begin{vmatrix} y_1(0) & y_2(0) \\ y_1(l) & y_2(l) \end{vmatrix}$$

будет равен нулю, где $y_1(x)$ и $y_2(x)$ — фундаментальная система решений уравнения (2.31).

Найдем фундаментальную систему решений уравнения (2.31). Полагая $X(x) = e^{rx}$, составим для этого уравнения характеристическое уравнение

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

Решая его, найдем $r_1 = r_2 = 0$, если $\lambda = 0$; $r_1 = \sqrt{\lambda}$; $r_2 = -\sqrt{\lambda}$, если $\lambda \neq 0$. Итак,

$$y_1(x) = 1; \ y_2 = x, \text{ если } \lambda = 0,$$

и

$$y_1(x) = e^{\sqrt{\lambda}x}; \ y_2(x) = e^{-\sqrt{\lambda}x}, \text{ если } \lambda \neq 0.$$

При $\lambda = 0$

$$D(0, l) = \begin{vmatrix} 1 & 0 \\ 1 & l \end{vmatrix} = l.$$

Но так как $l \neq 0$, то в этом случае краевая задача нетривиальных решений не имеет.

Найдем теперь те значения $\lambda \neq 0$, при которых рассматриваемая краевая задача будет иметь нетривиальные решения

$$D(0,\ l) = \begin{vmatrix} 1 & 1 \\ e^{\sqrt{\lambda} l} & e^{-\sqrt{\lambda} l} \end{vmatrix} = 0.$$

Отсюда $e^{-\sqrt{\lambda} l} - e^{\sqrt{\lambda} l} = 0$ или $e^{-\sqrt{\lambda} l}(1 - e^{2\sqrt{\lambda} l}) = 0$. Последнее равенство возможно только в случае, если $\lambda < 0$.

Итак, $\lambda < 0$. Положим $\lambda = -\mu^2$. В этом случае корни характеристического уравнения будут чисто мнимые, а фундаментальная система решений будет иметь вид

$$y_1(x) = \cos \mu x;\quad y_2(x) = \sin \mu x.$$

Для нахождения значений μ, при которых краевая задача будет иметь нетривиальное решение, составим определитель

$$D(0,\ l) = \begin{vmatrix} 1 & 0 \\ \cos \mu l & \sin \mu l \end{vmatrix} = 0.$$

Отсюда $\sin \mu l = 0$. Решая это тригонометрическое уравнение относительно μ, найдем

$$\mu l = k\pi,$$

или

$$\mu = \frac{k\pi}{l},\quad k = \pm 1,\ \pm 2,\ \pm 3,\ \ldots,$$

т. е.

$$\lambda = -\frac{k^2\pi^2}{l^2}.$$

Функция $y_2(x) = \sin \frac{k\pi}{l} x$ при фиксированном k является решением уравнения (2.31) и кроме того $y_2(0) = \sin \frac{k\pi}{l} 0 = 0$; $y_2(l) = \sin k\pi = 0$, т. е. удовлетворяет краевым условиям (2.32).

Значит, нетривиальные решения $X_k(x)$ рассматриваемой краевой задачи будут иметь вид

$$X_k(x) = \sin \frac{k\pi}{l} x,\quad k = 1,\ 2,\ 3,\ \ldots \tag{2.33}$$

Здесь k принимает только положительные значения по той простой причине, что при отрицательных k мы опять получаем решения уравнения точно такого же вида, отличающиеся от первых множителем -1, который здесь можно отбросить.

Напомним, что величины

$$\lambda_1 = \frac{\pi^2}{l^2}; \quad \lambda_2 = \frac{2^2\pi^2}{l^2}; \quad \ldots; \quad \lambda_k = \frac{k^2\pi^2}{l^2}; \quad \ldots$$

называются собственными числами (значениями) краевой задачи (2.31), (2.32), а соответствующие им функции

$$X_1(x) = \sin\frac{\pi x}{l}; \quad X_2(x) = \sin\frac{2\pi}{l}x; \quad \ldots; \quad X_k(x) = \sin\frac{k\pi}{l}x; \quad \ldots$$

собственными функциями этой задачи.

Найдем теперь функции $T(t)$. Для этого в уравнение (2.29) подставим найденное значение λ и решим его. В связи с тем, что каждому значению λ_k будет соответствовать вполне определенная функция $T_k(t)$ — решение уравнения (2.29), мы в это уравнение сразу введем новое обозначение. После подстановки λ получим

$$T_k''(t) + a^2\frac{k^2\pi^2}{l^2}T_k(t) = 0.$$

Здесь корни характеристического уравнения $r^2 + \frac{a^2k^2\pi^2}{l^2} = 0$ чисто мнимые $r_1 = \frac{ak\pi}{l}j$; $r_2 = -\frac{ak\pi}{l}j$. Поэтому общее решение уравнения будет иметь вид

$$T_k(t) = A_k\cos\frac{k\pi a}{l}t + B_k\sin\frac{k\pi a}{l}t, \tag{2.34}$$

где A_k и B_k — произвольные постоянные.

Подставляя (2.33) и (2.34) в равенство (2.27), получим бесчисленное множество нетривиальных решений уравнения (2.24), каждое из которых удовлетворяет краевым условиям (2.26):

$$u_k(x, t) = X_k(x)\,T_k(t) = \left(A_k\cos\frac{k\pi a}{l}t + B_k\sin\frac{k\pi a}{l}t\right)\sin\frac{k\pi}{l}x,$$

где $k = 1, 2, 3, \ldots$

Если бы удалось сейчас подобрать коэффициенты A_k и B_k найденных решений так, чтобы последние удовлетворяли начальным условиям (2.25), то задача была бы решена окончательно. К сожалению, это удается сделать лишь для очень узкого класса начальных функций $f(x)$ и $\varphi(x)$, что нас ни в коей мере удовлетворить не может. Поэтому перейдем ко второму этапу реализации намеченного плана: составим ряд из полученной последовательности частных решений уравнения

(2.24). Согласно принципу суперпозиции, сумма ряда $u(x, t)$ будет также удовлетворять однородному уравнению (2.24) и однородным краевым условиям (2.26).

$$u(x, t) = \sum_{k=1}^{\infty} u_k(x, t) = \sum_{k=1}^{\infty}\Big(A_k \cos\frac{k\pi a}{l}t + \\ + B_k \sin\frac{k\pi a}{l}t\Big)\sin\frac{k\pi}{l}x. \tag{2.35}$$

Теперь нам удастся подобрать коэффициенты A_k и B_k так, чтобы решение (2.35) удовлетворяло еще и начальным условиям (2.25) для довольно широкого класса начальных функций $f(x)$ и $\varphi(x)$.

Подставляя в (2.35) значение $t = 0$ и учитывая (2.25), получим равенство

$$u(x, 0) = \sum_{k=1}^{\infty} A_k \sin\frac{k\pi}{l}x = f(x), \tag{2.36}$$

которое должно выполняться для всех x из промежутка $(0, l)$. Из этого равенства попытаемся определить коэффициенты A_k, помня о том, что функция $f(x)$ известна. Равенство (2.36) представляет не что иное, как ряд Фурье, в который разложена функция $f(x)$ на промежутке $(0, l)$ по синусам. Для определения коэффициентов A_k можно воспользоваться известными формулами (см. ч. 1, стр. 47)

$$A_k = \frac{2}{l}\int_0^l f(x)\sin\frac{k\pi}{l}x\,dx.$$

Аналогичным приемом определяют и коэффициенты B_k. Только предварительно нужно продифференцировать (2.35) по t, а затем уже подставить $t = 0$.

$$\frac{\partial u(x, t)}{\partial t} = \sum_{k=1}^{\infty}\Big(-A_k\frac{k\pi a}{l}\sin\frac{k\pi a}{l}t + B_k\frac{k\pi a}{l}\cos\frac{k\pi a}{l}t\Big)\sin\frac{k\pi}{l}x.$$

После подстановки $t = 0$ с учетом второго начального условия (2.25), получим

$$\frac{\partial u(x, 0)}{\partial t} = \sum_{k=1}^{\infty} B_k\frac{k\pi a}{l}\sin\frac{k\pi}{l}x = \varphi(x).$$

Мы попадаем в ту же самую ситуацию, что и раньше: получен ряд Фурье для функции $\varphi(x)$, разложенный на промежутке $(0, l)$ по синусам. Только теперь роль коэффициентов разложения играют величины $B_k \frac{k\pi a}{l}$. Поэтому

$$B_k \frac{k\pi a}{l} = \frac{2}{l} \int_0^l \varphi(x) \sin \frac{k\pi}{l} x dx.$$

Отсюда

$$B_k = \frac{2}{k\pi a} \int_0^l \varphi(x) \sin \frac{k\pi}{l} x dx.$$

Остается подставить найденные выражения для коэффициентов A_k и B_k в ряд (2.35), чтобы получить решение $u(x, t)$ уравнения (2.24), удовлетворяющее краевым (2.26) и начальным условиям (2.25).

Все приведенные выше рассуждения носили очевидный формальный характер. Читателя не могло не насторожить то обстоятельство, что справедливость полученных выводов нигде не подтверждалось доказательствами, без которых существование и единственность решения краевой задачи в виде ряда (2.35) вызывает сомнение. Действительно, если, например, ряд (2.35) расходится или функция, определяемая этим рядом, не является дифференцируемой, то он, естественно, не может представлять решение рассматриваемой краевой задачи. Сформулированная ниже теорема призвана снять возникающие здесь вопросы и подвести своеобразный итог сказанному.

Теорема. *Если функция $f(x)$ на отрезке $[0, l]$ дважды непрерывно дифференцируема, имеет кусочно-непрерывную третью производную и удовлетворяет условиям*

$$f(0) = f(l) = 0; \; f''(0) = f''(l) = 0,$$

а функция $\varphi(x)$ непрерывно дифференцируема, имеет кусочно-непрерывную вторую производную и удовлетворяет условиям

$$\varphi(0) = \varphi(l) = 0,$$

то функция $u(x, t)$, определяемая рядом (2.35), имеет непрерывные производные второго порядка и удовлетворяет уравнению (2.24), начальным условиям (2.25) и краевым условиям (2.26). При этом возможно почленное дифференцирование ряда (2.35) по x и t два раза и полученные ряды сходятся абсолютно и равномерно при $0 \leqslant x \leqslant l$ и любом t.

Доказательство этой теоремы выходит за рамки пособия и мы его проводить не будем [35, стр. 92].

Представляет большой интерес физическая интерпретация решения (2.35), описывающего свободные колебания закрепленной струны.

В связи с тем, что решение (2.35) получается наложением отдельных колебаний, описываемых функциями $u_k(x, t)$, естественно начать анализ решения именно с этих функций. Для этого представим отдельный член ряда (2.35) в более удобном для анализа виде

$$u_k(x, t) = \left(A_k \cos\frac{k\pi a}{l} t + B_k \sin\frac{k\pi a}{l} t\right) \sin\frac{k\pi}{l} x =$$
$$= N_k \sin\frac{k\pi}{l} x \cdot \sin\left(\frac{k\pi a}{l} t + \psi_k\right),$$

где
$$N_k = \sqrt{A_k^2 + B_k^2}; \qquad \operatorname{tg}\psi_k = \frac{A_k}{B_k}.$$

Полученное представление позволяет сделать ряд выводов. Во-первых, каждая точка струны $x = x_0$ совершает гармоническое колебание

$$N_k \sin\frac{k\pi}{l} x_0 \sin\left(\frac{k\pi a}{l} t + \psi_k\right)$$

с амплитудой $N_k \sin\frac{k\pi}{l} x_0$, частотой $\omega_k = \frac{k\pi a}{l}$ и начальной фазой ψ_k. Причем, если амплитуда этого колебания зависит от абсциссы точки, то частота и фаза будут для всех точек одни и те же. Это означает, что все точки струны при своем движении одновременно (синхронно) проходят через положение равновесия, а также одновременно (синхронно) достигают своего максимального отклонения в ту или иную сторону. В отличие от бегущих волн, такие колебания струны называют *стоячими волнами*.

Во-вторых, профиль волны в любой момент времени представляет синусоиду

$$N_k \sin\left(\frac{k\pi a}{l} t_0 + \psi_k\right) \sin\frac{k\pi}{l} x,$$

навивающуюся на ось абсцисс. На рис. 2.11 изображены профили волн, соответствующие $k = 1, 2, 3$.

Обращают на себя внимание точки струны, которые остаются в покое во время движения. Так, при $k = 1$ такими точками являются концы струны, при $k = 2$ неподвижных точек уже будет три (к концам струны присоединяется еще ее середина), при $k = 3$ их будет четыре. Точки струны, которые остаются неподвижными при движении, т. е. амплитуда которых равна нулю, называются *узлами стоячей волны*.

Чтобы найти все узлы стоячей волны $u_k(x, t)$, нужно определить, очевидно, корни уравнения $\sin\frac{k\pi}{l}x = 0$ на промежутке $[0, l]$. Разрешая это уравнение, убеждаемся, что у волны $u_k(x, t)$ будет $k+1$ узел с абсциссами $0, \frac{l}{k}, \frac{2l}{k}, \ldots, \frac{(k-1)l}{k}, l$.

Представляют интерес и те точки струны, которые совершают колебания с максимальной амплитудой. Эти точки на-

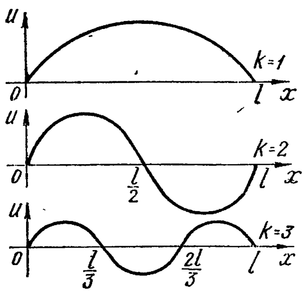

Рис. 2.11

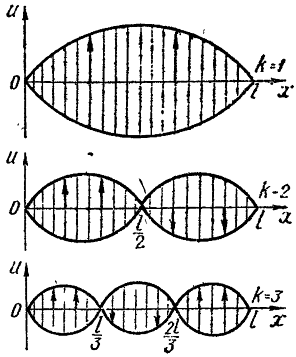

Рис. 2.12

зываются *пучностями стоячей волны*. Пучности стоячей волны располагаются посередине между узлами и могут быть найдены из условия

$$\sin\frac{k\pi}{l}x = 1.$$

Разрешая это уравнение, получим абсциссы пучностей стоячей волны $u_k(x, t)$: $\frac{l}{2k}, \frac{3l}{2k}, \ldots, \frac{2k-1}{2k}l$.

На рис. 2.12 схематически показан процесс колебания точек струны для стоячих волн, соответствующих $k = 1, 2, 3$. (Стрелками показаны направления одновременных отклонений точек струны.)

В повседневной жизни мы воспринимаем колебание струны по звуку, который она издает. Очевидно, звук струны является результатом наложения колебаний, соответствующих стоячим волнам $u_k(x, t)$. Сила звука, издаваемого струной, определяется ее амплитудой N_k, а высота звука зависит от частоты

$$\omega_k = \frac{k\pi}{l}a = \frac{k\pi}{l}\sqrt{\frac{T}{\rho}}, \quad a^2 = \frac{T}{\rho}.$$

Самый низкий тон, который может издавать струна, соответствует самой низкой частоте

$$\omega_1 = \frac{\pi}{l}\sqrt{\frac{T}{\rho}}. \qquad (2.37)$$

Этот тон называется *основным тоном струны*. Формула (2.37) красноречиво свидетельствует о том, что звук будет тем выше, чем будет больше натяжение T струны и чем короче струна.

Тона, соответствующие ω_2, ω_3, ω_4, ..., называются *обертонами*. Амплитуды соответствующих гармоник как правило быстро убывают ($N_k \to 0$ при $k \to \infty$), а поэтому эти тона влияют лишь на тембр звука.

Практика игры на таких струнных инструментах, как гитара, балалайка, скрипка и др., показывает, что звук, издаваемый струной, резко меняется, если ее, например, прижать точно в середине. Чем это объясняется? А дело в том, что в момент прикосновения к середине струны мы гасим стоячие волны, имеющие в этой точке пучности, и сохраняем те гармоники, которые имеют в этой точке узлы. Останутся только четные гармоники, и самой низкой частотой будет так называемая частота октавы основного тона

$$\omega_2 = \frac{2\pi}{l}\sqrt{\frac{T}{\rho}}.$$

Если теперь прижать струну на расстоянии $\frac{1}{3}$ ее длины от начала, то высота основного тона повышается втрое, так как при этом сохраняются те гармоники, которые имеют узлы в точке $x = \frac{l}{3}$.

Пример 1. Найти закон колебания струны длиной l, расположенной на отрезке $[0, l]$, если в начальный момент струне придана форма кривой $f(x) = H\sin\frac{2\pi}{l}x$, а затем она отпущена без начальной скорости. Струна закреплена на концах. Внешние силы отсутствуют.

Решение. Сформулируем краевую задачу, соответствующую условиям задачи.

Найти решение однородного волнового уравнения (внешние силы отсутствуют)

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial x^2},$$

удовлетворяющее начальным условиям

$$u(x, t)\big|_{t=0} = H\sin\frac{2\pi}{l}x; \qquad \left.\frac{\partial u(x, t)}{\partial t}\right|_{t=0} \equiv 0$$

и краевым условиям

$$u(0,\ t) = 0; \qquad u(l,\ t) = 0.$$

В связи с тем, что начальная скорость струны равна нулю, все коэффициенты B_k в решении (2.35) в свою очередь равны нулю, т. е. $B_k = 0$. Тогда искомое решение будет иметь вид

$$u(x,\ t) = \sum_{k=1}^{\infty} A_k \cos\frac{k\pi a}{l}\, t \sin\frac{k\pi}{l} x.$$

Находим коэффициенты A_k

$$A_k = \frac{2}{l}\int_0^l f(x)\sin\frac{k\pi}{l}\, x dx = \frac{2}{l}\int_0^l H\sin\frac{2\pi}{l}\, x\sin\frac{k\pi}{l}\, x dx.$$

Очевидно выписанный интеграл будет равен нулю для всех $k \neq 2$. При $k = 2$ получим

$$A_2 = \frac{2H}{l}\int_0^l \sin^2\frac{2\pi}{l}\, x dx = H.$$

Таким образом, решение будет иметь вид

$$u(x,\ t) = H\cos\frac{2\pi a}{l}\, t\sin\frac{2\pi}{l}\, x.$$

Как и следовало ожидать из физической постановки задачи, струна совершает стоячие колебания.

Пример 2. Найти закон колебания струны длиной l, если в начальный момент всем точкам струны сообщена скорость, равная v. Начальное отклонение отсутствует. Концы струны закреплены. Внешние силы отсутствуют.

Решение. Поставленную задачу можно сформулировать следующим образом. Найти решение однородного волнового уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial x^2},$$

удовлетворяющее начальным условиям

$$u(x,\ t)\,|_{t=0} = 0; \quad \left.\frac{\partial u(x,\ t)}{\partial t}\right|_{t=0} = v$$

и краевым условиям

$$u(0,\ t) = 0; \qquad u(l,\ t) = 0.$$

Очевидно, в общей формуле (2.35) все коэффициенты A_k равны нулю, т. е. $A_k = 0$. Тогда искомое решение примет вид

$$u(x,\ t) = \sum_{k=1}^{\infty} B_k \sin\frac{k\pi a}{l}\, t\sin\frac{k\pi}{l}\, x.$$

Находим коэффициенты

$$B_k = \frac{2}{k\pi a}\int_0^l v\sin\frac{k\pi}{l}x\,dx = \frac{2lv}{k^2\pi^2 a}(1-\cos k\pi) =$$

$$= \begin{cases} 0, & \text{если } k \text{ — четное;} \\ \dfrac{4lv}{k^2\pi^2 a}, & \text{если } k \text{ — нечетное.} \end{cases}$$

Окончательно имеем

$$u(x,\ t) = \frac{4lv}{\pi^2 a}\sum_{n=1}^{\infty}\frac{1}{(2n+1)^2}\sin\frac{(2n+1)\pi a}{l}t\ \ \sin\frac{(2n+1)\pi}{l}x.$$

2. 2. 2. Распространение тепла в ограниченном стержне

Пусть задан тонкий теплопроводящий стержень длиной l, боковая поверхность которого теплоизолирована и концы поддерживаются при нулевой температуре. Задача о распространении тепла в таком стержне может быть сформулирована следующим образом (см. § 1.5).

Найти решение $u(x,\ t)$ однородного уравнения теплопроводности

$$\frac{\partial u}{\partial t} = a^2\frac{\partial^2 u}{\partial x^2}, \qquad a^2 = \frac{k}{\rho c}, \tag{2.38}$$

удовлетворяющее начальному условию

$$u(x,\ t)\big|_{t=0} = f(x) \tag{2.39}$$

и однородным краевым условиям

$$u(x,\ t)\big|_{x=0} = u(0,\ t) = 0; \quad u(x,\ t)\big|_{x=l} = u(l,\ t) = 0. \tag{2.40}$$

Для решения этой задачи воспользуемся методом разделения переменных. Не останавливаясь на деталях, наметим ход решения задачи, который почти дословно повторяет рассуждения предыдущего пункта.

Следуя методу разделения переменных, ищем сначала нетривиальные частные решения уравнения (2.38) в виде

$$u(x,\ t) = X(x)\,T(t), \tag{2.41}$$

удовлетворяющие краевым условиям (2.40).

Подставим (2.41) в уравнение (2.38). После разделения переменных получим

$$\frac{1}{a^2}\frac{T'(t)}{T(t)}=\frac{X''(x)}{X(x)}=\lambda;$$

где $\lambda=\text{const}$. Как и ранее, получим два уравнения

$$T'(t)-a^2\lambda T(t)=0;$$
$$X''(t)-\lambda X(x)=0.$$

Краевые условия (2.40) дают

$$X(0)=0; \qquad X(l)=0.$$

Итак, для определения функции $X(x)$ мы опять получаем краевую задачу

$$X''(x)-\lambda X(x)=0;$$
$$X(0)=0; \qquad X(l)=0,$$

которая совпадает с краевой задачей (2.31), (2.32) предыдущего пункта. Как известно, собственными значениями этой задачи будут величины

$$\lambda_1=\frac{\pi^2}{l^2};\quad \lambda_2=\frac{2^2\pi^2}{l^2};\quad \ldots;\quad \lambda_k=\frac{k^2\pi^2}{l^2};\quad \ldots,$$

а соответствуюшие им собственные функции будут иметь вид

$$X_1(x)=\sin\frac{\pi}{l}x;\quad X_2(x)=\sin\frac{2\pi}{l}x;\quad \ldots;\quad X_k(x)=\sin\frac{k\pi}{l}x,\ \ldots$$

Найдем теперь функции $T(t)$. После подстановки $\lambda=-\mu^2=-\frac{k^2\pi^2}{l^2}$ в уравнение, определяющее эти функции, получим

$$T_k'(t)+a^2\frac{k^2\pi^2}{l^2}T_k(t)=0.$$

Общее решение этого уравнения будет иметь вид

$$T_k(t)=A_ke^{-\left(\frac{k\pi a}{l}\right)^2t},$$

где A_k — произвольные постоянные.

Подставляя найденные значения $X_k(x)$ и $T_k(x)$ в равенство (2.41), получим бесчисленное множество нетривиальных решений уравнения (2.38), каждое из которых удовлетворяет краевым условиям (2.40):

$$u_k(x,t)=X_k(x)T_k(t)=A_ke^{-\left(\frac{k\pi a}{l}\right)^2t}\sin\frac{k\pi}{l}x.$$

С целью удовлетворить начальному условию (2.39), составляем ряд

$$u(x, t) = \sum_{k=1}^{\infty} u_k(x, t) = \sum_{k=1}^{\infty} A_k e^{-\left(\frac{k\pi a}{l}\right)^2 t} \sin\frac{k\pi}{l} x. \tag{2.42}$$

Подставляя в (2.42) значение $t = 0$ и учитывая (2.39), получим условие, из которого определятся коэффициенты A_k

$$u(x, 0) = \sum_{k=1}^{\infty} A_k \sin\frac{k\pi}{l} x = f(x).$$

Очевидно, A_k являются коэффициентами Фурье функции $f(x)$ при разложении ее в ряд по синусам на промежутке $(0, l)$

$$A_k = \frac{2}{l} \int_0^l f(\eta) \sin\frac{k\pi}{l} \eta d\eta. \tag{2.43}$$

Остается подставить найденные выражения для коэффициентов в ряд (2.42).

Если функция $f(x)$ непрерывна, имеет кусочно-непрерывную производную и удовлетворяет условиям

$$f(0) = 0; \qquad f(l) = 0,$$

то ряд (2.42) определяет непрерывную при $t \geqslant 0$ функцию $u(x, t)$, являющуюся решением исходной краевой задачи.

Решение $u(x, t)$ можно записать в более компактном виде. Подставляя в выражение (2.42) значения A_k, найденные по формуле (2.43), получим

$$u(x, t) = \sum_{k=1}^{\infty} A_k e^{-\left(\frac{k\pi a}{l}\right)^2 t} \sin\frac{k\pi}{l} x = \sum_{k=1}^{\infty} \left[\frac{2}{l} \int_0^l f(\eta) \sin\frac{k\pi}{l} \eta d\eta\right] \times$$

$$\times e^{-\left(\frac{k\pi a}{l}\right)^2 t} \sin\frac{k\pi}{l} x = \int_0^l \left[\frac{2}{l} \sum_{k=1}^{\infty} e^{-\left(\frac{k\pi a}{l}\right)^2 t} \sin\frac{k\pi}{l} x \sin\frac{k\pi}{l} \eta\right] f(\eta)\, d\eta.$$

Введем обозначение

$$\Gamma(x, \eta; t) = \frac{2}{l} \sum_{k=1}^{\infty} e^{-\left(\frac{k\pi a}{l}\right)^2 t} \sin\frac{k\pi}{l} x \sin\frac{k\pi}{l} \eta.$$

В новых обозначениях функция $u(x, t)$ будет представима в виде

$$u(x, t) = \int_0^l \Gamma(x, \eta; t) f(\eta) d\eta. \qquad (2.44)$$

Функция $\Gamma(x, \eta; t)$ называется функцией мгновенного точечного источника или *функцией Грина.*

Функция $\Gamma(x, \eta; t)$ как функция от x представляет распределение температуры в стержне $0 \leqslant x \leqslant l$ в момент времени t, если температура в начальный момент времени $t = 0$ равна нулю и в этот момент в точке $x = \eta$ мгновенно выделяется некоторое количество тепла, в то время как на краях стержня все время поддерживается нулевая температура.

2. 2. 3. Стационарное распределение температуры внутри бесконечного цилиндра. Задача Дирихле для круга

Пусть задан бесконечный круговой однородный цилиндр, ось которого направлена по оси Oz, а радиус поперечного сечения равен R. Считаем, что температура внутри цилиндра уже установилась. Требуется найти закон распределения температуры внутри цилиндра в предположении, что на границе цилиндра в каждой точке поддерживается определенная, не меняющаяся с течением времени, температура. Задача, очевидно, теряет смысл, если не потребовать, чтобы в различных точках границы температура была различной. В противном случае (если на границе всюду будет поддерживаться одна и та же температура) температура внутри цилиндра будет со временем такой же, что и на границе его. Для упрощения задачи будем считать также, что распределение температуры внутри цилиндра не зависит от z. Тогда задача о распределении температуры внутри такого цилиндра может быть сформулирована следующим образом (см. § 1.5).

Найти решение $u(x, y)$ уравнения Лапласа

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \qquad (2.45)$$

удовлетворяющее краевому условию

$$u(R, \varphi) = f(\varphi) \qquad (2.46)$$

(распределение температуры на границе поперечного сечения цилиндра, т. е. на границе круга радиуса R, будет зависеть лишь от угла, который здесь обозначен через φ).

Как известно, сформулированная задача носит название внутренней задачи Дирихле для двумерного уравнения Лапласа.

Обратите еще раз внимание на постановку краевой задачи. Во-первых, бросается в глаза отсутствие начальных условий в ее формулировке. Это не должно вас удивлять, ибо рассматривается установившийся процесс, который от времени не зависит. Во-вторых, что наиболее существенно, краевые условия ранее задавались на осях декартовых координат, в которых и выписывалось само дифференциальное уравнение (и волновое, и теплопроводности). В связи с этим примененный там метод разделения переменных позволял свести краевую задачу, поставленную для уравнения в частных производных, к краевой задаче для обыкновенного дифференциального уравнения. Сейчас краевое условие задано на окружности, которая, очевидно, не совпадает с координатными линиями системы координат, в которой задано уравнение (2.45): ведь она декартова! В связи с этим примененный здесь метод разделения переменных может оказаться неэффективным. Выход из создавшегося положения будет найден, если уравнение Лапласа (2.45) записать в полярной системе координат. Теперь уже краевое условие (2.46) будет задано на координатной линии этой системы координат. Следует ожидать, что метод разделения переменных, примененный для преобразованного таким образом уравнения Лапласа, позволит нам до конца решить поставленную краевую задачу.

Будем считать, что план решения задачи уже намечен и остается приступить к его реализации.

В полярной системе координат $(r,\ \varphi)$ (с центром круга в начале координат) уравнение (2.45) будет иметь вид (см. ч. 2, стр. 79, 80)

$$\frac{\partial}{\partial r}\left(r\frac{\partial u}{\partial r}\right)+\frac{1}{r}\frac{\partial^2 u}{\partial \varphi^2}=0. \tag{2.47}$$

Теперь задача может быть сформулирована так: найти решение $u(r, \varphi)$ уравнения (2.47), удовлетворяющее краевому условию

$$u(r,\ \varphi)\,|_{r=R}=u(R,\ \varphi)=f(\varphi), \tag{2.48}$$

где $f(\varphi)$ — заданная функция.

Следуя методу разделения переменных, будем искать решение задачи в виде

$$u(r,\ \varphi)=P(r)\,\Phi(\varphi). \tag{2.49}$$

Подставляя (2.49) в (2.47), получим

$$\frac{d}{dr}\left(r\frac{dP(r)}{dr}\right)\Phi(\varphi)+\frac{1}{r}P(r)\frac{d^2\Phi(\varphi)}{d\varphi^2}=0$$

или после раскрытия скобок

$$\left(\frac{dP(r)}{dr}+r\frac{d^2P(r)}{dr^2}\right)\Phi(\varphi)+\frac{1}{r}P(r)\frac{d^2\Phi(\varphi)}{d\varphi^2}=0.$$

Теперь разделим переменные и приравняем каждую часть полученного равенства к $-\lambda$:

$$-\frac{r^2P''+rP'}{P}=\frac{\Phi''}{\Phi}=-\lambda,$$

где $\lambda=\mathrm{const}$.

Отсюда получаем два обыкновенных дифференциальных уравнения

$$P''+\frac{1}{r}P'-\frac{\lambda}{r^2}P=0 \qquad (2.50)$$

и

$$\Phi''+\lambda\Phi=0. \qquad (2.51)$$

Уравнение (2.51) нам уже встречалось неоднократно, его решение будет иметь вид

$$\Phi(\varphi)=A\cos\sqrt{\lambda}\,\varphi+B\sin\sqrt{\lambda}\,\varphi, \qquad (2.52)$$

где A и B — произвольные постоянные.

Покажем, что λ не может принимать произвольные значения. Действительно, искомая функция $u(r,\varphi)$ является периодической относительно φ с периодом 2π

$$u(r,\varphi+2\pi)=u(r,\varphi).$$

Это вытекает из характера полярной системы координат: увеличение φ на 2π возвращает точку (r,φ) в исходное положение. Но тогда и функция $\Phi(\varphi)$, как показывает равенство (2.49), обладает этим свойством, т. е.

$$\Phi(\varphi+2\pi)=\Phi(\varphi).$$

Вот это-то условие и позволяет определить λ. По формуле (2. 52) должны иметь

$$A\cos\sqrt{\lambda}(\varphi+2\pi)+B\sin\sqrt{\lambda}(\varphi+2\pi)\equiv A\cos\sqrt{\lambda}\,\varphi+B\sin\sqrt{\lambda}\,\varphi$$

или

$$A\cos(\sqrt{\lambda}\,\varphi+\sqrt{\lambda}\,2\pi)+B\sin(\sqrt{\lambda}\,\varphi+\sqrt{\lambda}\,2\pi)\equiv A\cos\sqrt{\lambda}\,\varphi+$$
$$+B\sin\sqrt{\lambda}\,\varphi.$$

В силу периодичности функции $\Phi(\varphi)$ последнее тождество возможно только тогда, когда $\sqrt{\lambda}$ есть целое число, т. е. $\lambda=k^2$, где $k=1, 2, 3, \ldots$

Таким образом, равенство (2.52) может быть переписано в виде

$$\Phi_k(\varphi) = A_k \cos k\varphi + B_k \sin k\varphi. \tag{2.53}$$

Кроме собственных значений

$$\lambda_1 = 1^2;\ \lambda_2 = 2^2;\ \ldots,\ \lambda_k = k^2,\ \ldots$$

рассматриваемая задача имеет еще одно собственное значение $\lambda_0 = 0$. Читатель без особого труда убедится, что этому собственному значению (в силу условия периодичности) соответствует только одно решение — константа. Пусть это решение будет равно $\frac{A_0}{2}$.

Итак, все собственные функции краевой задачи найдены.

Приступаем теперь к решению уравнения (2.50). После подстановки найденного значения $\lambda = k^2$, оно примет вид

$$P'' + \frac{1}{r}P' - \frac{k^2}{r^2}P = 0.$$

Полученное уравнение носит название уравнения Эйлера. Будем искать решение его в виде

$$P(r) = r^m.$$

После подстановки этого соотношения в уравнение, получим

$$m(m-1)r^{m-2} + mr^{m-2} - k^2r^{m-2} = 0$$

или

$$m(m-1) + m - k^2 = 0,$$

или

$$m^2 - k^2 = 0;\ m = \pm k,\ k > 0,$$

Следовательно, общее решение уравнения Эйлера имеет вид

$$P(r) = Cr^k + Dr^{-k},$$

где C и D — произвольные постоянные.

В полученном решении надо положить $D = 0$, иначе функция $u(r, \varphi) = P(r)\Phi(\varphi)$ при $r = 0$ (в центре круга) обращается в бесконечность и искомое решение не будет непрерывным (ограниченным). Окончательно получим

$$P_k(r) = Cr^k.$$

При $k = 0$ ($\lambda_0 = 0$) уравнение имеет решение, равное постоянной, которую можно принять за единицу.

Итак, мы получили бесчисленное множество частных решений уравнения (2.47)

$$u_0(r, \varphi) = \frac{A_0}{2}; \quad u_k(r, \varphi) = r^k (A_k \cos k\varphi + B_k \sin k\varphi),$$

$$k = 1, 2, 3, \ldots$$

Теперь, следуя принципу суперпозиции (в силу линейности и однородности уравнения Лапласа), составим решение в виде ряда

$$u(r, \varphi) = \frac{A_0}{2} + \sum_{k=1}^{\infty} r^k (A_k \cos k\varphi + B_k \sin k\varphi). \tag{2.54}$$

Коэффициенты этого ряда A_0, A_k, B_k подберем так, чтобы решение $u(r, \varphi)$ удовлетворяло краевому условию (2.48). В силу этого краевого условия получим

$$u(R, \varphi) = \frac{A_0}{2} + \sum_{k=1}^{\infty} R^k (A_k \cos k\varphi + B_k \sin k\varphi) = f(\varphi).$$

Последнее равенство (для удобства) может быть переписано следующим образом

$$f(\varphi) = \frac{A_0}{2} + \sum_{k=1}^{\infty} (A_k R^k \cos k\varphi + B_k R^k \sin k\varphi).$$

Оно представляет собой разложение функции $f(\varphi)$ в полный ряд Фурье на промежутке $[-\pi, \pi]$. Коэффициенты этого ряда A_0, A_k и B_k находятся по известным формулам (см. ч. 1, стр. 21)

$$A_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\varphi)\, d\varphi;$$

$$A_k R^k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\varphi) \cos k\varphi\, d\varphi;$$

$$B_k R^k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\varphi) \sin k\varphi\, d\varphi$$

или

$$A_k = \frac{1}{\pi R^k} \int_{-\pi}^{\pi} f(\tau) \cos k\tau d\tau;$$

$$B_k = \frac{1}{\pi R^k} \int_{-\pi}^{\pi} f(\tau) \sin k\tau d\tau.$$

Решение задачи можно считать законченным, если сейчас подставить найденные значения для коэффициентов A_0, A_k, B_k в формулу (2.54).

Мы получили, очевидно, формальное решение задачи Дирихле для круга в виде ряда (2.54).

Сейчас следовало бы доказать, что все наши рассуждения законны и ряд (2.54) действительно сходится к функции $u(r, \varphi)$, удовлетворяющей уравнению (2.47) и краевому условию (2.48). Оставляя в стороне эти доказательства, укажем лишь, что если функция $f(\varphi)$ будет непрерывна и дифференцируема, то ряд (2.54) дает непрерывное решение рассматриваемой краевой задачи.

В заключение заметим, что решение задачи может быть записано в другой форме. Получим соответствующую формулу. С этой целью подставим выражение для коэффициентов Фурье в формулу (2.54). Меняя порядок суммирования и интегрирования, получим

$$u(r, \varphi) = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\tau) \left\{ \frac{1}{2} + \sum_{k=1}^{\infty} \left(\frac{r}{R}\right)^k [\cos k\varphi \cos k\tau + \right.$$

$$\left. + \sin k\varphi \sin k\tau] \right\} d\tau = \frac{1}{\pi} \int_{-\pi}^{\pi} f(\tau) \left\{ \frac{1}{2} + \sum_{k=1}^{\infty} \left(\frac{r}{R}\right)^k \cos k(\varphi - \tau) \right\} d\tau.$$

По формуле Эйлера

$$\cos k(\varphi - \tau) = \frac{1}{2}\left(e^{jk(\varphi-\tau)} + e^{-jk(\varphi-\tau)}\right).$$

Полагая $z = \frac{r}{R}$, выражение в фигурных скобках преобразуем к виду

$$\frac{1}{2} + \sum_{k=1}^{\infty} z^k \cos k(\varphi - \tau) = \frac{1}{2} + \frac{1}{2} \sum_{k=1}^{\infty} z^k \left(e^{jk(\varphi-\tau)} + \right.$$

$$+ e^{-jk(\varphi-\tau)}) = \frac{1}{2}\left\{1 + \sum_{k=1}^{\infty}[(ze^{j(\varphi-\tau)})^k + (ze^{-j(\varphi-\tau)})^k]\right\} =$$

$$= \frac{1}{2}\left[1 + \frac{ze^{j(\varphi-\tau)}}{1 - ze^{j(\varphi-\tau)}} + \frac{ze^{-j(\varphi-\tau)}}{1 - ze^{-j(\varphi-\tau)}}\right] =$$

(здесь мы воспользовались формулой суммы членов бесконечно убывающей геометрической прогрессии $\sum\limits_{k=1}^{\infty} q^k = \frac{q}{1-q}$, $|q| < 1$)

$$= \frac{1}{2}\,\frac{1 - z^2}{1 - 2z\cos(\varphi - \tau) + z^2}.$$

Таким образом,

$$u(r, \varphi) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\tau)\,\frac{1 - \frac{r^2}{R^2}}{\frac{r^2}{R^2} - 2\frac{r}{R}\cos(\varphi - \tau) + 1}\,d\tau$$

или

$$u(r, \varphi) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(\tau)\,\frac{R^2 - r^2}{R^2 - 2Rr\cos(\varphi - \tau) + r^2}\,d\tau. \qquad (2.55)$$

Полученный интеграл называется интегралом Пуассона. Формула (2.55) позволяет выписать решение внутренней задачи Дирихле для круга в форме интеграла, зависящего от двух параметров r и φ. Интересно отметить, что представление решения в виде интеграла Пуассона позволяет значительно ослабить ограничения, накладываемые на функцию $f(\varphi)$. Теперь достаточно потребовать, чтобы эта функция была только непрерывна.

Приведенные здесь рассуждения могут быть с успехом применены и для решения других краевых задач для уравнения Лапласа (двумерного и трехмерного). При этом метод разделения переменных даст опять должный эффект, если область, на границе которой будет задаваться краевое условие, будет достаточно «хорошей»: граница ее должна совпадать с координатными линиями или поверхностями системы координат, в которой и задается уравнение Лапласа. К таким областям можно отнести прямоугольник, цилиндр, шар и др.

2. 2. 4. Общая схема метода разделения переменных

Здесь мы хотим подвести своеобразный итог сказанному выше. Оставляя в стороне вопросы строгого обоснования полученных результатов, сделаем акцент на ряде новых деталей, без которых понимание метода разделения переменных будет неполным.

Пусть задано линейное однородное уравнение второго порядка с постоянными коэффициентами

$$a_{11}\frac{\partial^2 u}{\partial x^2}+a_{22}\frac{\partial^2 u}{\partial t^2}+b_1\frac{\partial u}{\partial x}+b_2\frac{\partial u}{\partial t}+cu=0, \qquad (2.56)$$

в котором отсутствует член, содержащий смешанную производную. Как это следует из предыдущего, уравнение (2.56) может быть одного из трех типов (гиперболического, параболического или эллиптического). Безотносительно к какому типу принадлежит это уравнение, применим метод разделения переменных для решения краевой задачи, поставленной для этого уравнения.

Требуется найти решение $u(x, t)$ уравнения (2.56), удовлетворяющее однородному краевому условию

$$u(x, t)\,|_S = 0, \qquad (2.57)$$

где S — граница области, в которой отыскивается решение, и начальному условию*

$$u(x, t)\,|_{t=0} = f(x), \qquad (2.58)$$

где $f(x)$ — заданная функция.

Для решения этой задачи применим метод разделения переменных.

Первый этап решения краевой задачи по этому методу состоит в следующем.

Отыскиваются нетривиальные решения уравнения (2.56) в виде произведения

$$u(x, t) = X(x)\,T(t), \qquad (2.59)$$

удовлетворяющие только краевому условию (2.57). Подставляя произведение (2.59) в уравнение (2.56), получим

$$a_{11}X''(x)\,T(t)+a_{22}X(x)\,T''(t)+b_1X'(x)\,T(t)+$$
$$+b_2X(x)\,T'(t)+cX(x)\,T(t)=0.$$

* Если уравнение (2.56) эллиптического типа, то начальные условия отсутствуют.

Левая часть этого уравнения и однородность (!) его позволяют выполнить операцию разделения переменных

$$\frac{a_{11}X''(x)+b_1X'(x)+cX(x)}{X(x)}=\frac{-a_{22}T''(t)-b_2T'(t)}{T(t)}.$$

Последнее равенство возможно только тогда, когда эти отношения равны постоянной. Обозначим эту постоянную через $-\lambda$. Тогда получим два обыкновенных дифференциальных уравнения

$$a_{11}X''(x)+b_1X'(x)+(c+\lambda)X(x)=0; \tag{2.60}$$

$$a_{22}T''(t)+b_2T'(t)-\lambda T(t)=0. \tag{2.61}$$

Чтобы найти нетривиальные решения уравнения (2.56) вида (2.59), удовлетворяющие однородным (!) краевым условиям (2.57), необходимо получить соответствующие условия для одной из новых функций, например $X(x)$. Это удается сделать, если граница S области, на которой задаются однородные краевые условия, состоит из координатных линий системы координат xOt. В противном случае ожидаемого «разделения» краевых условий не получится, и прямое применение метода окажется невозможным. Из однородных краевых условий, наложенных на $u(x, t)$, следуют соответствующие краевые условия уже для функции $X(x)$, которые обозначим здесь через $X(x)|_S=0$.

Таким образом, мы приходим к следующей задаче о собственных значениях: найти такие значения параметра λ, при которых существуют нетривиальные решения уравнения (2.60), удовлетворяющие краевому условию $X(x)|_S=0$.

Нам уже пришлось убедиться в том, что эта краевая задача представляет самостоятельный интерес и имеет ряд трудностей (см. ч. 2). Во всяком случае она далеко не при всяком λ имеет нетривиальные решения. В связи с этим значение параметра λ_k, при котором краевая задача имеет нетривиальное решение, назовем собственным значением, а соответствующее ему решение — собственной функцией. Мы будем исходить из того, что собственные значения и собственные функции обладают рядом замечательных свойств. Перечислим основные из них.

1) Существует бесконечное множество собственных значений и соответствующих им собственных функций.

2) Каждому собственному значению λ_k соответствует только одна (с точностью до числового множителя) собственная функция $X_k(x)$.

3) Собственные функции $X_i(x)$ и $X_k(x)$, отвечающие различным собственным значениям λ_i и λ_k $(\lambda_i \neq \lambda_k)$, ортогональны на некотором промежутке (α, β); это значит, что

$$\int_{\alpha}^{\beta} X_i(x) X_k(x)\,dx = 0, \quad i \neq k.$$

При $i = k$ будем считать

$$\int_{\alpha}^{\beta} X_k^2(x)\,dx = 1.$$

В противном случае этого всегда можно добиться, умножив функцию $X_k(x)$ на соответственно подобранную постоянную, отчего она не перестанет удовлетворять уравнению (2.60) и условию $X(x)|_S = 0$.

4) ***Теорема Стеклова.**** *Всякая функция $F(x)$, имеющая непрерывную первую производную и кусочно-непрерывную вторую производную, разлагается в абсолютно и равномерно сходящийся ряд по собственным функциям $X_k(x)$ краевой задачи*

$$F(x) = \sum_{k=1}^{\infty} C_k X_k(x). \tag{2.62}$$

Последнее свойство имеет огромное принципиальное значение для решения краевых задач математической физики. Ведь любая функция, заданная в качестве начального условия, может быть теперь представлена в виде сходящегося ряда, коэффициенты которого легко определяются, если воспользоваться свойством ортогональности собственных функций. Действительно, умножим обе части равенства (2.62) на $X_i(x)$ и проинтегрируем полученное по промежутку (α, β). Тогда получим

$$\int_{\alpha}^{\beta} F(x) X_i(x)\,dx = \sum_{k=1}^{\infty} C_k \int_{\alpha}^{\beta} X_k(x) X_i(x)\,dx.$$

В сумме, стоящей справа, все члены при $k \neq i$ пропадут в силу свойства ортогональности, а коэффициент при C_i будет равен единице. Итак, мы получим

$$C_i = \int_{\alpha}^{\beta} F(x) X_i(x)\,dx.$$

---

* В. А. Стеклов (1864 — 1926) — выдающийся русский ученый-математик, давший строгое обоснование метода Фурье.

Таким образом, свойства собственных функций позволяют решить общую краевую задачу при любых начальных условиях. Это замечательный результат!

Вернемся к уравнению (2.61). Будем считать, что каждому найденному собственному значению λ_k соответствует только одно решение $T_k(t)$ этого уравнения. Тогда частными решениями уравнения (2.56), удовлетворяющими краевому условию (2.57), будут функции вида

$$u_k(x, t) = X_k(x)\, T_k(t). \tag{2.63}$$

Второй этап решения краевой задачи по методу разделения переменных состоит в составлении ряда, членами которого будут функции (2.63), а коэффициентами C_k — произвольные числа

$$u(x, t) = \sum_{k=1}^{\infty} C_k X_k(x)\, T_k(t). \tag{2.64}$$

Теперь нужно определить коэффициенты C_k этого ряда так, чтобы функция $u(x, t)$ удовлетворяла не только данному уравнению (2.56) и заданным однородным условиям (2.57) (это выполняется в силу принципа суперпозиции при любых коэффициентах C_k), но и заданному начальному условию (2.58).

Полагая в этом ряде $t = 0$ и учитывая начальное условие (2.58), получим

$$f(x) = \sum_{k=1}^{\infty} C_k T_k(0)\, X_k(x). \tag{2.65}$$

Равенство (2.65) можно рассматривать как разложение функции $f(x)$ в обобщенный ряд Фурье по собственным функциям $X_k(x)$ краевой задачи. Коэффициенты ряда $C_k T_k(0)$ находятся по известным формулам для коэффициентов Фурье, откуда и определяются величины C_k. Подставив найденные C_k в ряд (2.64), мы определим искомое решение краевой задачи.

Описанная здесь схема решения краевой задачи методом Фурье носит общий характер. В частном случае она может незначительно видоизменяться в сторону усложнения выкладок, но суть метода останется прежней.

Может показаться, что метод Фурье применяется только для решения однородных дифференциальных уравнений. Но это не так. Хотя прямое применение метода в случае неоднородного уравнения и невозможно, но эта трудность легко может быть обойдена простым приемом, который описан в следующем пункте.

Ранее подчеркивалось, что при применении метода Фурье существенно, чтобы краевые условия были однородными. Однородность краевых условий позволяет без помех решить краевую задачу на собственные значения для получающегося обыкновенного дифференциального уравнения и, используя принцип суперпозиции, составить из комбинации собственных функций решение исходного уравнения. Но при решении краевых задач часто приходится иметь дело с неоднородными краевыми условиями, что делает непосредственное применение метода Фурье невозможным. Тогда стараются свести задачу к такой, в которой краевые условия были бы однородными, после чего попадают в условия уже изученных задач. В дальнейшем мы приведем пример решения краевой задачи с неоднородными краевыми условиями.

Необходимо учитывать еще одно обстоятельство при пользовании методом разделения переменных. Искомая функция u в общем случае может зависеть не от двух, а от трех или большего числа переменных. Тогда ее надо искать в форме произведения не двух, а большего числа функций одного переменного. В дальнейшем мы продемонстрируем один из примеров подобного рода.

2. 2. 5. Вынужденные колебания закрепленной струны

Пусть задана однородная струна длиной l, закрепленная в точках $x = 0$ и $x = l$. Задача о вынужденных колебаниях такой струны может быть сформулирована следующим образом.

Найти решение $u(x, t)$ неоднородного уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + g(x, t),$$

удовлетворяющее начальным условиям

$$u(x, t)|_{t=0} = f(x); \quad \left.\frac{\partial u(x, t)}{\partial t}\right|_{t=0} = \varphi(x)$$

и однородным краевым условиям

$$u(x, t)|_{x=0} = u(0, t) = 0; \quad u(x, t)|_{x=l} = u(l, t) = 0.$$

Мы уже говорили о том, что непосредственное применение метода Фурье к задачам, в которых дифференциальное уравнение неоднородно, не представляется возможным по очевидным причинам. Тогда пользуются приемом, который встречался в теории линейных дифференциальных обыкновенных уравнений: решение отыскивают в виде суммы двух функций, одна из которых есть решение неоднородного уравнения (при удобно выбранных дополнительных условиях), а вторая — решение соответствующего однородного уравнения при заданных дополнительных условиях. Не останавливаясь на деталях, продемонстрируем этот прием на решении сформулированной выше задачи.

Отыскиваем решение $u(x, t)$ краевой задачи в виде суммы двух функций

$$u(x, t) = v(x, t) + w(x, t),$$

являющихся решениями следующих задач:

1) функция $v(x, t)$ есть решение неоднородного уравнения

$$\frac{\partial^2 v}{\partial t^2} = a^2 \frac{\partial^2 v}{\partial x^2} + g(x, t), \tag{2.66}$$

удовлетворяющее однородным начальным условиям

$$v|_{t=0} = 0; \quad \left.\frac{\partial v}{\partial t}\right|_{t=0} = 0$$

и однородным краевым условиям

$$v|_{x=0} = 0; \quad v|_{x=l} = 0;$$

2) функция $w(x, t)$ есть решение однородного уравнения

$$\frac{\partial^2 w}{\partial t^2} = a^2 \frac{\partial^2 w}{\partial x^2},$$

удовлетворяющее начальным условиям

$$w|_{t=0} = f(x); \quad \left.\frac{\partial w}{\partial t}\right|_{t=0} = \varphi(x)$$

и краевым условиям

$$w|_{x=0} = 0; \quad w|_{x=l} = 0.$$

Непосредственной проверкой убеждаемся, что составленная таким образом функция $u(x, t)$ будет решением заданной краевой задачи.

Решение $v(x, t)$ первой краевой задачи описывает вынужденные колебания струны, которые совершаются под действием внешней возмущающей силы при отсутствии начальных возмущений.

Решение $w(x, t)$ второй краевой задачи описывает, как известно, свободные колебания струны, которые происходят без действия внешней силы, а только вследствие придания точкам струны начальных отклонений и скоростей. Вторая краевая задача была подробно решена в п. 2. 2. 1, поэтому наши дальнейшие усилия будут направлены на отыскание решения неоднородного уравнения, т. е. функции $v(x, t)$.

Знание собственных функций соответствующей краевой задачи $\left(X_k(x) = \sin\frac{k\pi}{l}x\right)$, удовлетворяющих однородным краевым условиям $(X_k(0) = X_k(l) = 0)$, позволяет отыскать решение $v(x, t)$ неоднородной краевой задачи в виде ряда по этим собственным функциям

$$v(x, t) = \sum_{k=1}^{\infty} T_k(t) \sin\frac{k\pi}{l}x. \tag{2.67}$$

Теперь остается подобрать функции $T_k(t)$ таким образом, чтобы ряд (2.67) удовлетворял уравнению (2.66) и нулевым начальным условиям. Последние условия будут, очевидно, выполняться, если функции $T_k(t)$ подчинить условию

$$T_k(0) = 0; \quad T'_k(0) = 0.$$

Подставим ряд (2.67) в уравнение (2.66). Тогда получим

$$\sum_{k=1}^{\infty}\left[T''_k(t) + \frac{k^2\pi^2a^2}{l^2}T_k(t)\right]\sin\frac{k\pi}{l}x = g(x, t).$$

Фиксируя t, разложим функцию $g(x, t)$ в ряд Фурье по синусам на промежутке $(0, l)$

$$g(x, t) = \sum_{k=1}^{\infty} g_k(t) \sin\frac{k\pi}{l}x,$$

где

$$g_k(t) = \frac{2}{l}\int_0^l g(\eta, t)\sin\frac{k\pi}{l}\eta\, d\eta.$$

Подставляем это разложение в предыдущее равенство:

$$\sum_{k=1}^{\infty}\left[T''_k(t) + \frac{k^2\pi^2a^2}{l^2}T_k(t)\right]\sin\frac{k\pi}{l}x = \sum_{k=1}^{\infty} g_k(t)\sin\frac{k\pi}{l}x.$$

Последнее равенство возможно только тогда, когда равны коэффициенты разложений его правой и левой частей. Отсюда получим обыкновенное дифференциальное уравнение для определения неизвестных функций $T_k(t)$

$$T''_k(t) + \frac{k^2\pi^2a^2}{l^2}T_k(t) = g_k(t), \tag{2.68}$$

к которому нужно присоединить известные уже начальные условия $T_k(0)=0$; $T'_k(0)=0$.

Задача Коши, к которой свелось нахождение функции $v(x, t)$, может быть решена обычным приемом вариации произвольных постоянных, известным из общего курса высшей математики. Можно воспользоваться здесь и методом функции Коши (см. ч. 2, стр. 169). Известно, что частное решение неоднородного уравнения (2.68), удовлетворяющее нулевым начальным условиям, задается формулой

$$T_k(t)=\int_0^t K(t, s)\, g_k(s)\, ds,$$

где $K(t, s)$ — функция Коши соответствующего однородного уравнения

$$T''_k(t)+\frac{k^2\pi^2a^2}{l^2}T_k(t)=0.$$

Но функция Коши этого уравнения уже построена (см. пример 1 ч. 2, стр. 172) и равна

$$K(t, s)=\frac{l}{k\pi a}\sin\frac{k\pi a}{l}(t-s).$$

Таким образом,

$$T_k(t)=\frac{l}{k\pi a}\int_0^t \sin\frac{k\pi a}{l}(t-s)\, g_k(s)\, ds. \tag{2.69}$$

Подставляя найденные отсюда значения $T_k(t)$ в ряд (2.67), получим искомое решение $v(x, t)$.

Запишем решение заданной неоднородной краевой задачи, учитывая полученные здесь результаты и результаты п. 2. 2. 1:

$$u(x, t)=\sum_{k=1}^{\infty}T_k(t)\sin\frac{k\pi}{l}x+\sum_{k=1}^{\infty}\left(A_k\cos\frac{k\pi a}{l}t+\right.$$
$$\left.+B_k\sin\frac{k\pi a}{l}t\right)\sin\frac{k\pi}{l}x,$$

где коэффициенты $T_k(t)$ определяются по формуле (2.69), а

$$A_k=\frac{2}{l}\int_0^l f(x)\sin\frac{k\pi}{l}x\,dx;\quad B_k=\frac{2}{k\pi a}\int_0^l \varphi(x)\sin\frac{k\pi}{l}x\,dx.$$

Пример 3. На струну длиной l постоянно действует внешняя возмущающая сила, равная силе тяжести G. Найти закон колебания струны, если начальное отклонение и начальная скорость равны нулю, а концы струны закреплены.

Решение. Поставленную задачу можно сформулировать следующим образом. Найти решение $u(x, t)$ неоднородного волнового уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} - G,$$

удовлетворяющее однородным начальным условиям

$$u(x, t)\big|_{t=0} = 0; \quad \left.\frac{\partial u(x, t)}{\partial t}\right|_{t=0} = 0$$

и однородным краевым условиям

$$u(0, t) = 0; \quad u(l, t) = 0.$$

Решение этой задачи дается рядом

$$u(x, t) = \sum_{k=1}^{\infty} T_k(t) \sin \frac{k\pi}{l} x,$$

коэффициенты которого $T_k(t)$ определяются по формуле

$$T_k(t) = \frac{l}{k\pi a} \int_0^t \sin \frac{k\pi a}{l} (t-s)\, g_k(s)\, ds,$$

где

$$g_k(t) = -\frac{2}{l} \int_0^l G \sin \frac{k\pi}{l} \eta\, d\eta =$$

$$= -\frac{2G}{k\pi}(1 - \cos k\pi) = \begin{cases} 0, & \text{если } k \text{ — четное:} \\ -\frac{4G}{k\pi}, & \text{если } k \text{ — нечетное.} \end{cases}$$

Тогда

$$T_k(t) = -\frac{4Gl}{k^2\pi^2 a} \int_0^t \sin \frac{k\pi a}{l} (t-s)\, ds = -\frac{4Gl^2}{k^3\pi^3 a^2}\left(1 - \cos \frac{k\pi a}{l} t\right).$$

Подставим теперь найденное выражение для $T_k(t)$ в определяющий решение ряд

$$u(x, t) = \sum_{k=1}^{\infty} -\frac{4Gl^2}{k^3\pi^3 a^2}\left(1 - \cos \frac{k\pi a}{l} t\right) \sin \frac{k\pi}{l} x =$$

$$= -\frac{4Gl^2}{\pi^3 a^2} \sum_{k=1}^{\infty} \frac{1}{k^3}\left(1 - \cos \frac{k\pi a}{l} t\right) \sin \frac{k\pi}{l} x,$$

где $k = 1, 3, 5, \ldots, 2n-1, \ldots$

Анализ полученного решения может быть проведен точно так же, как мы это делали в п. 2.2.1.

2. 2. 6. Вынужденные колебания струны с подвижными концами

Наше представление о методе Фурье было бы неполным, если бы мы не привели пример на решение так называемой общей краевой задачи, когда не только само дифференциальное уравнение неоднородно, но неоднородны и краевые условия.

Рассмотрим вынужденные колебания однородной струны под действием внешней силы, когда концы струны не закреплены, а двигаются по заданному закону. Задача ставится таким образом.

Найти решение $u(x, t)$ неоднородного волнового уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2} + g(x, t), \tag{2.70}$$

удовлетворяющее начальным условиям

$$u(x, t)\,|_{t=0} = f(x); \quad \left.\frac{\partial u(x, t)}{\partial t}\right|_{t=0} = \varphi(x) \tag{2.71}$$

и неоднородным краевым условиям

$$u(x, t)\,|_{x=0} = \psi_1(t); \; u(x, t)\,|_{x=l} = \psi_2(t). \tag{2.72}$$

Идея решения этой задачи заключается в том, чтобы свести ее к задаче с однородными краевыми условиями, которую мы уже умеем решать для неоднородного уравнения.

С целью реализации этой идеи возьмем какую-нибудь функцию $p(x, t)$, которая удовлетворяет заданным краевым условиям (2.72). Обычно стараются выбрать эту функцию так, чтобы она была достаточно проста, но дифференцируема. Легко сообразить, что в рассматриваемом случае функцию $p(x, t)$ можно выбрать в следующем виде

$$p(x, t) = \psi_1(t) + \frac{\psi_2(t) - \psi_1(t)}{l}\, x. \tag{2.73}$$

Действительно, здесь $p(x, t)\,|_{x=0} = \psi_1(t)$ и $p(x, t)\,|_{x=l} = \psi_2(t)$. Тогда функция

$$v(x, t) = u(x, t) - p(x, t) \tag{2.74}$$

будет удовлетворять уже однородным краевым условиям. (Заметим, что если краевые условия (2.72) будут другого вида, то и функцию $p(x, t)$ следует подбирать в ином виде.)

Для функции $v(x, t)$, удовлетворяющей однородным краевым условиям, составим краевую задачу в силу заданного дифференциального уравнения (2.70) и начальных условий (2.71). Для этого подставим (2.74) в уравнение (2.70). Тогда получим

$$\frac{\partial^2 v}{\partial t^2} = a^2 \frac{\partial^2 v}{\partial x^2} + a^2 \frac{\partial^2 p}{\partial x^2} - \frac{\partial^2 p}{\partial t^2} + g(x, t)$$

или, с учетом (2.73)

$$\frac{\partial^2 v}{\partial t^2} = a^2 \frac{\partial^2 v}{\partial x^2} + \bar{g}(x, t),$$

где

$$\bar{g}(x, t) = g(x, t) - \psi_1''(t) - [\psi_2''(t) - \psi_1''(t)] \frac{x}{l}.$$

Найдем теперь начальные условия для функции $v(x, t)$

$$v(x, 0) = u(x, 0) - p(x, 0) = f(x) - \psi_1(0) - \frac{\psi_2(0) - \psi_1(0)}{l} x = \bar{f}(x),$$

$$v_t'(x, 0) = u_t'(x, 0) - p_t'(x, 0) = \varphi(x) - \psi_1'(0) - \frac{\psi_2'(0) - \psi_1'(0)}{l} x = \bar{\varphi}(x).$$

Таким образом, мы приходим к следующей краевой задаче для функции $v(x, t)$: найти решение неоднородного уравнения

$$\frac{\partial^2 v}{\partial t^2} = a^2 \frac{\partial^2 v}{\partial x^2} + \bar{g}(x, t),$$

удовлетворяющее начальным условиям

$$v(x, 0) = \bar{f}(x); \quad v_t'(x, 0) = \bar{\varphi}(x)$$

и однородным краевым условиям

$$v(0, t) = 0; \quad v(l, t) = 0.$$

Решив эту задачу методом, изложенным в п. 2.2.5 и подставив найденное значение $v(x, t)$ в (2.74), определим искомое решение $u(x, t)$.

Пример 4. Однородная струна, заданная на отрезке $(0, l)$, закреплена в левом конце, а правый конец колеблется по закону $A \sin \omega t$. Найти закон движения струны, считая, что начальное отклонение и начальная скорость в каждой внутренней точке струны равны нулю.

Решение. Задачу можно сформулировать следующим образом. Найти решение $u(x, t)$ однородного волнового уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\partial^2 u}{\partial x^2}, \tag{2.75}$$

удовлетворяющее начальным

$$u(x, 0) = 0; \quad \frac{\partial u(x, 0)}{\partial t} = 0 \tag{2.76}$$

и краевым условиям

$$u(0, t) = 0; \quad u(l, t) = A \sin \omega t. \tag{2.77}$$

Решение этой задачи будем искать методом, описанным выше. Сначала подберем функцию $p(x, t)$, удовлетворяющую заданным краевым условиям. В качестве такой функции здесь удобно взять следующую

$$p(x, t) = \frac{x}{l} A \sin \omega t.$$

Тогда функция

$$v(x, t) = u(x, t) - p(x, t)$$

будет удовлетворять однородным краевым условиям

$$v(0, t) = 0; \quad v(l, t) = 0.$$

Согласно теории, получим дифференциальное уравнение для функции $v(x, t)$. Оно будет иметь вид

$$\frac{\partial^2 v}{\partial t^2} = a^2 \frac{\partial^2 v}{\partial x^2} + \frac{x}{l} A \omega^2 \sin \omega t.$$

Находим начальные условия для функции $v(x, t)$:

$$v(x, 0) = u(x, 0) - p(x, 0) = 0; \quad v_t'(x, 0) = u_t'(x, 0) - p_t'(x, 0) = -\frac{A\omega x}{l}.$$

Таким образом, для того, чтобы найти функцию $v(x, t)$, надо решить полученное неоднородное уравнение при однородных краевых условиях и новых начальных условиях. После проведенных вычислений получим решение поставленной краевой задачи

$$u(x, t) = \frac{2A\omega}{\pi} \sum_{k=1}^{\infty} \frac{(-1)^k}{k\left(\frac{k^2\pi^2 a^2}{l^2} - \omega^2\right)} \left(\frac{ka\pi}{l} \sin \frac{ak\pi}{l} t - \omega \sin \omega t\right) \times$$

$$\times \sin \frac{k\pi}{l} x - \frac{x}{l} A \sin \omega t.$$

Формальное применение метода сведения краевой задачи с неоднородными краевыми условиями к задаче с однородными условиями, хотя и привело к цели, но считать его здесь оправданным сомнительно. Мы выиграли в одном, но проиграли в другом: до преобразования дифференциальное уравнение было однородным, после — стало неоднородным; начальные условия были нулевыми, после преобразования стали неоднородными. Это обстоятельство заставляет нас искать более простой путь для решения заданной задачи, который бы учитывал и характер заданного уравнения, и характер начальных условий.

Будем искать решение задачи в виде суммы

$$u(x, t) = v(x, t) + w(x, t),$$

где $w(x, t)$ — решение однородного уравнения (2.75), удовлетворяющее только краевым условиям (2.77), а $v(x, t)$ — решение того же уравнения, удовлетворяющее однородным краевым условиям

$$v(x, t)|_{x=0} = 0; \quad v(x, t)|_{x=l} = 0$$

и начальным условиям

$$v(x,t)|_{t=0} = -w(x,t)|_{t=0} = f(x); \quad \left.\frac{\partial v(x,t)}{\partial t}\right|_{t=0} = -\left.\frac{\partial w(x,t)}{\partial t}\right|_{t=0} = \varphi(x).$$

Легко проверить, что таким образом определенная функция $u(x, t)$ будет удовлетворять уравнению, начальным и краевым условиям. Учитывая характер краевых условий, ищем решение $w(x, t)$ в виде

$$w(x,t) = X(x)\sin \omega t.$$

После определения начальных условий, которым должна удовлетворять функция $v(x, t)$, надо решить краевую задачу для этой функции с однородными краевыми условиями методом разделения переменных. Это предлагаем осуществить читателю, проделав все необходимые здесь вычисления.

2. 2. 7. Свободные колебания прямоугольной мембраны

На примере свободных колебаний мембраны мы хотим проиллюстрировать применение метода Фурье в многомерном случае, т. е. когда искомая функция u зависит не от двух переменных, а от трех. И в этом случае схема метода разделения переменных остается в силе, хотя и появляются новые трудности, связанные в основном с увеличением объема вычислений. Заметим, что в соответствующих местах мы будем пользоваться готовыми результатами, полученными в предыдущих пунктах параграфа.

Известно, что двумерное однородное волновое уравнение описывает свободные колебания однородной мембраны. (Под мембраной мы понимаем упругую свободно изгибающуюся пленку.) Краевая задача для уравнения колебания прямоугольной мембраны, расположенной в состоянии покоя в плоскости xOy, занимающей область S: $0 \leqslant x \leqslant l$; $0 \leqslant y \leqslant m$, и закрепленной на границе этой области (рис. 2.13), ставится следующим образом. (Сравните с постановкой задачи для уравнения колебания закрепленной струны.)

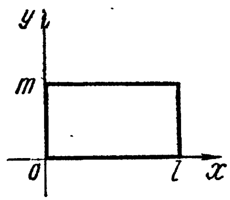

Рис. 2.13

Найти решение $u(x, y, t)$ однородного уравнения

$$\frac{\partial^2 u}{\partial t^2} = a^2\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right), \tag{2.78}$$

удовлетворяющее начальным условиям

$$u(x,y,t)|_{t=0} = f(x,y); \quad \left.\frac{\partial u(x,y,t)}{\partial t}\right|_{t=0} = \varphi(x,y) \tag{2.79}$$

и однородным краевым условиям

$$u(x, y, t)|_{x=0} = 0; \quad u(x, y, t)_{x=l} = 0; \quad u(x, y, t)|_{y=0} = 0;$$
$$u(x, y, t)|_{y=m} = 0. \qquad (2.80)$$

Выполняя первый этап метода разделения переменных, ищем нетривиальные решения уравнения (2.78), удовлетворяющие краевым условиям (2.80), в виде произведения трех функций, каждая из которых зависит только от одной переменной

$$u(x, y, t) = X(x) Y(y) T(t). \qquad (2.81)$$

Сразу находим краевые условия для искомых функций $X(x)$ и $Y(y)$. Последовательное использование условий (2.80) в равенстве (2.81) дает

$$X(0) = 0; \quad X(l) = 0; \quad Y(0) = 0; \quad Y(m) = 0.$$

Подставим теперь произведение (2.81) в уравнение (2.78)

$$XYT'' = a^2(X''YT + XY''T).$$

(Здесь аргументы у функций опущены для сокращения записи.) Разделим переменные, но так, чтобы одна из функций, для которой у нас известны краевые условия, была в одной части равенства, а оставшиеся функции — в другой. Пусть эта функция будет $X(x)$

$$XYT'' = a^2X''YT + a^2XY''T,$$
$$XYT'' - a^2XY''T = a^2X''YT.$$

Поделив обе части последнего равенства на произведение a^2XYT, мы добьемся желаемого

$$\frac{YT'' - a^2Y''T}{a^2YT} = \frac{X''}{X}.$$

Левая часть полученного равенства не зависит от переменной x, а правая часть — от переменных y и t. Поэтому это равенство может иметь место только при условии, что и левая и правая его части равны постоянной величине

$$\frac{YT'' - a^2Y''T}{a^2YT} = \frac{X''}{X} = -\lambda, \quad \lambda > 0.$$

Отсюда

$$X'' + \lambda X = 0,$$
$$YT'' - a^2Y''T + \lambda a^2YT = 0. \qquad (2.82)$$

Как известно, краевая задача

$$X'' + \lambda X = 0;$$
$$X(0) = 0; \; X(l) = 0$$

имеет собственные значения

$$\lambda_1 = \frac{\pi^2}{l^2}; \; \lambda_2 = \frac{2^2\pi^2}{l^2}; \; \ldots; \; \lambda_k = \frac{k^2\pi^2}{l^2}; \; \ldots$$

и соответствующие им собственные функции

$$X_1(x) = \sin\frac{\pi}{l}x; \; X_2(x) = \sin\frac{2\pi}{l}x; \; \ldots;$$
$$X_k(x) = \sin\frac{k\pi}{l}x, \; \ldots$$

Подставим найденное значение $\lambda = \lambda_k = \frac{k^2\pi^2}{l^2}$ в уравнение (2.82) и еще раз разделим переменные. Последовательно получаем

$$YT'' - a^2Y''T + \frac{k^2\pi^2a^2}{l^2}YT = 0,$$

$$\frac{T'' + \frac{k^2\pi^2a^2}{l^2}T}{a^2T} = \frac{Y''}{Y} = -\mu.$$

Отсюда

$$Y'' + \mu Y = 0,$$
$$T'' + \left(\frac{k^2\pi^2a^2}{l^2} + \mu a^2\right)T = 0. \qquad (2.83)$$

Краевая задача

$$Y'' + \mu Y = 0;$$
$$Y(0) = 0; \; Y(m) = 0$$

имеет собственные значения

$$\mu_1 = \frac{\pi^2}{m^2}; \; \mu_2 = \frac{2^2\pi^2}{m^2}; \; \ldots; \; \mu_n = \frac{n^2\pi^2}{m^2}, \; \ldots$$

и соответствующие им собственные функции

$$Y_1(y) = \sin\frac{\pi}{m}y; \; Y_2(y) = \sin\frac{2\pi}{m}y; \; \ldots;$$
$$Y_n(y) = \sin\frac{n\pi}{m}y, \; \ldots$$

Для определения функций $T(t)$ нужно подставить $\mu = \mu_n = \frac{n^2\pi^2}{m^2}$ в уравнение (2.83)

$$T''(t) + a^2\pi^2\left(\frac{k^2}{l^2} + \frac{n^2}{m^2}\right)T(t) = 0.$$

Решения этого уравнения обозначим $T_{k,n}(t)$, ибо каждой паре номеров k и n соответствует свое решение. Произвольные постоянные, входящие в общее решение, обозначим $A_{k,n}$ и $B_{k,n}$. Тогда

$$T_{k,n}(t) = A_{k,n}\cos\omega_{k,n}t + B_{k,n}\sin\omega_{k,n}t,$$

где величины $\omega_{k,n} = a\pi\sqrt{\frac{k^2}{l^2} + \frac{n^2}{m^2}}$ являются собственными частотами колебаний мембраны.

Найденные значения $X_k(x)$, $T_{k,n}(t)$ и $Y_n(y)$ подставим в равенство (2.81). Тогда мы получим нетривиальные решения уравнения (2.78), удовлетворяющие краевым условиям (2.80),

$$u_{k,n}(x, y, t) = (A_{k,n}\cos\omega_{k,n}t + B_{k,n}\sin\omega_{k,n}t)\sin\frac{k\pi}{l}x\sin\frac{n\pi}{m}y.$$

Переходим ко второму этапу метода Фурье.

Составим ряд из полученных частных решений заданного уравнения и подберем коэффициенты этого ряда так, чтобы сумма его удовлетворяла начальным условиям (2.79). В связи с тем, что каждое частное решение зависит от двух индексов k и n, придется составлять двойной ряд, когда индексы суммирования k и n пробегают независимо друг от друга все целые положительные числа

$$u(x, y, t) = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty}(A_{k,n}\cos\omega_{k,n}t + B_{k,n}\sin\omega_{k,n}t)\sin\frac{k\pi}{l}x\sin\frac{n\pi}{m}y. \tag{2.84}$$

Положим в этом равенстве $t = 0$ и учтем первое начальное условие

$$u(x, y, 0) = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty}A_{k,n}\sin\frac{k\pi}{l}x\sin\frac{n\pi}{m}y = f(x, y). \tag{2.85}$$

Дифференцируя ряд (2.84) по t и учитывая второе начальное условие, получим при $t = 0$

$$u_t'(x, y, 0) = \sum_{k=1}^{\infty}\sum_{n=1}^{\infty}\omega_{k,n}B_{k,n}\sin\frac{k\pi}{l}x\sin\frac{n\pi}{m}y = \varphi(x, y). \tag{2.86}$$

Равенства (2.85) и (2.86) представляют разложение функций двух переменных $f(x, y)$ и $\varphi(x, y)$ в так называемые двойные ряды Фурье. Наша задача сейчас заключается в том, чтобы получить формулы, по которым вычисляются коэффициенты $A_{k,n}$ и $B_{k,n}$ этих рядов. Но для этого придется убедиться в ортогональности системы функций $\sin\frac{k\pi}{l}x\cdot\sin\frac{n\pi}{m}y$ двух переменных в области S: $0\leqslant x\leqslant l$; $0\leqslant y\leqslant m$. По аналогии со случаем одной переменной систему этих функций будем называть ортогональной в области S, если двойной интеграл по этой области от произведения двух различных функций системы равен нулю. Итак,

$$\int_0^l\int_0^m \sin\frac{k_1\pi}{l}x\sin\frac{n_1\pi}{m}y\cdot\sin\frac{k_2\pi}{l}x\sin\frac{n_2\pi}{m}y\,dx\,dy =$$

$$=\left(\int_0^l \sin\frac{k_1\pi}{l}x\sin\frac{k_2\pi}{l}x\,dx\right)\left(\int_0^m \sin\frac{n_1\pi}{m}y\sin\frac{n_2\pi}{m}y\,dy\right)=0,$$

ибо в свою очередь система функций $\sin\frac{k\pi}{l}x$ ортогональна на промежутке $[0, l]$, а система функций $\sin\frac{n\pi}{m}y$ ортогональна на промежутке $[0, m]$, что обращает каждую из скобок в нуль.

Заметим, что в случае $k_1=k_2$ и $n_1=n_2$

$$\int_0^l\int_0^m \sin^2\frac{k_1\pi}{l}x\sin^2\frac{n_1\pi}{m}y\,dx\,dy =$$

$$=\int_0^l \sin^2\frac{k_1\pi}{l}x\,dx\int_0^m \sin^2\frac{n_1\pi}{m}y\,dy=\frac{l}{2}\cdot\frac{m}{2}=\frac{lm}{4}.$$

Вспоминая методику получения формул для вычисления коэффициентов Фурье в случае функции одной переменной (см. ч. 1, стр. 19), совершенно аналогично найдем соответствующие формулы для функций двух переменных. Они будут иметь вид

$$A_{k,n}=\frac{4}{lm}\int_0^l\int_0^m f(x, y)\sin\frac{k\pi}{l}x\sin\frac{n\pi}{m}y\,dx\,dy;$$

$$\omega_{k,n}B_{k,n}=\frac{4}{lm}\int_0^l\int_0^m \varphi(x, y)\sin\frac{k\pi}{l}x\sin\frac{n\pi}{m}y\,dx\,dy$$

или

$$B_{k,n} = \frac{4}{lm\omega_{n,k}} \int\limits_0^l \int\limits_0^m \varphi(x, y) \sin\frac{k\pi}{l} x \sin\frac{n\pi}{m} y \, dx\, dy.$$

Подставляя найденные значения $A_{k,n}$ и $B_{k,n}$ в ряд (2.84), получим искомое решение дифференциального уравнения (2.78), удовлетворяющее начальным условиям (2.79) и краевым условиям (2.80).

Мы рассмотрели, пожалуй, самый простой случай краевой задачи для двумерного волнового уравнения. Если мембрана имеет другую форму, например круглую, то краевая задача должна быть сформулирована уже не в декартовых координатах, а в полярных. Если же мембрана имеет произвольную форму, то, как это следует из предыдущего, непосредственное применение метода разделения переменных будет просто невозможным.

2. 2. 8. Заключение

Цель предыдущего анализа состояла в том, чтобы продемонстрировать метод Фурье при решении простейших краевых задач. Нас не должно смущать то обстоятельство, что подавляющее большинство рассмотренных здесь примеров было связано с волновым уравнением. В равной степени все сказанное выше относится и к уравнению теплопроводности, а частично — и к уравнению Лапласа.

Простота метода Фурье, его проникновение в физическую суть задачи делают этот метод одним из самых распространенных при решении задач математической физики.

§ 2. 3. МЕТОД ФУНКЦИИ ГРИНА

Изучение метода разделения переменных поставило перед нами целый ряд новых проблем. Одной из самых острых следует считать проблему решения краевой задачи с произвольными краевыми условиями. Ведь метод Фурье дает должный эффект только тогда, когда краевые условия задаются на линиях (поверхностях), являющихся координатными линиями (поверхностями) системы координат, в которой выписывается само дифференциальное уравнение. Читателю ясно, что только одно это обстоятельство значительно сужает класс задач, которые могут быть решены методом разделения переменных. Особенно остро этот недостаток метода Фурье ощущается при решении краевых задач для уравнения Лапласа, являющегося одним из самых важных уравнений математической физики.

Здесь будет изложен метод, который позволяет найти в явном виде решение краевой задачи при произвольных краевых условиях. Желая лучше подчеркнуть особенности нового метода, мы будем рассматривать лишь краевые задачи для уравнения Лапласа. Если читателя заинтересует применение метода функции Грина для решения краевых задач волнового уравнения или уравнения теплопроводности, то он сможет познакомиться с соответствующим материалом по более полному курсу [35].

Уравнение Лапласа привлекает наше внимание и по другой причине. Замечательные свойства его решений, вытекающие из ряда интегральных теорем, изученных нами в теории поля (см. ч. 2), их простота и наглядность делают раздел, посвященный уравнению Лапласа, наиболее изящным и стройным разделом всей теории уравнений в частных производных второго порядка. С элементами этой теории мы и познакомимся в этом параграфе.

Метод функции Грина уже знаком читателю по ч. 2 пособия. Он был приспособлен для решения краевых задач линейных обыкновенных дифференциальных уравнений. Основной результат, который был там получен, состоял в следующем: решение полуоднородной краевой задачи задавалось формулой

$$y(t)=\int_a^b \Gamma(t,\ s)\,f(s)\,ds, \tag{2.87}$$

где $\Gamma(t,\ s)$ — функция Грина рассматриваемой задачи $L[y]=f(t)$, $l_i[y;\ a,\ b]=0$ ($L[y]$ — линейный дифференциальный оператор, l_i — линейные формы, определяющие краевые условия). Таким образом, решение краевой задачи можно выписать в явном виде, если только известна функция Грина $\Gamma(t,\ s)$ этой задачи. Но определение функции Грина, ее построение представляет большие трудности, если даже можно доказать, что такая функция существует. Вообще говоря, каждая краевая задача имеет свою функцию Грина. В этом вы могли убедиться при рассмотрении примеров на построение функции Грина обыкновенных дифференциальных уравнений.

Наша конечная цель и будет состоять в том, чтобы получить формулу, аналогичную (2.87), но только уже для уравнения Лапласа

$$\Delta u=\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}=\varphi(x,\ y,\ z)$$

с заданными краевыми условиями. Построив, правда не без труда, функцию Грина для простейших областей, на границе которых будут задаваться краевые условия, мы тем самым решим рассматриваемую краевую задачу для этих областей.

Почти все результаты будут вытекать из трех интегральных формул, которые были получены ранее в теории поля (см. ч. 2, стр. 77). Ввиду важности этих формул для изучения дальнейшего материала пособия, мы еще раз напоминаем их читателю.

Непосредственным следствием формулы Остроградского является следующее равенство

$$\iiint_W v\Delta u dV + \iiint_W \left(\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\frac{\partial v}{\partial y} + \frac{\partial u}{\partial z}\frac{\partial v}{\partial z}\right) dV = \iint_S v\frac{\partial u}{\partial n} dS, \tag{2.88}$$

где W — некоторая замкнутая ограниченная область трехмерного пространства с кусочно-гладкой границей S, в которой заданы непрерывные функции $u(M)$ и $v(M)$, имеющие непрерывные частные производные до второго порядка включительно. Вычитая из (2.88) равенство, в котором мы поменяли местами u и v, получим формулу

$$\iiint_W (v\Delta u - u\Delta v)\, dV = \iint_S \left(v\frac{\partial u}{\partial n} - u\frac{\partial v}{\partial n}\right) dS, \tag{2.89}$$

из которой, в частности при $v(M) \equiv 1$, вытекает равенство

$$\iiint_W \Delta u dV = \iint_S \frac{\partial u}{\partial n} dS. \tag{2.90}$$

Формулы (2.88) и (2.89) называются *формулами Грина* и играют решающую роль для построения теории уравнения Лапласа.

2. 3. 1. Гармонические функции и их основные свойства

Каждый из нас хорошо знаком со свойствами линейной функции

$$u(x) = kx + b, \tag{2.91}$$

являющейся простейшей элементарной функцией. Одним из характерных свойств этой функции является наличие у нее непрерывной второй производной, обращающейся в нуль, т. е.

$$\frac{d^2u}{dx^2} = 0.$$

Таким образом, линейная функция является решением однородного одномерного уравнения Лапласа. Легко обнаруживаются и другие свойства этой функции.

а) Линейная функция не может принимать свое наибольшее либо наименьшее значение внутри некоторого отрезка $[c, d]$, на котором она задана. Как известно, наибольшее и наименьшее значения линейной на конечном отрезке функции достигаются на концах отрезка, за исключением того случая, когда $u(x) = \text{const}$ (рис. 2.14).

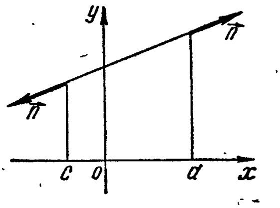

Рис. 2.14

Геометрическая интерпретация помогает нам высказать еще одно свойство рассматриваемой функции.

б) Линейная на конечном отрезке $[c, d]$ функция однозначно определяется своими значениями на концах этого отрезка, т. е. значения функции $u(x)$ можно найти для всех x из (c, d), зная ее значения в точках $x = c$ и $x = d$. Действительно,

$$u(x) = \frac{d-x}{d-c}\,u(c) + \frac{x-c}{d-c}\,u(d).$$

в) На рис. 2.14 изображено два направления внешней нормали $\vec{n}$ к границе отрезка, являющегося геометрическим изображением линейной функции (2.91) на промежутке $[c, d]$. Легко подсчитать, что сумма значений первых производных этой функции по направлению $\vec{n}$ в концах промежутка будет равна нулю.

г) Значение линейной функции (2.91) в центре любого промежутка $[c, d]$ равно ее среднему значению на $[c, d]$, а также среднему арифметическому значений функции (2.91) на концах этого промежутка, т. е.

$$u\left(\frac{c+d}{2}\right) = \frac{1}{d-c}\int_c^d u(x)\,dx = \frac{u(c)+u(d)}{2}.$$

Читателю предлагается самостоятельно проверить последнее свойство линейной функции.

Мы не спроста перечисляем очевидные свойства линейной функции. Дело в том, что аналогичными свойствами обладают и функции, являющиеся решениями однородного уравнения Лапласа. Ограничиваясь здесь случаем двух и трех переменных, перейдем к подробному изучению свойств этих функций.

Определение. *Функция $u(x, y, z)$ называется гармонической в конечной области W, если она в этой области дважды* *непрерывно дифференцируема и удовлетворяет однородному уравнению Лапласа*

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0.$$

Функция $u(x, y, z)$ называется гармонической в бесконечной области, если она обладает еще одним свойством: стремится к нулю при стремлении точки $M(x, y, z)$ в бесконечность.

Приведем простейшие примеры гармонических функций.

1) Функция $u(x, y, z) \equiv 1$ гармоническая в любой конечной области. В бесконечной области она гармонической не будет. (Почему?)

2) Функция $u(x, y) = \frac{x}{x^2 + y^2}$ гармоническая в любой области, которая не содержит начала координат. Действительно,

$$\frac{\partial^2 u}{\partial x^2} = \frac{2x(x^2 - 3y^2)}{(x^2 + y^2)^3}; \quad \frac{\partial^2 u}{\partial y^2} = -\frac{2x(x^2 - 3y^2)}{(x^2 + y^2)^3}.$$

Суммируя, находим

$$\Delta u = \Delta\left(\frac{x}{x^2 + y^2}\right) = 0.$$

3) Функция $u(x, y) = x^2 - y^2$ гармоническая в любой конечной области, а вот функция $u(x, y) = x^2 + y^2$ не является гармонической ни в какой области, так как она не удовлетворяет однородному уравнению Лапласа

$$\Delta u = \Delta(x^2 + y^2) \neq 0.$$

Хотелось бы выделить особо из огромного множества гармонических функций такие, которые имеют, так сказать, теоретическое значение и играют важную роль для построения теории уравнения Лапласа.

Пусть $M(x, y, z)$ и $M_0(x_0, y_0, z_0)$ — две точки трехмерного пространства. Обозначим через r расстояние между ними

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}.$$

Рассмотрим функцию

$$v(M) = v(x, y, z) = \frac{1}{r}, \tag{2.92}$$

считая точку M_0 фиксированной. Очевидно, функция $v(x, y, z)$ разрывна в точке M_0, т. е. при $x = x_0$, $y = y_0$, $z = z_0$.

Имеет место следующая теорема.

Теорема 1. *Функция (2.92) есть гармоническая в любой области, не содержащей точку $M_0(x_0, y_0, z_0)$.*

Доказательство. Покажем, что функция (2.92) удовлетворяет уравнению Лапласа $\Delta v = 0$. Имеем

$$\frac{\partial r}{\partial x} = \frac{x - x_0}{r}; \quad \frac{\partial r}{\partial y} = \frac{y - y_0}{r}; \quad \frac{\partial r}{\partial z} = \frac{z - z_0}{r}.$$

Тогда

$$\frac{\partial v}{\partial x} = -\frac{1}{r^2}\frac{\partial r}{\partial x} = -\frac{x - x_0}{r^3}$$

и

$$\frac{\partial^2 v}{\partial x^2} = -\frac{1}{r^3} + \frac{3(x - x_0)^2}{r^5}.$$

Совершенно аналогично находим

$$\frac{\partial^2 v}{\partial y^2} = -\frac{1}{r^3} + \frac{3(y - y_0)^2}{r^5};$$

$$\frac{\partial^2 v}{\partial z^2} = -\frac{1}{r^3} + \frac{3(z - z_0)^2}{r^5}.$$

Суммируя полученные частные производные второго порядка, получим

$$\Delta v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2} = -\frac{3}{r^3} + \frac{3}{r^5}[(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2] = -\frac{3}{r^3} + \frac{3r^2}{r^5} = 0,$$

что и требовалось доказать.

Функцию $v(x, y, z) = \frac{1}{r}$ называют *фундаментальным решением* уравнения Лапласа

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0. \tag{2.93}$$

Можно показать, что для двумерного уравнения Лапласа

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{2.94}$$

гармонической функцией, зависящей от величины r, будет служить следующая функция

$$v(M) = v(x, y) = \ln\frac{1}{r}, \tag{2.95}$$

где $r = \sqrt{(x - x_0)^2 + (y - y_0)^2}$.

Функцию (2.95) называют *фундаментальным решением* уравнения Лапласа (2.94).

Уместно заметить, что среди функций $u(M)$, значения которых зависят только от расстояния r между M и некоторой фиксированной точкой M_0, других гармонических функций, кроме функций вида $Av(M) + B$, не существует.

Теперь мы переходим к доказательству основной теоремы, из которой затем будет получен ряд следствий, имеющих для нас первостепенное значение.

Теорема 2. *Если функция $u(x, y, z)$ непрерывна, имеет непрерывные частные производные первого и второго порядков в замкнутой области W, то имеет место формула*

$$u(x_0, y_0, z_0) = \frac{1}{4\pi}\iint\limits_S \left(\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial \frac{1}{r}}{\partial n}\right) dS - \frac{1}{4\pi}\iiint\limits_W \frac{1}{r}\Delta u dV, \tag{2.96}$$

где r — расстояние от фиксированной точки $M_0(x_0, y_0, z_0)$, лежащей внутри W, до переменной точки $M(x, y, z)$, n — внешняя нормаль к поверхности S.

Доказательство. Для доказательства теоремы мы хотим воспользоваться формулой Грина (2.89), полагая в ней $v = \frac{1}{r}$. Но эта функция обращается в бесконечность, если точка M совпадает с точкой M_0. Поэтому непосредственное применение формулы Грина к области W здесь невозможно. Придется вместо области W рассмотреть область W_ε, полученную из замкнутой области W удалением некоторого шара достаточно малого радиуса ε с центром в точке M_0 (рис. 2.15). Но тогда и граница области изменится: граница области W_ε будет состоять теперь из двух поверхностей: поверхности S, которая ограничивала область W, и сферической поверхности S_ε. В связи с этим вместо одного интеграла по поверхности S в правой части равенства (2.89) придется написать сумму двух интегралов по S и по S_ε. В связи с тем, что $v = \frac{1}{r}$ в области W_ε — гармоническая функция ($\Delta v = 0$), мы получим

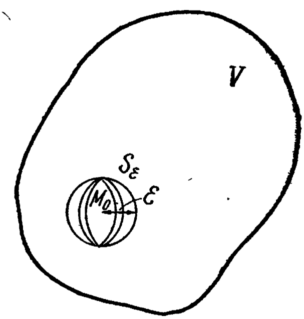

Рис. 2.15

$$\iiint\limits_{W_\varepsilon} \frac{1}{r}\Delta u dV = \iint\limits_S \left(\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial \frac{1}{r}}{\partial n}\right) dS + \iint\limits_{S_\varepsilon}\left(\frac{1}{r}\frac{\partial u}{\partial n} - u\frac{\partial \frac{1}{r}}{\partial n}\right) dS.$$

Перейдем теперь в полученном равенстве к пределу, когда $\varepsilon \to 0$. Тогда слева получим точно такой же интеграл, но уже распространенный на всю область W. Первое слагаемое правой части нисколько не изменится, так как оно от ε не зависит. Для доказательства теоремы достаточно будет показать, что второе слагаемое этой части равенства стремится к пределу $4\pi u(x_0, y_0, z_0)$.

Перепишем это слагаемое в виде

$$\iint\limits_{S_\varepsilon} \left(\frac{1}{r} \frac{\partial u}{\partial n} - u \frac{\partial \frac{1}{r}}{\partial n} \right) dS = \iint\limits_{S_\varepsilon} \frac{1}{r} \frac{\partial u}{\partial n} dS - \iint\limits_{S_\varepsilon} u \frac{\partial \frac{1}{r}}{\partial n} dS.$$

Так как переменная точка $M(x, y, z)$ находится на поверхности шара, то для всех точек поверхности S_ε величина r имеет постоянное значение, равное радиусу шара ε. С другой стороны, первые частные производные функции $u(x, y, z)$ по условию теоремы непрерывны в замкнутой области W. Это означает, что они будут ограничены в этой области вместе с производной $\frac{\partial u}{\partial n}$. Другими словами существует такое число N, что $\left| \frac{\partial u}{\partial n} \right| < N$. Тогда получим

$$\left| \iint\limits_{S_\varepsilon} \frac{1}{r} \frac{\partial u}{\partial n} dS \right| \leqslant \iint\limits_{S_\varepsilon} \frac{1}{\varepsilon} \left| \frac{\partial u}{\partial n} \right| dS < \frac{1}{\varepsilon} N \iint\limits_{S_\varepsilon} dS = \frac{1}{\varepsilon} N \, 4\pi\varepsilon^2 =$$
$$= 4\pi N\varepsilon \to 0 \text{ при } \varepsilon \to 0.$$

Переходим к оценке второго интеграла по поверхности S_ε. Для этого нам нужно предварительно вычислить производную $\frac{\partial \frac{1}{r}}{\partial n}$. В рассматриваемом случае нормаль n направлена в прямо противоположную сторону по отношению к радиусу шара (нормаль берется внешняя!). Поэтому

$$\left. \frac{\partial \frac{1}{r}}{\partial n} \right|_{S_\varepsilon} = - \left. \frac{\partial \frac{1}{r}}{\partial r} \right|_{S_\varepsilon} = \frac{1}{\varepsilon^2},$$

$$\iint\limits_{S_\varepsilon} u \frac{\partial \frac{1}{r}}{\partial n} dS = \frac{1}{\varepsilon^2} \iint\limits_{S_\varepsilon} u \, dS = \frac{1}{\varepsilon^2} u(M_\varepsilon) 4\pi\varepsilon^2.$$

Здесь мы воспользовались теоремой о среднем значении для поверхностного интеграла.

Окончательно

$$\lim_{\varepsilon\to 0}\iint_{S_\varepsilon} u\frac{\partial\frac{1}{r}}{\partial n}\,dS=\lim_{\varepsilon\to 0}4\pi\,u(M_\varepsilon)=4\pi\,u(M_0).$$

На этом доказательство теоремы заканчивается.

Если исходить из соответствующего фундаментального решения для двумерного уравнения Лапласа (2.94) $v=\ln\frac{1}{r}$ и повторить предшествующий вывод, то можно получить формулу

$$u(x_0,\ y_0)=\frac{1}{2\pi}\oint_l\left(\ln\frac{1}{r}\,\frac{\partial u}{\partial n}-u\frac{\partial\ln\frac{1}{r}}{\partial n}\right)dl-\frac{1}{2\pi}\iint_D\ln\frac{1}{r}\,\Delta u d\sigma,\tag{2.97}$$

которая дает интегральное представление функции $u(x,\ y)$ в случае двух независимых переменных.

Если функция u гармоническая в замкнутой области W, т. е. $\Delta u=0$, то формула (2.96) приобретает особенно простой вид

$$u(x_0,\ y_0,\ z_0)=\frac{1}{4\pi}\iint_S\left(\frac{1}{r}\,\frac{\partial u}{\partial n}-u\frac{\partial\frac{1}{r}}{\partial n}\right)dS.\tag{2.98}$$

В случае гармонической функции u двух независимых переменных получаем

$$u(x_0,\ y_0)=\frac{1}{2\pi}\oint_l\left(\ln\frac{1}{r}\,\frac{\partial u}{\partial n}-u\frac{\partial\ln\frac{1}{r}}{\partial n}\right)dl.\tag{2.99}$$

Формулы (2.98) и (2.99) дают возможность вычислить значение гармонической функции u в любой внутренней точке M_0 некоторой замкнутой области, если известны ее значения и значения ее нормальной производной на границе этой области. Если вспомнить, что задача Дирихле состоит в том, чтобы найти гармоническую функцию в некоторой области, принимающую на границе этой области заданное значение, то полученные формулы (2.98) и (2.99) вносят большой вклад в дело решения краевых задач для уравнения Лапласа.

Перечислим основные свойства гармонических функций.

1) Интеграл от нормальной производной гармонической функции по границе области равен нулю, т. е.

$$\iint_S\frac{\partial u}{\partial n}\,dS=0.\tag{2.100}$$

Этот результат является следствием формулы (2.90), где надо положить $\Delta u = 0$ (функция u — гармоническая).

Можно показать, что условие (2.100) является необходимым и достаточным для того, чтобы непрерывная функция $u(M)$, имеющая в области W непрерывные частные производные до второго порядка включительно, была гармонической в этой области.

Линейная функция (2.91), будучи гармонической, обладает, как мы установили, свойством, выраженным в пункте «в», где сумма значений первых производных по направлению $\vec{n}$ в концах промежутка равна нулю.

2) Всякие две функции, гармонические в замкнутой ограниченной области W с кусочно-гладкой границей, принимающие одинаковые значения на границе S, совпадают всюду в W.

Это свойство вытекает уже из формулы (2.88). Полагая в ней $v(M) = u(M)$, получим

$$\iiint\limits_W \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2\right] dV = \iint\limits_S u \frac{\partial u}{\partial n} dS.$$

Но если $\frac{\partial u}{\partial n} = 0$ в точках на границе S, то $\frac{\partial u}{\partial x} = 0$; $\frac{\partial u}{\partial y} = 0$ и $\frac{\partial u}{\partial z} = 0$ всюду в W, т. е. $u(M) = \text{const}$ в W. Отсюда следует, что если $u(M) = 0$ на S, то $u(M) \equiv 0$ в W. Но разность двух гармонических функций является функцией гармонической. Применяя к этой разности полученный результат, мы убедимся в справедливости сформулированного выше следствия. Для линейной функции (2.91) это свойство, как легко увидеть, соответствует пункту «б».

3) ***Теорема Гаусса*** *Значение гармонической функции в центре шара равно среднему арифметическому ее значений на границе этого шара.*

Этот результат получается из формулы (2.98), если предположить, что область W — шар радиуса R с центром в точке $M_0(x_0, y_0, z_0)$, а S — граница этой области, т. е. сферическая поверхность. Действительно, в рассматриваемом случае внешняя нормаль к поверхности шара направлена по радиусу. Поэтому производная по нормали равна производной по радиусу, т. е.

$$\frac{\partial \frac{1}{r}}{\partial n} = \frac{\partial \frac{1}{r}}{\partial r} = -\frac{1}{r^2}.$$

Если еще учесть, что всюду на поверхности шара $r = R$, то формулу (2.98) можно переписать в виде

$$u(x_0, y_0, z_0) = \frac{1}{4\pi}\iint\limits_S \frac{1}{R}\frac{\partial u}{dn}\,dS - \frac{1}{4\pi}\iint\limits_S\left(-\frac{1}{R^2}\right)u\,dS =$$

$$= \frac{1}{4\pi R}\iint\limits_S \frac{\partial u}{\partial n}\,dS + \frac{1}{4\pi R^2}\iint\limits_S u\,dS = \frac{1}{4\pi R^2}\iint\limits_S u\,dS,$$

ибо интеграл $\iint\limits_S \frac{\partial u}{\partial n}\,dS = 0$ (по первому свойству гармонических функций). Итак, окончательно

$$u(x_0, y_0, z_0) = \frac{1}{4\pi R^2}\iint\limits_S u\,dS. \tag{2.101}$$

(Напоминаем, что здесь $S = 4\pi R^2$ — площадь поверхности шара радиуса R. Поэтому правая часть равенства (2.101) называется средним арифметическим значений функции u, которые просуммированы интегралом на поверхности S.)

Заметим, что равенство (2.101) теряет силу, если функция $u(M)$ не является гармонической в центре $M_0(x_0, y_0, z_0)$ шара.

Интересно отметить, что свойство гармонических функций, выраженное теоремой Гаусса, является характеристическим и может быть принято в качестве определения гармоничности.

4) ***Принцип максимума и минимума для гармонических функций.***

Если функция $u(M)$ является гармонической в области W, то она не может нигде внутри этой области принимать свое наибольшее либо наименьшее в этой области значение, за исключением того случая, когда $u(M) = \text{const}$.

Напоминаем, что непрерывная функция $u(M)$, каковой является гармоническая функция по определению, в замкнутой области W должна достигать своего наибольшего и наименьшего значения (свойство непрерывной функции в замкнутой области). Теорема утверждает, что внутри области W эта функция не может принимать своего наибольшего и наименьшего значения. Значит, этих значений гармоническая функция достигает на границе рассматриваемой области. В этом и состоит смысл сформулированной теоремы.

Доказательство этого предложения проведем от противного. Предположим, что функция $u(M)$ достигает своего наибольшего значения в точке M_0, лежащей внутри области W, т. е. $u(M) \leqslant u(M_0)$ для всех точек M из W. Так как точка M_0 есть внутренняя точка области W, то она принадлежит этой области вместе с некоторым шаром, центр которого в точке M_0

и радиус равен ε. Покажем, что во всех точках M, принадлежащих границе S_ε шара, значение функции $u(M)$ совпадает с $u(M_0)$. Для этого воспользуемся теоремой Гаусса, выраженной равенством

$$u(M_0) = \frac{1}{4\pi\varepsilon^2} \iint\limits_{S_\varepsilon} u(M)\, dS.$$

Имеем очевидное равенство

$$u(M_0) = \frac{1}{4\pi\varepsilon^2} \iint\limits_{S_\varepsilon} u(M_0)\, dS.$$

Вычитая теперь почленно из первого равенства второе, получим

$$0 = \frac{1}{4\pi\varepsilon^2} \iint\limits_{S_\varepsilon} [u(M) - u(M_0)]\, dS,$$

откуда

$$\iint\limits_{S_\varepsilon} [u(M) - u(M_0)]\, dS = 0.$$

Подынтегральная функция непрерывна и по предположению принимает неположительные значения. Но в этом случае интеграл может обратиться в нуль только при условии равенства нулю подынтегральной функции, т. е.

$$u(M) - u(M_0) = 0$$

или $u(M) = u(M_0)$ на S_ε.

Таким образом, во всех точках на границе шара значение функции $u(M)$ равно $u(M_0)$. А так как все приведенные вычисления можно применять к шару любого радиуса, меньшего ε, то

$$u(M) = u(M_0)$$

всюду внутри шара радиуса ε, т. е. в этом шаре функция $u(M)$ постоянна. Используя этот результат, легко показать, что аналогичным свойством будет обладать функция $u(M)$ не только в окрестности точки M_0, но и всюду внутри W.

Из принципа максимума и мунимума вытекает, что если $u(M)$ — функция гармоническая внутри замкнутой области W, а k_1 и k_2 — соответственно нижняя и верхняя грани функции $u(M)$ на границе S области W, то всюду в области W выполняются неравенства

$$k_1 \leqslant u(M) \leqslant k_2,$$

ибо в противном случае наибольшее или наименьшее значение этой функции в области W достигалось бы в некоторой точке M_0, лежащей внутри W.

Отсюда сразу следует свойство 2 гармонических функций. Действительно, если $k_1 = k_2 = 0$, то $u(M) = 0$ в области W. Другими словами, всякие две гармонические функции тождественно совпадают всюду в W, если они принимают одинаковые значения на границе S.

В заключение приведем несколько примеров гармонических функций, которые часто встречаются в приложениях.

Запас гармонических функций, которым мы с вами располагаем, очевидно, не исчерпывает всего множества функций, обладающих свойством гармоничности. Действительно, используя фундаментальные решения однородного уравнения Лапласа $u(x, y) = \ln\frac{1}{r}$ и $u(x, y, z) = \frac{1}{r}$, где r — расстояние между точками M_0 и M, мы, учитывая линейность этого уравнения, можем построить бесчисленное множество гармонических функций, являющихся линейной комбинацией уже известных гармонических функций. Отсюда, в частности, следует, что какой бы ни была система конечного числа точек M_1, M_2, ..., M_k и чисел e_1, e_2, ..., e_k, функция

$$u(x, y) = u(M) = e_1 \ln\frac{1}{r_{MM_1}} + e_2 \ln\frac{1}{r_{MM_2}} + \ldots + e_k \ln\frac{1}{r_{MM_k}} \tag{2.102}$$

и функция

$$u(x, y, z) = u(M) = \frac{e_1}{r_{MM_1}} + \frac{e_2}{r_{MM_2}} + \ldots + \frac{e_k}{r_{MM_k}} = \sum_{i=1}^{k} \frac{e_i}{r_{MM_i}} \tag{2.103}$$

будут гармоническими внутри любой области W, не содержащей точек M_i, $i = 1, 2, 3, \ldots, k$.

Значение функции (2.103) в точке M имеет физический смысл потенциала тяготения (потенциала электростатического поля) в этой точке, создаваемого системой точечных масс (зарядов) e_1, e_2, ..., e_k, сосредоточенных в точках M_1, M_2, ..., M_k. В связи с этим, функцию (2.103) называют *ньютоновским потенциалом* конечной системы точечных масс, а функцию (2.102) по аналогии — *логарифмическим потенциалом* такой системы.

Пусть в области W распределены массы (заряды) с плотностью $\rho(P)$. В малом объеме dV_p, содержащем точку P, заключена масса (заряд) величины $\rho(P)dV_p$. Тогда потенциал поля, создаваемого этой массой (зарядом), равен

$$\frac{\rho(P)}{r_{MP}}dV_p.$$

Потенциал поля, создаваемого массами (зарядами), содержащимися в области W, по аналогии с предыдущим, равен

$$u(M)=\iiint\limits_{W}\frac{\rho(P)}{r_{MP}}dV_p. \tag{2.104}$$

Интеграл (2.104) называется *объемным потенциалом*. Для двумерного пространства (плоскости) объемный потенциал имеет вид

$$u(M)=\iint\limits_{D}\rho(P)\ln\left(\frac{1}{r_{MP}}\right)d\sigma_p. \tag{2.105}$$

Интегралы (2.104) и (2.105) представляют собой интегралы с параметром. Используя теорему о дифференцировании по параметру под знаком интеграла, для любой точки M, не принадлежащей области W, получим

$$\Delta u(M)=\iiint\limits_{W}\Delta\left(\frac{1}{r_{MP}}\right)\rho(P)dV_p=0.$$

Для двумерной области D аналогично

$$\Delta u(M)=\iint\limits_{D}\Delta\left(\ln\frac{1}{r_{MP}}\right)\rho(P)d\sigma_p=0.$$

Таким образом, функции (2.104) и (2.105) являются гармоническими во всех точках, внешних к области W.

Наряду с объемным потенциалом можно привести и другие примеры гармонических функций подобного рода, рассматривая соответствующие поверхностные интегралы.

Из теории функций комплексной переменной известна тесная связь гармонических функций двух переменных с аналитическими функциями комплексной переменной. Вещественная и мнимая части $u(x, y)$ и $v(x, y)$ аналитической функции $f(z)$ комплексной переменной $z=x+jy$ в силу условий Коши—Римана $\frac{\partial u}{\partial x}=\frac{\partial v}{\partial y}$; $\frac{\partial u}{\partial y}=-\frac{\partial v}{\partial x}$ (см. ч. 1, стр. 123) являются гармоническими функциями. Действительно, дифференцируя первое из этих равенств по x, а второе — по y и складывая, получим

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0.$$

Дифференцируя первое из тех же равенств по y, а второе — по x и вычитая, найдем

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0.$$

При этом для любой гармонической функции $u(x, y)$ можно всегда указать вторую гармоническую функцию $v(x, y)$, которая называется *сопряженной* по отношению к $u(x, y)$, такую, что функция

$$f(z) = u(x, y) + jv(x, y)$$

будет аналитической. Таким образом, все многообразие гармонических функций двух переменных определяется запасом аналитических функций комплексной переменной.

2. 3. 2. Задачи Дирихле и Неймана для уравнения Лапласа

Вернемся к проблеме интегрирования уравнения Лапласа, а точнее, к вопросу решения краевых задач для этого уравнения. В § 1.5 были перечислены основные краевые задачи для уравнения Лапласа. Сейчас мы напомним некоторые относящиеся сюда факты.

При постановке краевых задач приходится иметь дело с двумя типами областей: конечными и бесконечными. В дальнейшем, вне зависимости от того к какому типу принадлежит рассматриваемая область, границу ее будем считать конечной, состоящей из конечного числа кусочно-гладких поверхностей*.

Краевая задача для уравнения Лапласа называется внутренней, если искомая функция определяется в конечной области, и внешней, если эта функция должна быть определена в бесконечной области.

Важнейшими краевыми задачами для уравнения Лапласа являются задача Дирихле и задача Неймана.

1) Внутренняя задача Дирихле. Найти функцию $u(x, y, z)$, гармоническую в конечной области W и принимающую на границе S этой области заданные значения

$$u(x, y, z)|_S = f(x, y, z),$$

где $f(x, y, z)$ — непрерывная на S функция.

2) Внутренняя задача Неймана. Найти функцию $u(x, y, z)$, гармоническую в конечной области W, такую, чтобы ее произ-

*) Поверхность называется *гладкой*, если в каждой ее точке можно провести касательную плоскость, непрерывно изменяющую свое положение при непрерывном движении точки касания. Поверхность, состоящая из нескольких гладких кусков, называется *кусочно-гладкой*.

водная $\frac{\partial u}{\partial n}$ по направлению внешней нормали в каждой точке поверхности S равнялась значению в этой точке заданной функции $\varphi(x, y, z)$:

$$\left.\frac{\partial u(x, y, z)}{\partial n}\right|_S = \varphi(x, y, z),$$

где $\varphi(x, y, z)$ — непрерывная на S функция.

Внешние задачи отличаются от соответствующих внутренних только тем, что гармоническая функция отыскивается в бесконечной области. Как известно, в этом случае гармоническая функция $u(M)$ будет стремиться к нулю при стремлении точки M в бесконечность.

Непосредственным следствием рассматриваемых в предыдущем пункте свойств гармонической функции является следующая теорема.

Теорема. *Решение задачи Дирихле (как внутренней, так и внешней) единственно.*

В справедливости этой теоремы для внутренней задачи Дирихле вы убедитесь, привлекая второе свойство гармонической функции: всякие две функции, гармонические в замкнутой ограниченной области W, принимающие одинаковые значения на границе S, совпадают всюду в W.

Доказательство единственности внешней задачи Дирихле мы здесь опускаем.

Обратимся теперь к задаче Неймана.

Ранее мы установили (§ 1.5), что задача Неймана в общем случае неразрешима. Действительно, по первому свойству гармонической функции $\iint\limits_S \frac{\partial u}{\partial n} dS = 0$. Отсюда следует, что необходимым условием разрешимости сформулированной задачи Неймана является выполнение условия

$$\iint\limits_S \varphi(x, y, z)\, dS = 0.$$

О единственности решения задачи Неймана для уравнения Лапласа можно сказать следующее: два решения внутренней задачи Неймана могут отличаться только на постоянное слагаемое. Действительно, если функция $u(x, y, z)$ решает эту краевую задачу, то в силу линейности уравнения Лапласа и характера краевых условий, функция $u(x, y, z) + C$, где C — произвольная постоянная, решает ту же задачу. Используя интегральную формулу (2.88), можно показать, что выражение

$u(x, y, z)+C$ исчерпывает все решения внутренней задачи Неймана.

В случае рассматриваемого трехмерного пространства решение внешней задачи Неймана единственно.

2. 3. 3. Метод функции Грина. Функция Грина для уравнения Лапласа

Излагая метод функции Грина при решении задачи Дирихле для уравнения Лапласа, мы будем следовать тому плану, который был намечен в ч. 2 пособия при решении краевых задач для обыкновенного дифференциального уравнения.

Рассмотрим сначала вспомогательную задачу, которая приведет нас к определению функции Грина, играющей решающую роль в образовании искомого решения.

Эта задача ставится так: найти функцию $\gamma_{M_0}=\gamma_{M_0}(x, y, z)$, гармоническую в конечной области W и принимающую на границе S этой области заданные значения, равные $-\frac{1}{4\pi r}$.

Можно доказать, что решение этой задачи Дирихле со специальным заданием значений искомой функции на границе рассматриваемой области существует и единственно. Предположим, что задача решена и функция $\gamma_{M_0}(M)$ найдена.

Введение функции $\gamma_{M_0}(M)$ помогает получить формулу, выражающую значения гармонической функции внутри области через ее значения на границе.

Пусть $u(M)$ — гармоническая функция в замкнутой области W, ограниченной поверхностью S. Тогда имеет место уже знакомая нам формула

$$u(M_0)=\frac{1}{4\pi}\iint\limits_S\left(\frac{1}{r}\frac{\partial u}{\partial n}-u\frac{\partial\frac{1}{r}}{\partial n}\right)dS, \tag{2.98}$$

выражающая значения функции $u(M)$ внутри области через ее значения и значения ее нормальной производной на границе S области W.

Теперь применим формулу Грина (2.89) к паре гармонических функций $u(M)$ и $\gamma_{M_0}(M)$. Тогда получим

$$\iint\limits_S\left[u(M)\frac{\partial\gamma_{M_0}(M)}{\partial n}-\gamma_{M_0}(M)\frac{\partial u(M)}{\partial n}\right]dS=0$$

или, в силу краевых условий для функции $\gamma_{M_0}(M)$,

$$\iint\limits_S\left[u(M)\frac{\partial\gamma_{M_0}(M)}{\partial n}+\frac{1}{4\pi r}\frac{\partial u(M)}{\partial n}\right]dS=0.$$

Вычитая это равенство из (2.98), избавимся под знаком интеграла от производной $\frac{\partial u}{\partial n}$:

$$u(M_0) = \iint\limits_S \left(\frac{1}{4\pi r} \frac{\partial u}{\partial n} - \frac{1}{4\pi} u \frac{\partial \frac{1}{r}}{\partial n} - u \frac{\partial \gamma_{M_0}(M)}{\partial n} - \right.$$

$$\left. - \frac{1}{4\pi r} \frac{\partial u}{\partial n} \right) dS = - \iint\limits_S u(M) \frac{\partial}{\partial n} \left[\frac{1}{4\pi r} + \gamma_{M_0}(M) \right] dS.$$

Итак, мы добились цели: формула

$$u(M_0) = - \iint\limits_S u(M) \frac{\partial}{\partial n} \left[\frac{1}{4\pi r} + \gamma_{M_0}(M) \right] dS \qquad (2.106)$$

выражает значение гармонической функции внутри области через ее значения на границе этой области.

Введем в рассмотрение новую функцию $\Gamma(M, M_0)$.

Определение. *Функцией Грина задачи Дирихле для уравнения Лапласа называется функция $\Gamma(M, M_0)$, удовлетворяющая следующим условиям:*

1) функция $\Gamma(M, M_0)$ есть гармоническая функция точки M во всей области W, исключая точку M_0, где она обращается в бесконечность;

2) функция $\Gamma(M, M_0)$ как функция точки M удовлетворяет условию

$$\Gamma(M, M_0)|_S = 0;$$

3) в области W функция $\Gamma(M, M_0)$ допускает представление

$$\Gamma(M, M_0) = \frac{1}{4\pi r} + \gamma_{M_0}(M),$$

где r есть расстояние между M и M_0, а $\gamma_{M_0}(M)$ — гармоническая функция везде внутри W.

Из этого определения следует, что о функции Грина для области W можно говорить лишь тогда, когда существует функция $\gamma_{M_0}(M)$ с указанными выше свойствами.

С помощью функции Грина равенство (2.106) приводится к следующему виду

$$u(M_0) = -\frac{1}{4\pi} \iint\limits_S u(M) \frac{\partial \Gamma(M, M_0)}{\partial n} dS.$$

Предположим, что функция Грина известна. Тогда решение внутренней задачи Дирихле, если оно существует, дается формулой

$$u(M_0) = -\frac{1}{4\pi}\iint\limits_S f(M)\frac{\partial\Gamma(M, M_0)}{\partial n}\,dS. \qquad (2.107)$$

Подробное исследование формулы (2.107) показывает, что для поверхностей, обладающих некоторыми специальными свойствами (поверхностей Ляпунова) [35, стр. 346], эта формула представляет решение задачи Дирихле при любом выборе непрерывной функции $f(M)$, входящей в краевое условие.

Формула (2.107) не открывает широкую дорогу для решения задачи Дирихле, ибо определение функции Грина $\Gamma(M, M_0)$ часто оказывается не менее трудным, чем решение задачи Дирихле с общими значениями на границе. Однако в некоторых случаях удается определить $\Gamma(M, M_0)$ непосредственно, и тогда формула (2.107) дает такой простой метод решения задачи Дирихле, с которым не идут ни в какое сравнение все другие методы решения этой задачи.

Теперь мы хотим выполнить обещанное и получить формулу, дающую решение неоднородного уравнения Лапласа с однородными краевыми условиями, аналогичную формуле (2.87).

Вместо задачи о функции $u(M)$, гармонической внутри области W и совпадающей на границе S с функцией $f(M)$, можно рассмотреть эквивалентную в некотором роде задачу о дважды непрерывно дифференцируемой в W функции $v(M)$, удовлетворяющей неоднородному уравнению Лапласа

$$\Delta v = \varphi(M)$$

и обращающейся в нуль на границе S.

Для этого введем разность

$$v(M) = u(M) - f(M),$$

где $f(M)$ заданная, дважды непрерывно дифференцируемая функция. Подставляя эту разность в однородное уравнение Лапласа (2.93), получим

$$\Delta v = \Delta u - \Delta f = -\Delta f,$$

ибо функция $u(M)$ — гармоническая.

Если сейчас оператор Лапласа от заданной функции $f(M)$ обозначить через $\varphi(M)$ и учесть, что

$$v(M)|_S = u(M)|_S - f(M)|_S = f(M) - f(M) = 0,$$

то полуоднородная задача Дирихле для уравнения Лапласа может быть сформулирована следующим образом.

Найти решение $v(M)$ неоднородного уравнения Лапласа (уравнение Пуассона) в конечной области W

$$\Delta v = \varphi(M), \tag{2.108}$$

удовлетворяющее однородным краевым условиям на границе S этой области

$$v(M)|_S = 0. \tag{2.109}$$

Для решения этой задачи воспользуемся методом функции Грина.

Интегральную форму (2.96), справедливую для любой дважды непрерывно дифференцируемой в W функции, применим для функции $v(M)$: Тогда получим

$$v(M_0) = \iint\limits_S \left(\frac{1}{4\pi r}\frac{\partial v}{\partial n} - \frac{1}{4\pi} v \frac{\partial \frac{1}{r}}{\partial n}\right) dS - \iiint\limits_W \frac{1}{4\pi r} \Delta v dV. \tag{2.110}$$

Желая ввести в выражение, стоящее в правой части этого равенства, функцию Грина

$$\Gamma(M, M_0) = \frac{1}{4\pi r} + \gamma_{M_0}(M),$$

воспользуемся формулой Грина (2.89), применяя ее к паре функций $\gamma_{M_0}(M)$ и $v(M)$. Получим

$$\iiint\limits_W - \gamma_{M_0}(M) \Delta v dV = \iint\limits_S \left(v \frac{\partial \gamma_{M_0}(M)}{\partial n} - \gamma_{M_0}(M) \frac{\partial v}{\partial n}\right) dS$$

или

$$0 = \iint\limits_S \left(v \frac{\partial \gamma_{M_0}(M)}{\partial n} - \gamma_{M_0}(M) \frac{\partial v}{\partial n}\right) dS + \iiint\limits_W \gamma_{M_0}(M) \Delta v dV.$$

Вычитая из (2.110) последнее равенство, находим

$$v(M_0) = \iint\limits_S \left(\frac{1}{4\pi r} + \gamma_{M_0}(M)\right) \frac{\partial v}{\partial n} - v \frac{\partial}{\partial n}\left(\frac{1}{4\pi r} \gamma_{M_0}(M)\right) dS -$$

$$- \iiint\limits_W \left(\frac{1}{4\pi r} + \gamma_{M_0}(M)\right) \Delta v dV,$$

что может быть переписано с учетом определения функции Грина

$$v(M_0) = \iint\limits_S \left\{\Gamma(M, M_0)\frac{\partial v}{\partial n} - v(M)\frac{\partial \Gamma(M, M_0)}{\partial n}\right\} dS -$$

$$- \iiint\limits_W \Gamma(M, M_0)\, \Delta v dV.$$

В связи с тем, что функция $v(M)$ обращается в нуль на границе S области W, а $\Gamma(M, M_0) = 0$ для точек M на S, последнее равенство значительно упростится

$$v(M_0) = - \iiint\limits_W \Gamma(M, M_0)\, \Delta v dV.$$

Но $\Delta v = \varphi(M)$, что окончательно дает

$$v(M_0) = - \iiint\limits_W \Gamma(M, M_0)\, \varphi(M)\, dV. \qquad (2.111)$$

Интеграл в правой части формулы (2.111), как функция от M_0, рассматриваемая внутри области W, в общем случае для любой непрерывно дифференцируемой в замкнутой области W функции $\varphi(M)$ есть решение задачи, т. е. для точек M_0 внутри W он удовлетворяет уравнению Пуассона (2.108) и стремится к нулю, когда точка M_0 стремится к точке M на границе области W, т. е. удовлетворяет краевому условию (2.109).

Формула (2.111) является очевидным обобщением формулы (2.87) на случай трехмерного пространства.

2. 3. 4. Решение задачи Дирихле для шара

В заключение параграфа приведем пример на построение функции Грина для одной из простейших областей — шара.

Пусть дан шар радиуса R с центром O в начале координат. Требуется найти функцию $u(M)$, гармоническую в шаре и удовлетворяющую краевому условию

$$u(M)\,|_S = f(M),$$

где S — граница шара и $f(M)$ — функция, заданная и непрерывная на сфере S.

Желая воспользоваться формулой (2.107), построим функцию Грина этой задачи.

Пусть M_0 — какая-нибудь фиксированная точка внутри шара. Найдем теперь точку M_1, сопряженную точке M_0

(рис. 2.16), т. е. такую точку, которая лежит на продолжении луча OM_0 и для которой имеет место равенство

$$OM_0 \cdot OM_1 = R^2. \qquad (2.112)$$

Обозначим через ρ расстояние от центра шара до M_0, а через r и r_1 расстояния от некоторой точки M, лежащей внутри или на поверхности шара, соответственно до точек M_0 и M_1. Найдем соотношение между r и r_1, когда точка M находится на поверхности шара. Для этого рассмотрим треугольники OM_0M и OM_1M. Они подобны, так как имеют один общий угол с вершиной в точке O и пропорциональные стороны, заключающие этот угол. Действительно, равенство (2.112) можно переписать в виде

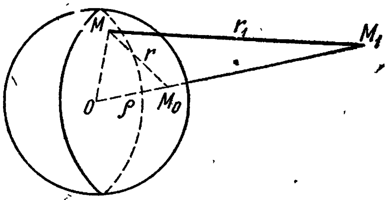

Рис. 2.16

$$\frac{OM_0}{R} = \frac{R}{OM_1},$$

т. е. $\frac{OM_0}{OM} = \frac{OM}{OM_1}$, что и подтверждает сказанное выше: $\triangle OM_0M \sim \triangle OM_1M$. Из подобия этих треугольников следует, что

$$\frac{MM_0}{MM_1} = \frac{OM_0}{OM} \text{ или } \frac{r}{r_1} = \frac{\rho}{R}.$$

Последняя пропорция может быть переписана в следующем виде

$$\frac{1}{r} - \frac{R}{\rho} \frac{1}{r_1} = 0. \qquad (2.113)$$

Напоминаем: равенство (2.113) получено в предположении, что точка M находится на поверхности S шара.

Легко догадаться, что за искомую функцию Грина здесь можно взять

$$\Gamma(M, M_0) = \frac{1}{4\pi r} - \frac{1}{4\pi} \frac{R}{\rho} \frac{1}{r_1}. \qquad (2.114)$$

Действительно, эта функция обладает всеми свойствами функции Грина (см. определение):

1) функция (2.114) есть гармоническая функция точки M во всем шаре, исключая точку M_0, где она обращается в бесконечность. В гармоничности этой функции можно убедиться

непосредственным дифференцированием, если учесть, что R и ρ будут постоянными, а

$$\frac{1}{r_1}=\frac{1}{\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}},$$

где x_1, y_1, z_1 — координаты фиксированной точки M_1, x, y, z — координаты переменной точки M;

2) функция (2.114) как функция точки M удовлетворяет условию

$$\Gamma(M,\ M_0)\,|_S=0,$$

что следует из (2.113);

3) функция (2.114) имеет нужное представление.

Чтобы найти решение заданной задачи, остается подставить (2.114) в формулу (2.107)

$$u(M_0)=-\frac{1}{4\pi}\iint\limits_S f(M)\frac{\partial\left(\frac{1}{r}-\frac{R}{\rho}\frac{1}{r_1}\right)}{\partial n}dS.$$

Полученную формулу можно преобразовать к более удобному виду

$$u(M_0)=\frac{1}{4\pi R}\iint\limits_S f(M)\frac{R^2-\rho^2}{r^3}\,dS. \qquad (2.115)$$

Она носит название *формулы Пуассона* и дает решение внутренней задаче Дирихле для шара.

§ 2.4. МЕТОД ИНТЕГРАЛЬНЫХ ПРЕОБРАЗОВАНИЙ

Рассмотренные методы интегрирования уравнений в частных производных обычно называются классическими. Как мы видели, эти методы, приспособленные для решения различных краевых задач математической физики, обладают рядом недостатков, которые затрудняют их использование в практике. Решения, получаемые классическими методами, очень часто нуждаются в дальнейшей доработке с целью получения упрощенных приближенных соотношений, в которых интересующие исследователя параметры изучаемого процесса должны быть отделены от других физических характеристик задачи. Иногда эти соотношения получаются с таким трудом, что инженеру приходится искать другие пути решения поставленной задачи.

За последнее время широкое признание среди физиков и инженеров получил метод интегральных преобразований, который нам знаком по ч. 1 пособия, где при помощи этого метода решались линейные обыкновенные дифференциальные уравнения с постоянными коэффициентами. Суть метода состояла в том, что применяя преобразование Фурье или преобразование Лапласа мы сводили решение заданного дифференциального уравнения к решению алгебраического уравнения, которое, очевидно, в математическом отношении представляет задачу несравненно более простую, чем первоначальная. Но не только в этом сказывалось преимущество метода интегральных преобразований (операционного метода) перед классическим. Наряду со значительным сокращением вычислительной работы, он позволял получить решение в такой форме, которая была наиболее удобной для исследователя, так как давала значительно большую обозримость полученного решения. Особенно ярко все эти преимущества описываемого метода проявлялись при решении системы дифференциальных уравнений, когда каждая неизвестная функция определялась сама по себе, независимо от вычисления остальных неизвестных функций.

Но еще в большей степени операционный метод импонировал инженерам своей простотой, связанной с унифицированной последовательностью действий, которая быстро и без помех приводила к решению поставленной задачи. Хотя такие рецептурные приемы решения дифференциальных уравнений приводили нередко к грубым ошибкам, но глубокое проникновение в физическую суть задачи сопутствовало успеху, давая возможность обойти возникающие здесь подводные камни.

Как мы уже говорили, в основу описываемого здесь метода были положены два интегральных преобразования: преобразование Фурье и преобразование Лапласа. В связи с тем, что в дальнейшем нам придется неоднократно пользоваться этими преобразованиями, уместно напомнить читателю относящиеся сюда основные факты, отсылая за подробностями к ч. 1 пособия.

Интегральным преобразованием Фурье называется такое преобразование заданной функции $f(x)$ на промежутке $(-\infty, +\infty)$, которое осуществляется интегралом

$$F(p) = \int_{-\infty}^{+\infty} e^{-ipx} f(x)\,dx. \tag{2.116}$$

Формула (2.116) позволяет по функции $f(x)$, абсолютно интегрируемой на всей числовой оси и удовлетворяющей условиям Дирихле на любом конечном промежутке этой оси, найти функцию $F(p)$. При этом функция $f(x)$ называется *оригиналом*, а $F(p)$ — *изображением*.

Формула

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ipx}F(p)\,dp \qquad (2.117)$$

позволяет решить обратную задачу: по изображению $F(p)$ найти оригинал. Поэтому она называется обратным преобразованием Фурье, в то время как формула (2.116) называется прямым преобразованием. Можно сказать, что эти формулы представляют два взаимно связанных интегральных уравнения, каждое из которых является решением другого. Они дают возможность преобразовать формально функцию переменной x ($f(x)$) в функцию переменной p ($F(p)$) и обратно.

Если функция $f(x)$ задана на промежутке $[0, +\infty)$ и удовлетворяет на нем упомянутым выше условиям, то можно говорить о так называемых синус- и косинус-преобразованиях Фурье этой функции на промежутке $[0, +\infty)$.

Прямое *синус-преобразование Фурье* имеет вид

$$F(p)=\int_{0}^{+\infty}\sin px f(x)\,dx. \qquad (2.118)$$

Обратное преобразование выражается интегралом

$$f(x)=\frac{2}{\pi}\int_{0}^{\infty}\sin pxF(p)\,dp. \qquad (2.119)$$

Прямое и обратное *косинус-преобразования Фурье* выражаются формулами (см. ч. 1, стр. 78)

$$F(p)=\int_{0}^{\infty}\cos px f(x)\,dx, \qquad (2.120)$$

$$f(x)=\frac{2}{\pi}\int_{0}^{\infty}\cos px\,F(p)\,dp. \qquad (2.121)$$

Условие абсолютной интегрируемости значительно сужает класс функций, представимых интегралом Фурье. С целью расширения этого класса функций вводят в рассмотрение *преобразование Лапласа*, положенное в основу операционного исчисления:

$$F(p)=\int_{0}^{\infty}e^{-px}f(x)\,dx, \qquad (2.122)$$

где x — действительная переменная; $p = s + j\omega$ — комплексная переменная. Функция $f(x)$ предполагается непрерывной вместе со своими производными, за возможным исключением конечного числа точек разрыва первого рода на интервале конечной длины, а также возрастающей не быстрее показательной функции.

Обратное преобразование Лапласа, позволяющее по изображению $F(p)$ находить оригинал, имеет вид

$$f(x) = \frac{1}{2\pi j} \int\limits_{s-j\infty}^{s+j\infty} e^{px} F(p)\, dp. \qquad (2.123)$$

Интегрирование в формуле (2.123) происходит в области комплексной переменной $p = s + j\omega$ вдоль прямой, параллельной мнимой оси и расположенной справа от всех особых точек подынтегральной функции.

Формулы (2.117) и (2.123) называются *формулами обращения*, соответственно преобразования Фурье и преобразования Лапласа. Аналогичные названия имеют и формулы (2.119) и (2.121).

Преобразования Фурье и Лапласа имеют целый ряд замечательных свойств, благодаря чему они используются в приложениях, в частности, при решении обыкновенных дифференциальных уравнений и их систем.

В этом параграфе мы будем применять метод интегральных преобразований для решения дифференциальных уравнений в частных производных второго порядка при заданных начальных и краевых условиях. Точно так же, как и в случае интегрирования обыкновенных дифференциальных уравнений, рассматриваемый метод будет служить эффективным средством решения краевых задач математической физики.

Общая схема применения метода интегральных преобразований для решения дифференциальных уравнений в частных производных состоит в следующем. Как и в случае обыкновенного дифференциального уравнения, выбранное интегральное преобразование применяют к заданному уравнению, тем самым временно исключая одну из независимых переменных. Таким образом, задача сразу упрощается: теперь уже приходится интегрировать уравнение в частных производных, которое содержит на единицу меньше независимых переменных, чем заданное уравнение. Если, например, искомая функция зависела от двух переменных, то после преобразования исходное уравнение станет обыкновенным дифференциальным уравнением, с методами интегрирования которого мы хорошо знакомы.

Предположим, что нам удалось найти решение преобразованного или, как говорят, изображающего, уравнения. Тогда это решение будет функцией p и остальных переменных. Теперь для получения решения заданного уравнения (в пространстве оригиналов) нужно по найденному изображению найти оригинал по одной из формул обращения. Если искомая функция зависит от трех (или четырех) переменных, то иногда интегральное преобразование применяют повторно, последовательно исключая дифференциальные операции сначала по одной переменной, а затем по другой. Решив полученное уравнение в пространстве изображений, последовательно осуществляют обратные преобразования, тем самым находят искомое решение задачи. Выбор интегрального преобразования обусловлен целым рядом причин. Здесь приходится учитывать не только тип самого дифференциального уравнения, но и характер начальных и краевых условий: ведь кроме заданного уравнения приходится преобразовывать и условия задачи. В одном случае эффективным является применение преобразования Лапласа, в другом случае — преобразования Фурье. Иногда условия задачи таковы, что приходится использовать и ряд других преобразований, из которых наибольшее распространение получили преобразование Ханкеля и преобразование Меллина [36, стр. 67]. Если преобразование Лапласа, как правило, применяется по отношению к временной координате, то преобразование Фурье и ряд других преобразований применяются по отношению к пространственной координате. В настоящее время продолжают разрабатываться методы интегрального преобразования по пространственным координатам в соответствии с характером постановки краевых задач математической физики.

В дальнейшем под *интегральным преобразованием с ядром* $K(x, p)$ мы будем понимать преобразование вида

$$F(p) = \int_a^b K(x, p) f(x)\, dx, \tag{2.124}$$

которым функции $f(x)$ переменной x сопоставляется функция $F(p)$ переменной p.

Очевидно, всякое интегральное преобразование определяется ядром $K(x, p)$, промежутком интегрирования (a, b) и множеством функций $f(x)$, к которым оно применяется. Задать интегральное преобразование (2.124) — это значит задать все эти данные. Так, например, интегральное преобразование Фурье (2.116) имеет ядро $K(x, p) = e^{-ipx}$, промежуток интегрирования — всю числовую ось и множество функций $f(x)$, обладающих уже известными свойствами. Косинус-преобразование Фурье имеет ядро $K(x, p) = \cos px$, промежуток интегрирования $(0, +\infty)$ и известный уже нам класс функций $f(x)$.

Все приведенные выше преобразования были такими, когда a равно нулю, а b — бесконечности, за исключением преобразования (2.116), где оба предела интегрирования равны бесконечности. Если иметь в виду преобразование по пространственной координате, то применение рассмотренных интегральных преобразований успешно только к задачам для полуограниченных тел. Это значительно суживает диапазон применения описываемого метода. Большое число задач для тел, имеющих конечные размеры, не может быть решено при помощи интегральных преобразований с бесконечными пределами, что приводит к необходимости рассмотрения таких интегральных преобразований (2.124), когда оба предела a и b — конечны. Такие интегральные преобразования получили название *конечных интегральных преобразований.* Конечные интегральные преобразования представляют наибольший интерес, поскольку они совместно с преобразованием Лапласа по временной координате дают возможность решать разнообразные задачи математической физики для тел конечных размеров. Хотя применение преобразования указанного типа не дает возможности решать новые задачи, которые не могут быть решены классическими методами с помощью рядов Фурье, тем не менее решения с помощью конечных интегральных преобразований обладают всеми преимуществами рассматриваемого метода, которые указаны выше для случая интегральных преобразовиний с бесконечными пределами интегрирования.

Оставляя пока в стороне элементы общей теории интегральных преобразований, рассмотрим ряд примеров, иллюстрирующих метод интегральных преобразований при решении краевых задач математической физики.

2. 4. 1. Распространение тепла в неограниченном стержне

Пусть задан тонкий теплопроводящий стержень, боковая поверхность которого теплоизолирована. Будем считать его настолько длинным, что влиянием температурных условий на концах стержня можно пренебречь. Другими словами, основным фактором, влияющим на распределение температуры вдоль стержня, будет начальная температура, которая предполагается заданной.

Задача о распространении тепла в таком стержне может быть сформулирована следующим образом.

Задача Коши. Найти решение $u(x, t)$ однородного уравнения теплопроводности

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad \infty < x < +\infty, \quad t > 0, \tag{2.125}$$

удовлетворяющее начальному условию

$$u(x, t)\,|_{t=0} = f(x), \tag{2.126}$$

где функция $f(x)$ задана на всей числовой оси.

Для решения этой задачи воспользуемся методом интегральных преобразований.

Характер задачи показывает (стержень бесконечен в обе стороны), что удобнее всего воспользоваться интегральным преобразованием Фурье (2.116).

Введем в рассмотрение преобразованную температуру при помощи интеграла Фурье

$$U(p, t) = \int_{-\infty}^{+\infty} e^{-jpx} u(x, t)\, dx.$$

Теперь уравнение (2.125) подвергнем преобразованию Фурье. Для этого умножим обе части уравнения на e^{-jpx} и проинтегрируем по x в пределах от $-\infty$ до $+\infty$:

$$\int_{-\infty}^{+\infty} e^{-jpx} \frac{\partial u}{\partial t}\, dx = a^2 \int_{-\infty}^{+\infty} e^{-jpx} \frac{\partial^2 u}{\partial x^2}\, dx.$$

Мы будем считать, что операция дифференцирования по t и операция интегрирования по x могут меняться местами. Другими словами

$$\int_{-\infty}^{+\infty} e^{-jpx} \frac{\partial u}{\partial t}\, dx = \frac{\partial}{\partial t} \int_{-\infty}^{+\infty} e^{-jpx} u\, dx = \frac{\partial U(p, t)}{\partial t}.$$

Второй интеграл преобразуем при помощи интегрирования по частям

$$\int_{-\infty}^{+\infty} e^{-jpx} \frac{\partial^2 u}{\partial x^2}\, dx = e^{-jpx} \frac{\partial u}{\partial x}\Big|_{-\infty}^{+\infty} + jp \int_{-\infty}^{+\infty} e^{-jpx} \frac{\partial u}{\partial x}\, dx.$$

В связи с тем, что из физических соображений $\lim\limits_{x\to\infty} \frac{\partial u}{\partial x} = 0$, внеинтегральный член обратится в нуль в правой части последнего равенства. Выполняя операцию интегрирования по частям еще раз, окончательно получим

$$\int_{-\infty}^{+\infty} e^{-jpx} \frac{\partial^2 u}{\partial x^2} dx = jpe^{-jpx} u(x, t)\Big|_{-\infty}^{+\infty} - p^2 \int_{-\infty}^{+\infty} e^{-jpx} u\, dx =$$
$$= -p^2 U(p, t),$$

ибо $\lim\limits_{x\to\infty} u(x, t) = 0$.

Таким образом, в пространстве изображений уравнение (2.125) будет иметь вид

$$\frac{\partial U(p,\ t)}{\partial t}=-a^2p^2U(p,\ t)\ \text{или}\ \frac{\partial U}{\partial t}=-a^2p^2U. \qquad (2.127)$$

Подвергнем теперь интегральному преобразованию Фурье начальную функцию (2.126):

$$\left(\int_{-\infty}^{+\infty} e^{-ipx}u(x,\ t)\,dx\right)\Bigg|_{t=0}=\int_{-\infty}^{+\infty} e^{-ipx}f(x)\,dx=F(p)$$

или

$$U(p,\ t)\,|_{t=0}=F(p), \qquad (2.128)$$

где $F(p)$ — изображение функции $f(x)$.

Итак, мы свели краевую задачу (2.125), (2.126) к решению обыкновенного дифференциального уравнения первого порядка (2.127) при начальном условии (2.128).

Разделяя переменные в уравнении (2.127) и интегрируя, получим

$$U(p,\ t)=F(p)\,e^{-a^2p^2t}. \qquad (2.129)$$

Теперь остается выполнить последний шаг метода интегральных преобразований: перейти от найденного изображения (2.129) обратно к оригиналу — решению заданного уравнения. Обычно этот этап метода наиболее труден и часто требует от исследователя большой настойчивости и изобретательности.

В рассматриваемом случае прямое применение формулы обращения (2.117), очевидно, ничего не даст. Но правую часть равенства (2.129) можно рассматривать как произведение двух изображений $F(p)$ и $e^{-a^2p^2t}$. Известно (см. ч. 1, стр. 90), что произведению изображений соответствует свертка оригиналов этих функций. Другими словами, если оригиналы функций $F(p)$ и $e^{-a^2p^2t}$ обозначить соответственно через $f(x)$ и $\varphi(x)$, то произведению $F(p)\cdot e^{-a^2p^2t}$ будет соответствовать функция $\int_{-\infty}^{+\infty} f(x)\,\varphi(x-\tau)\,d\tau$. Этим фактом мы здесь и воспользуемся.

Для реализации намеченного плана найдем сначала оригинал функции $e^{-a^2p^2t}$, воспользовавшись формулой обращения (2.117)

$$\varphi(x,\ t)=\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{ipx}e^{-a^2p^2t}\,dp=\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-a^2p^2t+ipx}\,dp.$$

Последний интеграл легко сводится к интегралу Пуассона $\int\limits_{-\infty}^{+\infty} e^{-\alpha^2}\,d\alpha = \sqrt{\pi}$. Действительно,

$$\varphi(x, t) = \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} e^{-a^2p^2t + jpx}\,dx =$$

$$= \frac{1}{2\pi}\int\limits_{-\infty}^{+\infty} e^{-\left(a\sqrt{t}\,p - \frac{jx}{2a\sqrt{t}}\right)^2} e^{-\frac{x^2}{4a^2t}}\,dp = \frac{1}{2a\sqrt{\pi t}}\,e^{-\frac{x^2}{2a^2t}}.$$

Итак, оригиналом изображения $U(p, t)$ будет служить функция

$$u(x, t) = \frac{1}{2a\sqrt{\pi t}}\int\limits_{-\infty}^{+\infty} f(x)\,e^{-\frac{(x-\tau)^2}{2a^2t}}\,d\tau.$$

На этом решение задачи заканчивается.

Интересно отметить, что функция

$$\Gamma(x, t; \tau) = \frac{1}{2a\sqrt{\pi t}}\,e^{-\frac{(x-\tau)^2}{2a^2t}}, \tag{2.130}$$

рассматриваемая как функция от переменных x и t, является решением уравнения (2.125). Функцию (2.130) называют *фундаментальным решением уравнения теплопроводности* (2.125). Фундаментальное решение дает распределение температуры, которое вызывается мгновенным точечным источником тепла, помещенным в начальный момент времени $t = 0$ в точке $x = \tau$ стержня.

2. 4. 2. Распространение тепла в полуограниченном стержне

В качестве второго примера применения метода интегральных преобразований к решению дифференциальных уравнений в частных производных рассмотрим задачу о распространении тепла в стержне, который ограничен с одной стороны, а в другую сторону простирается в бесконечность. Как и раньше будем считать, что боковая поверхность стержня теплоизолирована. Распространение тепла в таком стержне изучается при условии, что конец стержня $x = 0$ поддерживается при постоянной температуре u_0, а начальная температура равна нулю. Таким образом, задача о распространении тепла в полуограниченном стержне сводится к следующей.

Найти решение $u(x, t)$ однородного уравнения теплопроводности

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad x > 0, \ t > 0, \tag{2.131}$$

удовлетворяющее начальному условию

$$u(x, t)\big|_{t=0} = 0 \tag{2.132}$$

и краевому условию

$$u(x, t)\big|_{x=0} = u_0. \tag{2.133}$$

Характер условий задачи показывает, что для ее решения можно пользоваться и преобразованием Лапласа (2.122) и синус-преобразованием Фурье (2.118). Рассмотрим сначала способ решения задачи, связанный с использованием преобразования Лапласа.

Искомая функция $u(x, t)$ преобразуется при помощи интеграла (2.122) в функцию $U(x, p)$ по формуле

$$U(x, p) = \int_0^\infty e^{-pt} u(x, t)\, dt.$$

Здесь преобразование уже применяется не к пространственной координате x, а к временной t. Получим преобразующее уравнение. Для этого обе части уравнения (2.131) умножим на ядро преобразования e^{-pt} и проинтегрируем по t в пределах от 0 до ∞

$$\int_0^\infty e^{-pt} \frac{\partial u}{\partial t}\, dt = a^2 \int_0^\infty e^{-pt} \frac{\partial^2 u}{\partial x^2}\, dt.$$

Первый интеграл вычисляем методом интегрирования по частям

$$\int_0^\infty e^{-pt} \frac{\partial u}{\partial t}\, dt = e^{-pt} u(x, t)\Big|_0^\infty + p \int_0^\infty e^{-pt} u\, dt = pU(x, p).$$

Внеинтегральный член обращается в нуль в силу того, что $\lim\limits_{t\to+\infty} e^{-pt} u(x, t) = 0$ и благодаря начальному условию (2.132).

Как и ранее будем считать, что операция дифференцированная по x и операция интегрирования по t могут меняться местами. Поэтому

$$a^2 \int_0^\infty e^{-pt} \frac{\partial^2 u}{\partial x^2}\, dt = a^2 \frac{\partial^2}{\partial x^2} \int_0^\infty e^{-pt} u\, dt = a^2 \frac{\partial^2 U(x, p)}{\partial x^2}.$$

Таким образом изображающее уравнение будет иметь вид

$$pU = a^2 \frac{\partial^2 U}{\partial x^2}. \tag{2.134}$$

Применим теперь преобразование Лапласа к краевому условию (2.133)

$$\left(\int_0^\infty e^{-pt} u(x, t)\, dt \right)\Bigg|_{x=0} = \int_0^\infty e^{-pt} u_0\, dt$$

или

$$U(x, t)\Big|_{x=0} = \frac{u_0}{p}, \tag{2.135}$$

ибо

$$\int_0^\infty e^{-pt} u_0 dt = u_0 \int_0^\infty e^{-pt}\, dt = \frac{u_0}{p}.$$

Итак, рассматриваемая краевая задача свелась к решению обыкновенного дифференциального уравнения (2.134) при условии (2.135). Интегрирование уравнения (2.134) не представляет труда. С учетом (2.135) ограниченное решение будет иметь вид

$$U(x, p) = \frac{u_0}{p} e^{-\frac{x}{a}\sqrt{p}}. \tag{2.136}$$

Перейдем от найденного изображения искомого решения к оригиналу. По таблице оригиналов и соответствующих им изображений для преобразования Лапласа находим

$$u(x, t) = u_0 erfc\left(\frac{x}{2a\sqrt{t}}\right) = \frac{2u_0}{\sqrt{\pi}} \int_{\frac{x}{2a\sqrt{t}}}^\infty e^{-\alpha^2}\, d\alpha.$$

Если ввести в рассмотрение интеграл вероятностей

$$\Phi(\tau) = \frac{2}{\sqrt{\pi}} \int_0^\tau e^{-\alpha^2}\, d\alpha,$$

значения которого табулированы (имеются подробные таблицы значений этой функции), то решение $u(x, t)$ задачи можно представить в виде

$$u(x, t) = u_0 \left[1 - \Phi\left(\frac{x}{2a\sqrt{t}}\right) \right]. \tag{2.137}$$

Покажем, как применяется синус-преобразование Фурье к решению рассматриваемой задачи теплопроводности. Сначала сделаем несколько замечаний общего характера.

Выбор синус- или косинус-преобразования Фурье для решения той или иной краевой задачи определяется видом краевых условий на нижнем пределе переменной, подлежащей исключению. Так, применение синус-преобразования к производной $\frac{\partial^2 u}{\partial x^2}$, входящей в дифференциальное уравнение, целесообразно в том случае, когда задана величина $u(x, t)$ при $x=0$, а применение косинус-преобразования целесообразно в том случае, когда задано значение $\frac{\partial u(x, t)}{\partial x}$ при $x=0$. Действительно, умножая $\frac{\partial^2 u}{\partial x^2}$ сначала на $\sin px$ и интегрируя по частям полученное выражение по x в пределах от 0 до ∞, найдем

$$\int_0^\infty \frac{\partial^2 u}{\partial x^2} \sin px\,dx = \left(\frac{\partial u}{\partial x} \sin px\right)\Big|_0^\infty - p\int_0^\infty \cos px \frac{\partial u}{\partial x}\,dx.$$

В силу физических условий задачи $\lim\limits_{x\to\infty} \frac{\partial u}{\partial x} = 0$. Значит внеинтегральный член обращается в нуль. После повторного интегрирования по частям получим

$$\int_0^\infty \frac{\partial^2 u}{\partial x^2} \sin px\,dx = -p\,(u\cos px)\Big|_0^\infty - p^2\int_0^\infty u \sin px\,dx.$$

Предполагаем, что $\lim\limits_{x\to\infty} u = 0$, тогда

$$\int_0^\infty \frac{\partial^2 u}{\partial x^2} \sin px\,dx = p\left(u(x, t)\big|_{x=0}\right) - p^2 U(p, t).$$

Если теперь воспользоваться косинус-преобразованием Фурье, то рассуждая аналогично, получим

$$\int_0^\infty \frac{\partial^2 u}{\partial x^2} \cos px\,dx = -\left(\frac{\partial u}{\partial x}\right)\Big|_{x=0} - p^2 U(p, t).$$

Еще одно важное замечание. Если в дифференциальное уравнение входит производная $\frac{\partial u}{\partial x}$, то исключить ее при помощи синус- или косинус-преобразования Фурье по переменной x не удастся. Это следует из того, что интегрирование по частям интеграла $\int\limits_0^\infty \frac{\partial u}{\partial x}\sin pxdx\left(\int\limits_0^\infty \frac{\partial u}{\partial x}\cos pxdx\right)$ приводит к интегралу $\int\limits_0^\infty u\cos pxdx\ \left(\int\limits_0^\infty u\sin pxdx\right)$, который не выражается через $U(p,\ x)$. Это большой недостаток рассматриваемых преобразований по сравнению, например, с преобразованием Лапласа, которое в этом случае быстро приводит к цели. Но у синус- и косинус-преобразований есть и серьезное преимущество по сравнению с преобразованием Лапласа. Нахождение оригинала по известному изображению в первом случае значительно проще, чем во втором. Ведь в первом случае интегралы обращения (2.119) и (2.121) вещественны, а во втором приходится иметь дело с контурным интегралом (2.123) от функции комплексной переменной.

После этих замечаний приступим к решению рассматриваемой задачи при помощи синус-преобразования Фурье. Читателю ясно, что выбор этого преобразования обусловлен характером поставленной задачи.

$$U(p,\ t)=\int\limits_0^\infty \sin pxu(x,\ t)\,dx.$$

Умножим обе части дифференциального уравнения (2.131) на ядро преобразования $\sin px$ и проинтегрируем по x в пределах от 0 до ∞. Тогда с учетом предыдущего получим изображающее уравнение в виде

$$\frac{dU}{dt}=a^2(pu_0-p^2U). \tag{2.138}$$

Применим теперь преобразование к начальному условию (2.132). Очевидно,

$$U(p,\ t)\big|_{t=0}=0. \tag{2.139}$$

Итак, задача свелась к интегрированию обыкновенного дифференциального уравнения (2.138) с начальным условием (2.139).

После решения уравнения (2.138), которое может быть переписано в виде

$$\frac{dU}{dt}+a^2p^2U=a^2pu_0,$$

и перехода в пространство оригиналов мы получим ответ задачи в виде (2.137).

2. 4. 3. Конечные интегральные преобразования и их применение

При решении краевых задач для тел конечных размеров широко применяются интегральные преобразования (2.124), когда пределы интегрирования a и b — конечные числа. Но для практического использования конечных преобразований необходимо знать формулы обращения, аналогичные формулам (2.117) и (2.123) преобразований Фурье и Лапласа, которые позволяли бы по изображению находить оригинал, т. е. выполнять последний этап метода интегральных преобразований. Формулы обращения для конечных интегральных преобразований, как правило, находятся при помощи разложения искомой функции в ряд по ортогональным функциям соответствующей краевой задачи. Таким образом, если для преобразований с бесконечными пределами обратное преобразование является интегральным, то для конечных интегральных преобразований обратное преобразование интегральным уже не будет.

Проиллюстрируем сказанное примерами.

Рассмотрим два конечных интегральных преобразования, которые будем называть конечными синус- и косинус-преобразованиями Фурье

$$F(p) = \int_0^{\pi} \sin pxf(x)\,dx \tag{2.140}$$

и

$$F(p) = \int_0^{\pi} \cos pxf(x)\,dx, \tag{2.141}$$

где p — натуральное число. Выбор пределов интегрирования объясняется только соображениями удобства, так как путем соответствующей замены переменной промежуток интегрирования может быть преобразован в другой.

Получим сначала формулы обращения для преобразований (2.140) и (2.141). Это нам поможет сделать теория рядов Фурье. Действительно, если функция $f(x)$ обладает известными условиями на промежутке $(0, \pi)$, то она может быть разложена на этом промежутке в ряд Фурье или по синусам, или по косинусам кратных дуг. Причем коэффициенты этих разложений b_p и a_p, как известно, определяются соответственно по формулам

$$b_p = \frac{2}{\pi}\int_0^{\pi} \sin pxf(x)\,dx,$$

$$a_p = \frac{2}{\pi} \int_0^{\pi} \cos pxf(x)\,dx.$$

Сравнивая выписанные формулы с (2.140) и (2.141), мы видим, что эти коэффициенты с точностью до множителя $\frac{2}{\pi}$ совпадают с изображениями изучаемых интегральных преобразований. Отсюда следует, что формулами обращения для преобразований (2.140) и (2.141) будут служить следующие

$$f(x) = \frac{2}{\pi} \sum_{p=1}^{\infty} \sin pxF(p) \tag{2.142}$$

и

$$f(x) = \frac{1}{\pi} F(0) + \frac{2}{\pi} \sum_{p=1}^{\infty} \cos pxF(p), \tag{2.143}$$

т. е. разложения оригиналов в ряды Фурье по соответствующей ортогональной системе тригонометрических функций.

Чем же определяется выбор ядра конечного интегрального преобразования при решении краевой задачи? На этот вопрос мы уже, по существу, ответили в предыдущем пункте параграфа. Выбор ядра преобразования определяется в первую очередь характером краевых условий по той переменной, которая исключается из уравнения. Например, при применении конечного синус-преобразования Фурье к дифференциальному уравнению, содержащему производную $\frac{\partial^2 u}{\partial x^2}$, мы приходим к равенству

$$\int_0^{\pi} \sin px \frac{\partial^2 u}{\partial x^2}\,dx = \left(\frac{\partial u}{\partial x} \sin px\right)\Big|_0^{\pi} - p\int_0^{\pi} \cos px \frac{\partial u}{\partial x}\,dx.$$

Внеинтегральный член обращается в нуль за счет множителя $\sin px$ (не забывайте, что здесь p — число натуральное). Производная $\frac{\partial u}{\partial x}$ предполагается величиной ограниченной. Интегрируя еще раз по частям, получим

$$\int_0^{\pi} \sin px \frac{\partial^2 u}{\partial x^2}\,dx = -p(u \cos px)\big|_0^{\pi} - p^2 \int_0^{\pi} \sin px\, u\,dx.$$

Если ввести обозначения

$$u(x, t)\big|_{x=0} = u_0 \text{ и } u(x, t)\big|_{x=\pi} = u_\pi,$$

то последнее равенство можно переписать в виде

$$\int_0^\pi \sin px \frac{\partial^2 u}{\partial x^2}\,dx = p\left[u_0 - (-1)^p u_\pi\right] - p^2 U(p, t). \qquad (2.144)$$

При применении конечного косинус-преобразования Фурье в рассматриваемом случае мы придем к равенству

$$\int_0^\pi \cos px \frac{\partial^2 u}{\partial x^2}\,dx = (-1)^p \left(\frac{\partial u}{\partial x}\right)\Big|_{x=\pi} - \left(\frac{\partial u}{\partial x}\right)_{x=0} - p^2 U(p, t).$$

Читатель без труда сделает очевидные здесь выводы: если известны значения функции $u(x, t)$ при $x = 0$ и $x = \pi$, то производная $\frac{\partial^2 u}{\partial x^2}$ может быть исключена из дифференциального уравнения при помощи конечного синус-преобразования Фурье; если же известны значения производной $\frac{\partial u(x, t)}{\partial x}$ при $x = \pi$ и $x = 0$, то надо применять конечное косинус-преобразование Фурье.

Необходимо отметить, что если в дифференциальное уравнение входит производная $\frac{\partial u}{\partial x}$ (или любая производная нечетного порядка), то исключить ее при помощи конечных синус- или косинус-преобразований Фурье по переменной x не удается по той же самой причине, что и в случае соответствующих бесконечных преобразований. В этом сказывается один из недостатков метода интегральных преобразований.

В качестве примера применения метода конечных интегральных преобразований рассмотрим задачу об определении потенциала внутри квадратной, со стороной равной π, пластины однородной электропроводности, три стороны которой поддерживаются на уровне потенциала земли, а четвертая поддерживается при потенциале, равном u_0. Эта задача может быть сформулирована следующим образом.

Найти решение $u(x, t)$ однородного уравнения Лапласа

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0, \quad 0 < x < \pi, \quad 0 < y < \pi, \qquad (2.145)$$

удовлетворяющее краевым условиям

$$u(x, y)\,|_{x=0} = 0; \quad u(x, y)\,|_{x=\pi} = 0;$$
$$u(x, y)\,|_{y=0} = 0; \quad u(x, y)\,|_{y=\pi} = u_0.$$

Заметим, что здесь сторона квадрата выбрана равной π только из соображений удобства, что не может повлиять на общность получаемых здесь результатов: простая замена переменных приводит к решению задачи для квадрата или прямоугольника любых размеров

$$\xi = \frac{ax}{\pi}; \quad \eta = \frac{by}{\pi}.$$

Так как потенциал задан при $x = 0$ и $x = \pi$, то мы можем воспользоваться конечным синус-преобразованием Фурье

$$U(p, y) = \int_0^\pi \sin px\, u(x, y)\, dx,$$

где p — натуральное число.

Умножим обе части уравнения (2.145) на $\sin px$ и проинтегрируем полученные выражения по x в пределах от 0 до π. Тогда получим, воспользовавшись соотношением (2.144) и заданными краевыми условиями:

$$\int_0^\pi \sin px \frac{\partial^2 u}{\partial x^2}\, dx = -p^2 U(p, y).$$

Учитывая, что операция дифференцирования по y перестановочна с операцией интегрирования по x, найдем

$$\int_0^\pi \sin px \frac{\partial^2 u}{\partial y^2}\, dx = \frac{\partial^2}{\partial y^2} \int_0^\pi \sin px\, u\, dx = \frac{\partial^2 U(p, y)}{\partial y^2}.$$

Таким образом, преобразованное уравнение будет представлять собой обыкновенное дифференциальное уравнение второго порядка

$$\frac{d^2 U}{dy^2} - p^2 U = 0 \tag{2.146}$$

с постоянными коэффициентами.

Преобразуем теперь краевые условия

$$0 = \int_0^\pi \sin px\, u(x, y)\,|_{y=0}\, dx = \left(\int_0^\pi \sin px\, u(x, y)\, dx \right) \Bigg|_{y=0} =$$

$$= U(p, y)\,|_{y=0}; \tag{2.147}$$

$$\left(\int_0^\pi \sin px\, u\, dx \right) \Bigg|_{y=\pi} = \int_0^\pi \sin px\, u_0\, dx = \begin{cases} 0, & \text{если } p \text{ — четное;} \\ \dfrac{2u_0}{p}, & \text{если } p \text{ — нечетное.} \end{cases}$$

Итак,

$$U(p, y)|_{y=\pi} = \begin{cases} 0, & \text{если } p \text{ — четное;} \\ \dfrac{2u_0}{p}, & \text{если } p \text{ — нечетное.} \end{cases} \quad (2.148)$$

Задача свелась к краевой задаче для обыкновенного дифференциального уравнения: найти решение уравнения (2.146), удовлетворяющее краевым условиям (2.147), (2.148).

Общее решение уравнения (2.146) имеет вид

$$U(p, y) = C_1e^{py} + C_2e^{-py}.$$

Используя краевые условия (2.147), (2.148), получим

$$U(p, y) = \begin{cases} 0, & \text{если } p \text{ — четное;} \\ \dfrac{2u_0}{p} \operatorname{cth} p\pi \operatorname{sh} py, & \text{если } p \text{ — нечетное.} \end{cases}$$

Воспользовавшись формулой обращения (2.142), окончательно находим

$$u(x, y) = \frac{4u_0}{\pi} \sum_{k=0}^{\infty} \frac{1}{2k+1} \operatorname{cth}(2k+1)\pi \operatorname{sh}(2k+1) y \sin(2k+1) x.$$

2. 4. 4. Об условиях, обеспечивающих возможность интегральных преобразований

Полученные ранее результаты носили в основном частный характер. Не останавливаясь на строгом обосновании выводов ряда положений общей теории интегральных преобразований, попробуем сделать ряд обобщений.

Пусть задано линейное уравнение второго порядка с постоянными коэффициентами

$$a_{11}\frac{\partial^2 u}{\partial x^2} + a_{22}\frac{\partial^2 u}{\partial y^2} + b_1\frac{\partial u}{\partial x} + b_2\frac{\partial u}{\partial y} + cu = f(x, y) \quad (2.149)$$

вместе с некоторыми дополнительными условиями, в которые могут входить начальные и краевые условия.

Получим достаточные условия, обеспечивающие интегральное преобразование краевой задачи к виду, при котором она не содержит дифференциальных операций по переменной преобразования.

Подвергнем заданное уравнение (2.149) интегральному преобразованию

$$U(p, y) = \int_a^b K(p, x)\, u(x, y)\, dx. \quad (2.150)$$

Очевидно, пределы интегрирования a и b следует выбрать так, чтобы они совпали с пределами изменения переменной x, т. е. той переменной, по которой и происходит преобразование, осуществляемое интегралом (2.150).

Заметим также, что кроме уравнения, преобразованию должны быть подвергнуты те дополнительные условия, которые входят в краевую задачу.

Умножим обе части уравнения (2.149) на ядро преобразования $K(p, x)$ и проинтегрируем полученные выражения по x в пределах от a до b. Тогда получим

$$\int_a^b \left(a_{11}\frac{\partial^2 u}{\partial x^2}+b_1\frac{\partial u}{\partial x}+cu\right)K(p, x)\,dx+\int_a^b \left(a_{22}\frac{\partial^2 u}{\partial y^2}+b_2\frac{\partial u}{\partial y}\right)K(p, x)\,dx=$$
$$=\int_a^b f(x, y)\;K(p, x)\,dx$$

или, переходя в пространство изображений,

$$\int_a^b \left(a_{11}\frac{\partial^2 u}{\partial x^2}+b_1\frac{\partial u}{\partial x}+cu\right)K(p, x)\,dx+a_{22}\frac{\partial^2 U}{\partial y^2}+b_2\frac{\partial U}{\partial y}=F(p, y),$$

где $F(p, y)$ — изображение свободного члена $f(x, y)$.

Чтобы добиться намеченной цели, оставшийся интеграл преобразуем методом интегрирования по частям. После элементарных вычислений получим

$$\int_a^b \left(a_{11}\frac{\partial^2 K}{\partial x^2}-b_1\frac{\partial K}{\partial x}+cK\right)u\,dx+\left[\left(a_{11}\frac{\partial u}{\partial x}+b_1 u\right)K-a_{11}u\frac{\partial K}{\partial x}\right]\Bigg|_a^b+$$
$$+a_{22}\frac{\partial^2 U}{\partial y^2}+b_2\frac{\partial U}{\partial y}=F(p, y).$$

Изображающее уравнение не будет содержать интегрального члена, если

$$a_{11}\frac{\partial^2 K}{\partial x^2}-b_1\frac{\partial K}{\partial x}+cK=-\lambda^2 K, \qquad (2.151)$$

где λ — постоянная величина. Действительно,

$$\int_a^b \left(a_{11}\frac{\partial^2 K}{\partial x^2}-b_1\frac{\partial K}{\partial x}+cK\right)u\,dx=-\lambda^2\int_a^b Ku\,dx=-\lambda^2 U(p, y).$$

Соотношение (2.151) можно рассматривать как уравнение для определения ядра $K(p, x)$. Его можно переписать в виде

$$a_{11}\frac{d^2 K}{dx^2}-b_1\frac{dK}{dx}+(c+\lambda^2)K=0. \qquad (2.152)$$

Таким образом, если ядро преобразования $K(p, x)$ является решением уравнения (2.152), то изображающее уравнение примет вид

$$-\lambda^2 U + \left[\left(a_{11}\frac{\partial u}{\partial x} + b_1 u\right)K - a_{11}u\frac{\partial K}{\partial x}\right]\Big|_a^b + a_{22}\frac{\partial^2 U}{\partial y^2} + \\ + b_2\frac{\partial U}{\partial y} = F(p, y).$$

Теперь наши усилия должны быть направлены на то, чтобы вычислить выражение

$$\left[\left(a_{11}\frac{\partial u}{\partial x} + b_1 u\right)K - a_{11}u\frac{\partial K}{\partial x}\right]\Big|_a^b, \tag{2.153}$$

в которое входят значения искомой функции $u(x, y)$ и ее производной $\frac{\partial u(x, y)}{\partial x}$ при $x = a$ и $x = b$. Это сделать не так просто, как кажется на первый взгляд. Во-первых, необходимо, чтобы условия задачи (начальные, краевые и другие) допускали выражение условий по переменной x только через известные функции переменной y, иначе выражение (2.153) внесло бы в изображающее уравнение кроме функции $U(p, y)$ еще и другие неизвестные функции. Во-вторых, так как $u(x, y)$ и $\frac{\partial u(x, y)}{\partial x}$ при двух различных значениях x не могут быть заданы произвольно (благодаря условиям задачи), они не могут одновременно фигурировать в условиях задачи, что является необходимым для вычисления выражения (2.153). Отсюда следует, что за ядро выбранного интегрального преобразования нельзя принять произвольное решение уравнения (2.152), так как в этом случае выражение (2.153) вычислить не удастся.

Разобранные примеры показывают, что ядро $K(p, x)$ преобразования нужно выбрать таким, чтобы можно было обойти возникающую здесь трудность. Обычно это делают следующим образом, предполагая, что преобразование конечное. Учитывая характер краевых условий, составляют краевую задачу для обыкновенного дифференциального уравнения (2.152), которая имеет бесчисленное множество собственных значений

$$\lambda_1, \lambda_2, \lambda_3, \ldots, \lambda_n, \ldots,$$

каждому из которых соответствует собственная функция

$$K_1, K_2, K_3, \ldots, K_n, \ldots$$

Причем, как правило, собственные функции K_i оказываются ортогональными на промежутке (a, b) и допускают представление в виде сходящегося ряда

$$\sum_{i=1}^{\infty} A_i K_i$$

для любой функции $f(x)$, т. е.

$$f(x) = \sum_{i=1}^{\infty} A_i K_i,$$

где A_i — коэффициенты ряда.

Если, например, краевые условия для уравнения (2.149) имеют самый общий вид

$$\left[\alpha_a \frac{\partial u}{\partial x} + \beta_a u\right]_{x=a} = \varphi_a; \quad \left[\alpha_b \frac{\partial u}{\partial x} + \beta_b u\right]_{x=b} = \varphi_b,$$

то для построения ядра конечного интегрального преобразования выбирают такую систему собственных функций (решений уравнения (2.152)), которая удовлетворяет однородным краевым условиям

$$\left[\alpha_a \frac{\partial K}{\partial x} - \beta_a K\right]_{x=a} = 0; \quad \left[\alpha_b \frac{\partial K}{\partial x} + \beta_b K\right]_{x=b} = 0.$$

При этом выражение (2.153) будет вычислено и изображающее уравнение можно считать преобразованным до конца. Остается только преобразовать заданные дополнительные условия, чтобы преобразованная задача вообще не содержала переменной x.

После решения изображающего уравнения необходимо перейти обратно в пространство оригиналов. Как было указано ранее, формула обращения для конечного интегрального преобразования (2.150) будет иметь вид

$$u(x, y) = \sum_{i=1}^{\infty} U_i K_i,$$

где K_i — система собственных функций соответствующей краевой задачи для уравнения (2.152).

В случае бесконечного интегрального преобразования поступают точно так же. Ядро преобразования должно быть решением уравнения (2.152), притом таким, чтобы выражение (2.153) могло быть вычислено. Если это ядро является одним из знакомых ядер, то мы сможем провести прямое и обратное преобразование по известным формулам, приведенным в этом параграфе.

§ 2. 5. ПОНЯТИЕ О ПРИБЛИЖЕННОМ РЕШЕНИИ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ В ЧАСТНЫХ ПРОИЗВОДНЫХ МЕТОДОМ КОНЕЧНЫХ РАЗНОСТЕЙ *

В предыдущих параграфах мы познакомились с аналитическими методами интегрирования уравнений в частных производных. Каждый раз в результате применения того или иного метода решение рассматриваемой краевой задачи выписывалось в виде ряда или интеграла, т. е. в виде конечной формулы. Но, к сожалению, значительно чаще встречаются такие задачи, когда получить решение в явном виде известными методами не удается. Попробуйте применить метод Фурье, например, к уравнению теплопроводности вида

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left[\rho(x, t)\frac{\partial u}{\partial x}\right],$$

когда $\rho(x, t)$ не представима в виде произведения двух функций, одна из которых зависит только от x, а другая только от t. Вы очень скоро убедитесь в несостоятельности своей попытки: выписать решение в явном виде не удастся. А ведь на первый взгляд это уравнение кажется совсем «безобидным»: оно линейное, очень близко напоминает уже изученное ранее уравнение теплопроводности, которое без особого труда было нами проинтегрировано.

Особую группу задач математической физики составляют такие задачи, которые приводят к нелинейным уравнениям в частных производных. Только в исключительных случаях можно представить решение нелинейного уравнения в аналитической форме, т. е. проинтегрировать его в квадратурах.

В связи с этим возникает необходимость в разработке методов *приближенного* решения различных задач для дифференциальных уравнений в частных производных. Значительное развитие приближенные методы получили за последние годы в связи с бурным ростом вычислительной техники.

Приближенные методы решения уравнений в частных производных условно можно разделить на две большие группы. К первой группе следует отнести методы, в которых приближенное решение получается в виде отрезка некоторого ряда, т. е. в аналитической форме. Описанный в предыдущих параграфах метод Фурье может служить примером подобного рода. Как известно, при применении этого метода точное решение

* Перед чтением § 2.5 рекомендуем повторить материал § 3.11 ч. 1 пособия.

дифференциального уравнения получается, вообще говоря, в виде бесконечного сходящегося ряда. Так как суммирование этого ряда оказывается, как правило, невозможным, то на практике в качестве решения рассматриваемой задачи берут сумму некоторого числа первых членов ряда-решения. Так приходят к приближенному решению дифференциального уравнения в частных производных.

Широкую известность за последнее время получил также метод, разработанный академиком Б. Г. Галеркиным и предложенный им для приближенного интегрирования дифференциальных уравнений. Этот метод с успехом может быть применен к дифференциальным уравнениям различных типов и порядков. Суть его состоит в следующем. Пусть отыскивается решение $u(x, y)$ уравнения

$$L[u] = 0,$$

где L — некоторый дифференциальный оператор с двумя переменными, удовлетворяющее однородным краевым условиям. Неизвестную функцию ищут приближенно в виде конечного ряда

$$U(x, y) \approx a_1\varphi_1(x, y) + a_2\varphi_2(x, y) + \ldots + a_n\varphi_n(x, y) = \sum_{i=1}^{n} a_i\varphi_i(x, y), \qquad (*)$$

где $\varphi_i(x, y)$ — произвольные, заранее выбранные функции, удовлетворяющие краевым условиям, а a_i — неопределенные коэффициенты.

Естественно предполагаемое решение $U(x, y)$ подставить в левую часть заданного уравнения. Тогда получим приближенное равенство

$$L[U] = L\left[\sum_{i=1}^{n} a_i\varphi_i(x, y)\right] \approx 0.$$

Умножим обе части полученного равенства на $\varphi_i(x, y)$ и проинтегрируем по области W, в которой и отыскивается решение уравнения $L[u]$:

$$\iint\limits_W L[U(x, y)]\,\varphi_j(x, y)\,dx\,dy = \iint\limits_W L\left[\sum_{i=1}^{n} a_i\varphi_i(x, y)\right]\varphi_j(x, y) \approx 0,$$

$$j = 1, 2, 3, \ldots, n.$$

В связи с тем, что функции $\varphi_i(x, y)$ предполагаются известными, интегралы

$$\iint\limits_W \varphi_i(x, y)\,\varphi_j(x, y)\,dx\,dy$$

можно вычислить, причем каждой функции $\varphi_i(x, y)$ будет соответствовать свое уравнение с неопределенными пока коэффициентами $a_1, a_2, \ldots, a_n$. Таким образом можно получить систему n алгебраических уравнений первой степени с n неизвестными $a_1, a_2, \ldots, a_n$, решив которую и подставив найденные значения неизвестных в (*), мы получим приближенное решение исходного дифференциального уравнения. Очевидно, описываемый здесь метод академика Б. Г. Галеркина следует отнести к первой группе приближенных методов интегрирования дифференциальных уравнений в частных производных. Обычно к этой же группе относят так называемые вариационные методы решения краевых задач, широко использующие основные понятия вариационного исчисления.

Ко второй группе приближенных методов решения уравнений в частных производных следует отнести те методы, которые позволяют получить таблицу приближенных значений искомого решения в точках области, где и отыскивается решение задачи. Такие методы принято называть численными методами интегрирования уравнений математической физики.

По существу все численные методы содержат схемы приближенных вычислений, с помощью которых за конечное число шагов можно получить достаточно хорошее приближение (аппроксимацию) решения заданного уравнения. При этом предусматривается, что каждый из вычислительных шагов может быть выполнен с некоторой точностью человеком или машиной. Особое место среди численных методов интегрирования дифференциальных уравнений занимает метод конечных разностей (метод сеток). Универсальность, возможность применения в линейных и в нелинейных задачах делает метод сеток самым распространенным из применяемых в настоящее время приближенных методов решения уравнений математической физики. Но не только чрезвычайная общность метода сеток привлекает исследователей. Пожалуй, это наиболее удобный и прозрачный численный метод, при помощи которого почти всегда можно получить представление об искомом решении. Хотя достигаемая этим методом точность как правило невелика, но перечисленные преимущества ставят его в ряду численных методов на первое место.

Передать идею метода сеток можно буквально в нескольких словах. Представьте себе, что перед вами стоит задача о нахождении приближенного решения некоторого дифференциального уравнения, связывающего производные различных порядков от искомой функции $u(x, y)$. Встает вопрос: нельзя ли упростить заданное уравнение, «избавившись» каким-нибудь образом от производных искомой функции $u(x, y)$? Самый естественный путь в решении этой задачи содержится в формуле, хорошо всем известной (определение производной!)

$$\frac{\partial u}{\partial x}=\lim_{h\to 0}\frac{u(x+h, y)-u(x, y)}{h} \quad \left(\frac{\partial u}{\partial y}=\lim_{l\to 0}\frac{u(x, y+l)-u(x, y)}{l}\right).$$

Действительно, если отбросить в этом равенстве знак предела, то получим приближенную формулу

$$\frac{\partial u}{\partial x}\approx\frac{u(x+h, y)-u(x, y)}{h} \quad \left(\frac{\partial u}{\partial y}\approx\frac{u(x, y+l)-u(x, y)}{l}\right),$$

позволяющую заменить производную $\frac{\partial u}{\partial x}\left(\frac{\partial u}{\partial y}\right)$ разностным отношением

$$\frac{u(x+h, y)-u(x, y)}{h} \quad \left(\frac{u(x, y+l)-u(x, y)}{l}\right).$$

Оказывается и производные более высокого порядка могут быть аппроксимированы разностными отношениями подобного рода.

Таким образом, если теперь все производные от искомой функции $u(x, y)$, входящие в дифференциальное уравнение, заменить соответствующими разностными соотношениями, аппроксимирующими эти производные, то получим конечно-разностное уравнение, которое, вообще говоря, в математическом отношении представляет задачу более простую, чем первоначальная (см. ч. 1, стр. 194). Решение $U(x, y)$ полученного конечно-разностного уравнения, аппроксимирующего в известном нам смысле дифференциальное уравнение, следует считать приближенным решением исходного дифференциального уравнения. Читателю ясно, что от полученного таким образом приближенного решения $U(x, y)$ необходимо требовать, чтобы при стремлении к нулю приращений неизвестных переменных ($h\to 0$, $l\to 0$) оно приближалось (сходилось) к решению исходного дифференциального уравнения $u(x, y)$.

Легко увидеть, что в конечно-разностном уравнении, аппроксимирующем дифференциальное, x и y являются переменными непрерывными. В связи с этим и само это уравнение называется *уравнением с непрерывными аргументами*. Решить такое уравнение бывает очень трудно, а порой и просто невозможно. В связи с этим вместо конечно-разностного уравнения с непрерывными аргументами чаще всего рассматривают соответствующее ему конечно-разностное уравнение с дискретными аргументами, когда независимые переменные x и y пробегают дискретное множество значений, например

$$x_i = ih,\quad y_k = kl,\quad i, k = 0,\ \pm 1,\ \pm 2,\ \ldots$$

Полученное таким образом конечно-разностное уравнение с дискретными аргументами является подходящей аппроксимацией исходного дифференциального уравнения. Решение этого уравнения U_{ik} позволяет получить таблицу приближенных значений для искомого решения $u(x, y)$.

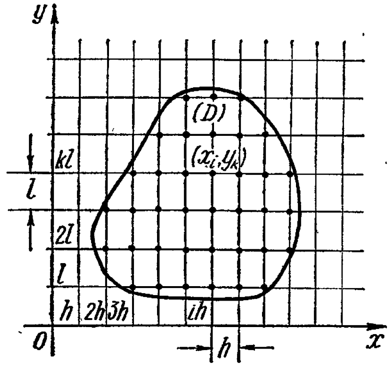

Рис. 2.17

Обычно дискретизация исходного дифференциального уравнения проходит по следующей схеме. Область D непрерывного изменения аргументов x и y заменяется конечным (дискретным) множеством точек (узлов), которые могут быть получены как результат пересечения параллельных прямых:

$$x = ih,\quad i = 0;\ \pm 1,\ \pm 2,\ \ldots$$

и

$$y = kl;\quad k = 0;\ \pm 1;\ \pm 2,\ \ldots$$

(рис. 2.17). Это семейство прямых называют *сеткой*. Все функции непрерывного аргумента, входящие в дифференциальное уравнение, заменяются функциями дискретного аргумента, определенными в узлах сетки. Например, значения решения $u(x, y)$ дифференциального уравнения в области D заменяются значениями этой функции в узлах сетки (x_i, y_k), принадлежащих области D, которые мы обозначим через u_{ik}. Производные, входящие в дифференциальное уравнение, аппроксимируются соответствующими разностными отношениями, но уже для функций дискретного аргумента. Например,

$$\left(\frac{\partial u}{\partial x}\right)_{i,k} \approx \frac{u(x_i + h, y_k) - u(x_i, y_k)}{h} = \frac{u_{i+1,k} - u_{i,k}}{h}$$

$$\left(\left(\frac{\partial u}{\partial y}\right)_{i,k} \approx \frac{u(x_i, y_k + l) - u(x_i, y_k)}{l} = \frac{u_{i,k+1} - u_{i,k}}{l}\right).$$

Производные более высоких порядков от функции $u(x, y)$ также удается выразить в виде линейных комбинаций значений дискретной функции $u_{i,k}$ в нескольких узлах сетки.

В результате дифференциальное уравнение заменяется конечно-разностным уравнением с дискретными аргументами. Легко увидеть, что такое конечно-разностное уравнение можно истолковать как систему линейных алгебраических уравнений относительно неизвестных $u_{i,k}$, в которой число уравнений равно числу неизвестных. Разрешая полученную систему алгебраических уравнений, получают приближенные значения искомого решения $u(x, y)$. Если кроме дифференциального уравнения были заданы начальные и краевые условия, то они тоже должны быть заменены разностными начальными и краевыми условиями для функции $u_{i,k}$, которые должны быть рассмотрены совместно с полученной ранее системой алгебраических уравнений.

Теперь мы хотим обратить внимание читателя на одну любопытную деталь, приглашая его сделать небольшой экскурс в прошлое. В § 1.1 пособия отмечалось, что в задачах теории поля переменные x и y являются непрерывно меняющимися независимыми переменными, а уравнения, описывающие систему с распределенными параметрами — дифференциальными уравнениями в частных производных. Указывалось на принципиальное отличие задач с сосредоточенными параметрами от задач с распределенными параметрами, т. е. задач теории поля. Что же означает с физической точки зрения замена дифференциального уравнения в частных производных, описывающего систему с распределенными параметрами, аппроксимирующим его конечно-разностным уравнением? Ведь такая замена предусматривает переход от непрерывного поля к дискретному полю, напоминающему систему с сосредоточенными параметрами! Можно ли считать такую замену оправданной с физической точки зрения? При построении конечно-разностного приближения мы заменяем систему с распределенными параметрами набором дискретных элементов, причем так, что характеристики первоначально заданного поля приближенно остаются неизменными. Короче говоря, процесс дискретизации можно считать вполне оправданным, если только расстояние между соседними узлами сетки достаточно мало. Посмотрите на кусок ткани с большого расстояния! Он вам покажется непрерывным. Но если на него посмотреть через увеличительное стекло, то окажется, что ткань состоит из множества отдельных точек, которые в свою очередь составлены из большого числа волокон. Если для приготовления ткани использовано достаточное количество ниток, она будет иметь все характерные черты системы с непрерывно распределенными параметрами. Этот пример может служить иллюстрацией процесса дискретизации непрерывных систем, где аппроксимации вносят очевидные ошибки в решение. Следует ожидать, что при достаточном уменьшении промежутков между узлами ошибки, порожденные аппроксимацией, станут настолько малыми, что не окажут существенного влияния на общую точность решения.

Неискушенному читателю может показаться, что описываемый здесь в общих чертах численный метод интегрирования дифференциального уравнения в частных производных, состоящий в построении конечно-разностного уравнения и решении его, обладает такой силой, что каждый раз применяя этот метод для нахождения приближенного решения заданного дифференциального уравнения можно получить желаемый результат. К сожалению, это не совсем так. Даже в самых простых случаях приходится сталкиваться с нежелательными явлениями.

Представьте себе, что конечно-разностное уравнение, аппроксимирующее заданное дифференциальное, построено. (Это построение может быть выполнено, очевидно, неоднозначно, в чем вы убедитесь в дальнейшем.) Удалось разрешить полученную систему линейных алгебраических уравнений. И вдруг оказывается, что полученное приближенное решение не сходится при увеличении числа шагов выбранной сетки к решению исследуемой краевой задачи. Такое приближенное решение никакой практической ценности, конечно, не имеет. В чем дело? От чего зависит сходимость решения, полученного конечно-разностным методом: от характера сетки или от вида конечно-разностного уравнения? Все эти вопросы требуют внимательного рассмотрения и решения.

При проведении вычислительной работы нам приходится учитывать ошибки, встречающиеся, например, при округлении промежуточных результатов. Может оказаться, что ошибка, допущенная на определенном этапе работы, распространяется и растет при каждом последовательном расчете, иногда превышая само искомое решение. Такой вычислительный процесс нельзя считать, очевидно, пригодным для практических целей. С другой стороны, может быть и такой случай, когда описываемая выше ошибка в процессе дальнейшего вычисления будет становиться все меньше и меньше и ее влияние или совсем не скажется на окончательном результате, или же скажется совершенно незначительно.

Было замечено, что вычислительные схемы, полученные по методу сеток очень часто подвержены этим явлениям. В одном случае (для одной разностной схемы) вычислительные ошибки имеют способность расти, «раскачиваться», в другом случае (для другой разностной схемы этого же уравнения) такие ошибки уменьшаются, «затухают». В первом случае соответствующее конечно-разностное уравнение называют *неустойчивым*, а во втором случае — *устойчивым*. Читателю ясно, что только устойчивое конечно-разностное уравнение имеет смысл рассматривать с целью получения приближенного решения заданного дифференциального уравнения. Неустойчивые уравнения практической ценности не имеют. В связи с этим среди множества проблем численного интегрирования дифференциальных уравнений выяснение условий, при которых конечно-разностное уравнение будет устойчивым, занимает одно из первых мест.

Интересно отметить, что вопрос сходимости решения конечно-разностного уравнения к решению дифференциального тесно связан с вопросом устойчивости рассматриваемого аппроксимирующего уравнения. Действительно, решение дифференциального уравнения можно представить как сумму решения аппроксимирующего его конечно-разностного уравнения и некоторой ошибки, ошибки аппроксимации. Если конечно-разностное уравнение устойчиво, то следует ожидать, что ошибка аппроксимации будет уменьшаться, затухать. Другими словами, решение конечно-разностного уравнения будет мало отличаться от решения дифференциального и при измельчании сетки будет к нему стремиться. Если же конечно-разностное уравнение неустойчиво, то ошибка аппроксимации будет, вообще говоря, увеличиваться, «раскачиваться», т. е. решение конечно-разностного уравнения уже не будет сходиться к решению дифференциального. Грубо говоря, из устойчивости следует сходимость. Приведенные здесь важные факты могут быть установлены совершенно строго. Мы рекомендуем внимательному читателю обратиться к одному из специальных руководств, где он легко отыщет соответствующие доказательства.

Применение метода сеток мы продемонстрируем на ряде простейших примеров.

В свое время мы указывали на принципиальное отличие эллиптических уравнений, описывающих стационарные процессы, от параболических и гиперболических, которые описывают процессы нестационарные. В связи с этим и применение метода сеток к этим двум категориям уравнений глубоко различно. Чтобы подчеркнуть это обстоятельство, мы разбили последующий материал параграфа на три пункта. В первом пункте рассмотрены некоторые первоначальные понятия метода сеток. Во втором приведены примеры на нахождение приближенного решения эллиптического уравнения (уравнения Лапласа). Третий пункт содержит пример на интегрирование уравнения параболического типа (уравнения теплопроводности).

2. 5. 1. Аппроксимация дифференциального уравнения. Конечно-разностное уравнение

Пусть задано уравнение второго порядка с постоянными коэффициентами

$$L[u] \equiv a_{11} \frac{\partial^2 u}{\partial x^2} + a_{22} \frac{\partial^2 u}{\partial y^2} + b_1 \frac{\partial u}{\partial x} + b_2 \frac{\partial u}{\partial y} + cu = f, \quad (2.154)$$

где f — заданная функция независимых переменных x и y. В этом уравнении отсутствует член, содержащий смешанную производную. Очевидно, выписанное уравнение может принадлежать и гиперболическому, и параболическому, и эллиптическому типу. Безотносительно к какому типу принадлежит это уравнение, применим метод сеток для нахождения приближенного решения краевой задачи, поставленной для этого уравнения.

Требуется найти решение $u(x, y)$ уравнения (2.154), удовлетворяющее краевому условию

$$u(x, y)|_S = \varphi(x, y),$$

где S — граница области W, в которой отыскивается решение, и начальному условию

$$u(x, y)|_{y=0} = \psi(x). \, ^*$$

Желая ярче подчеркнуть идею аппроксимации дифференциального уравнения (2.154) конечно-разностным уравнением, ограничимся пока рассмотрением случая, когда область W представляет собой часть плоскости xOy, ограниченной прямоугольником S: $0 \leqslant x \leqslant a$; $0 \leqslant y \leqslant b$ (рис. 2.18).

Первый этап при нахождении приближенного решения сформулированной краевой задачи методом сеток состоит в выборе особого набора точек, в которых будут определяться значения искомой функции. Отрезок $[0, a]$ изменения переменной x при помощи точек $0 = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = a$ разобъем на n одинаковых частей с шагом $h = \frac{a}{n}$. Тогда $x_i = ih$, $i = 0, 1, 2, \ldots, n$. Совершенно аналогично отрезок $[0, b]$ изменения переменной y при помощи точек $0 = y_0 < y_1 < y_2 < \ldots < y_{r-1} < y_r = b$ разобьем на r одинаковых частей с шагом $l = \frac{b}{r}$. Очевидно, $y_k = kl$, $k = 0, 1, 2, \ldots, r$. Через полученные точки деления проведем прямые, параллельные осям координат. Точки пересечения этих прямых носят названия *узлов* полу-

---

* См. сноску на стр. 104.

ченной прямоугольной сетки и имеют координаты (x_i, y_k) (см. рис. 2.19).

Выбор шагов сетки h и l диктуется обычно следующими соображениями. Во-первых, учитывается характер искомого решения: чем медленнее (более плавно) изменяется функция $u(x, y)$, тем больше (более крупными) могут быть взяты шаги h и l. Во-вторых, учитывается точность, с которой отыскивается

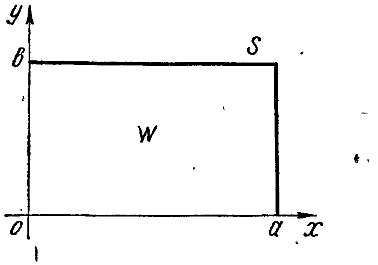

Рис. 2.18

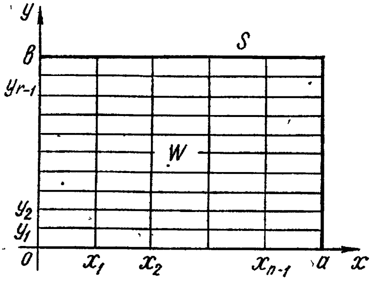

Рис. 2.19

решение задачи. Чем выше требуемая точность, тем мельче нужно брать шаги h и l. Но ведь слишком большие шаги h и l, очевидно, являются и причиной большей погрешности решения. С другой стороны, слишком мелкие требуют большей затраты вычислительного труда. Поэтому должны существовать некоторые «оптимальные» шаги h и l, удовлетворяющие как первому, так и второму требованию. Обычно на практике для выбора более или менее оптимальных шагов применяют принцип просчета с двойными и половинными шагами. Кроме просчета, например, с шагом h, задача еще раз просчитывается с шагом $2h$. Если окажется, что результат, полученный с шагом $2h$, очень мало отличается от результата, полученного с шагом h, то можно считать, что выбранный шаг h не больше оптимального. (В противном случае величину шага h надо уменьшить.) Теперь выполняют просчет с шагом $\frac{h}{2}$. Если результаты, полученные при просчете с шагом h и $\frac{h}{2}$, практически совпадают, то шаг h можно считать более или менее оптимальным.

Уместно отметить, что шаг l, как правило, опеделяется выбором h. Если h выбран, то l должен быть взят таким, чтобы не была нарушена устойчивость полученной конечно-разностной аппроксимации. Оказывается, и вы в этом убедитесь в дальнейшем, устойчивость конечно-разностного уравнения существенно зависит от соотношения между шагами сетки, взятыми вдоль оси Ox и вдоль оси Oy, т. е. между h и l.

Второй этап метода сеток состоит в аппроксимации производных, входящих в дифференциальное уравнение (2.154), краевых условий конечно-разностными отношениями в сеточной области W_h.

Условимся обозначать значения функции $u(x, y)$ в узловых точках (x_i, y_k) выбранной сетки с помощью индексов i, k. Тогда $u(x_i, y_k) = u(ih, kl) = u_{i,k}$. Приближенное значение для $u_{i,k}$ в той же точке обозначим через $U_{i,k}$.

Заменим теперь каждую частную производную $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$, $\frac{\partial^2 u}{\partial x^2}$ и $\frac{\partial^2 u}{\partial y^2}$ разностным отношением. Как уже указывалось, такая замена может быть произведена различными способами. Укажем на один из возможных здесь вариантов

$$\left(\frac{\partial u}{\partial x}\right)_{i,k} \approx \frac{u_{i+1,k} - u_{i,k}}{h}, \tag{2.155}$$
$$\left(\frac{\partial u}{\partial y}\right)_{i,k} \approx \frac{u_{i,k+1} - u_{i,k}}{l}.$$

Частную производную второго порядка $\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,k}$ можно аппроксимировать при помощи разности второго порядка

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,k} \approx \frac{\dfrac{u_{i+1,k} - u_{i,k}}{h} - \dfrac{u_{i,k} - u_{i-1,k}}{h}}{h} =$$
$$= \frac{u_{i+1,k} - 2u_{i,k} + u_{i-1,k}}{h^2}. \tag{2.156}$$

Совершенно аналогично получим

$$\left(\frac{\partial^2 u}{\partial y^2}\right)_{i,k} \approx \frac{u_{i,k+1} - 2u_{i,k} + u_{i,k-1}}{l^2}. \tag{2.156'}$$

Подставляя полученные разностные отношения вместо соответствующих производных в дифференциальное уравнение (2.154) и заменяя значения функции $f(x, y)$ ее значениями в узлах сетки $f(x_i, y_k) = f(ih, kl) = f_{i,k}$, найдем конечно-разностное уравнение, аппроксимирующее заданное дифференциальное

$$(L_h[U])_{i,k} \equiv a_{11}\frac{U_{i+1,k} - 2U_{i,k} + U_{i-1,k}}{h^2} +$$
$$+ a_{22}\frac{U_{i,k+1} - 2U_{i,k} + U_{i,k-1}}{l^2} + b_1\frac{U_{i+1,k} - U_{i,k}}{h} +$$

$$+b_2\frac{U_{i,k+1}-U_{i,k}}{l}+cU_{i,k}=f_{i,k}. \qquad (2.157)$$

Такое конечно-разностное уравнение можно записать для каждого узла (x_i, y_k), где $i=1, 2, \ldots, n-1$; $k=1, 2, \ldots, r-1$.

Присоединяя к полученному разностному уравнению краевые и начальные условия, выписанные в соответствующих узлах сетки, получим разностную краевую задачу.

Конечно-разностное уравнение (2.157) является одним из простейших сеточных уравнений, соответствующих дифференциальному уравнению (2.154). Дело в том, что производная $\left(\frac{\partial u}{\partial x}\right)_{i,k}$ может быть заменена другим разностным отношением, отличным от (2.155). Действительно,

$$\left(\frac{\partial u}{\partial x}\right)_{i,k}\approx\frac{u_{i,k}-u_{i-1,k}}{h} \qquad (2.158)$$

или

$$\left(\frac{\partial u}{\partial x}\right)_{i,k}\approx\frac{u_{i+1,k}-u_{i-1,k}}{2h}. \qquad (2.159)$$

При использовании формулы (2.158) или (2.159) мы получим, очевидно, конечно-разностное уравнение, аппроксимирующее дифференциальное уравнение (2.154), которое будет отличаться от уравнения (2.157). Встает вопрос, какую разностную схему выбрать при аппроксимации дифференциального уравнения конечно-разностным? Читателю ясно, что предпочтение следует отдать тому разностному отношению, которое лучше, более точно аппроксимирует соответствующую производную. Следует ожидать, что такой подход дает возможность получить более точное приближенное решение поставленной выше краевой задачи: чем более точным выражением заменена производная, тем более точным получим решение!

Остановимся пока на решении первой половины вопроса: какое из трех выписанных разностных отношений (2.155), (2.158) или (2.159) более точно аппроксимирует производную $\left(\frac{\partial u}{\partial x}\right)_{i,k}$. Для решения этой проблемы нам придется воспользоваться формулой Тейлора для функции двух переменных. Выпишем ее

$$u(x+h,\ y+l)=u(x,\ y)+du(x,\ y)+\frac{1}{2!}d^2u(x,\ y)+$$
$$+\frac{1}{3!}d^3u(x,\ y)+\frac{1}{4!}d^4u(x+\theta h,\ y+\theta l),\ 0<\theta<1.$$

Применим эту формулу к интересующему нас случаю: заменим $u(x+h, y)$ и $u(x-h, y)$ их тейлоровскими разложениями. Тогда получим

$$u(x+h, y) = u(x, y) + \frac{\partial u(x, y)}{\partial x} h + \frac{1}{2} \frac{\partial^2 u(x, y)}{\partial x^2} h^2 + \frac{1}{6} \frac{\partial^3 u(x, y)}{\partial x^3} h^3 + O(h^4),$$

$$u(x-h, y) = u(x, y) - \frac{\partial u(x, y)}{\partial x} h + \frac{1}{2} \frac{\partial^2 u(x, y)}{\partial x^2} h^2 - \frac{1}{6} \frac{\partial^3 u(x, y)}{\partial x^3} h^3 + O(h^4),$$

где выражение $O(h^4)$ означает величину бесконечно малую при $h \to 0$ четвертого порядка малости относительно h.*

Полученные соотношения могут быть переписаны в следующем виде

$$u_{i+1,k} = u_{i,k} + \left(\frac{\partial u}{\partial x}\right)_{i,k} h + \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,k} h^2 + \frac{1}{6}\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,k} h^3 + O(h^4);$$

$$u_{i-1,k} = u_{i,k} - \left(\frac{\partial u}{\partial x}\right)_{i,k} h + \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,k} h^2 - \frac{1}{6}\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,k} h^3 + O(h^4).$$

(Здесь, как и ранее, $u(x_i+h, y_k) = u_{i+1,k}$; $u(x_i-h, y_k) = u_{i-1,k}$.) Отсюда

$$u_{i+1,k} - u_{i,k} = \left(\frac{\partial u}{\partial x}\right)_{i,k} h + \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,k} h^2 + \frac{1}{6}\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,k} h^3 + O(h^4);$$

$$u_{i-1,k} - u_{i,k} = -\left(\frac{\partial u}{\partial x}\right)_{i,k} h + \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,k} h^2 - \frac{1}{6}\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,k} h^3 + O(h^4);$$

$$u_{i+1,k} - u_{i-1,k} = 2h\left(\frac{\partial u}{\partial x}\right)_{i,k} + \frac{1}{3}\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,k} h^3 + O(h^4).$$

Теперь легко отыскать точные выражения для правых частей формул (2.155), (2.158) и (2.159)

$$\frac{u_{i+1,k} - u_{i,k}}{h} = \left(\frac{\partial u}{\partial x}\right)_{i,k} + \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,k} h + \frac{1}{6}\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,k} h^2 + O(h^3);$$

$$\frac{u_{i-1,k} - u_{i,k}}{h} = -\left(\frac{\partial u}{\partial x}\right)_{i,k} + \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,k} h - \frac{1}{6}\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,k} h^2 + O(h^3);$$

---

* Выражение $O(\alpha^n)$, $n > 0$, означает величину, для которой $\lim\limits_{\alpha \to 0} \frac{O(\alpha^n)}{\alpha^n} = C$, где C — постоянная, отличная от нуля величина.

$$\frac{u_{i+1,k}-u_{i-1,k}}{2h}=\left(\frac{\partial u}{\partial x}\right)_{i,k}+\frac{1}{6}\left(\frac{\partial^3 u}{\partial x^3}\right)_{i,k}h^2+O(h^3).$$

Таким образом, погрешность, полученная за счет аппроксимации производной $\left(\frac{\partial u}{\partial x}\right)_{i,k}$ соответствующим разностным отношением, будет стремиться к нулю при $h\to 0$. Но если в первом и во втором случаях величина погрешности имеет порядок первой степени шага, то в третьем случае — второй степени шага. А это говорит о том, что скорость сходимости разностной схемы, соответствующей (2.159), будет более высокой, чем скорость сходимости схем, соответствующих (2.155) и (2.158). Поэтому, схема (2.159) будет точнее аппроксимировать производную $\left(\frac{\partial u}{\partial x}\right)_{i,k}$, чем схемы (2.155) и (2.158). Говорят, что схема (2.159) имеет второй порядок точности, а схемы (2.155) и (2.158) только первый порядок точности.

Воспользовавшись разностным отношением (2.159), мы придем к конечно-разностному уравнению

$$(L_h[U])_{i,k}\equiv a_{11}\frac{U_{i+1,k}-2U_{i,k}+U_{i-1,k}}{h^2}+$$
$$+a_{22}\frac{U_{i,k+1}-2U_{i,k}+U_{i,k-1}}{l^2}+b_1\frac{U_{i+1,k}-U_{i-1,k}}{h}+$$
$$+b_2\frac{U_{i,k+1}-U_{i,k-1}}{l}+cU_{i,k}=f_{i,k}, \qquad (2.160)$$

которое более точно, чем уравнение (2.157), аппроксимирует рассматриваемое дифференциальное уравнение (2.154).

Можно получить и другие виды конечно-разностных уравнений, аппроксимирующих заданное дифференциальное.

Итак, одно и то же дифференциальное уравнение (2.154) может быть аппроксимировано разными конечно-разностными уравнениями с одинаковым или же нет порядком точности.

Для того, чтобы можно было легко запомнить ту или иную разностную схему, принято сопоставлять каждому конечно-разностному уравнению рисунок, на котором указано в каких точках сетки значения неизвестной функции связаны конечно-разностным уравнением. Так, конечно-разностному уравнению (2.157) можно сопоставить рис. 2.20.

В нашем случае конечно-разностное уравнение связывает значения искомой функции в пяти точках, изображенных на рис. 2.20. Легко увидеть, что уравнение имеет вид

$$(L_h[U])_{i,k}\equiv d_0U_{i,k}+d_1U_{i+1,k}+d_2U_{i-1,k}+d_3U_{i,k+1}+$$
$$+d_4U_{i,k-1}=f_{i,k}, \qquad (2.161)$$

где

$$d_0 = c - \frac{2a_{11}}{h^2} - \frac{2a_{22}}{l^2} - \frac{b_1}{h} - \frac{b_2}{l};$$

$$d_1 = \frac{a_{11}}{h^2} + \frac{b_1}{h}; \quad d_2 = \frac{a_{11}}{h^2};$$

$$d_3 = \frac{a_{22}}{l^2} + \frac{b_2}{l}; \quad d_4 = \frac{a_{22}}{l^2}.$$

Наиболее общий способ построения конечно-разностного уравнения состоит в том, что приближается соответствующим разностным отношением не каждая производная в отдельности, а сразу весь дифференциальный оператор $L[u]$. Схема такого построения состоит в следующем. Сначала выбирают N узлов, расположенных определенным образом около точки (i, k). Например, так, как это изображено на рис. 2.20. Вообще говоря, могут быть и другие варианты такого выбора. При этом количество выбранных узлов не может быть меньше определенного числа. Минимальный набор узлов определяется в первую очередь порядком заданного дифференциального уравнения. Легко увидеть, что минимальное число узлов для первой производной — 2, для второй несмешанной — 3, для второй смешанной — 4 и т. д. Например, уравнение

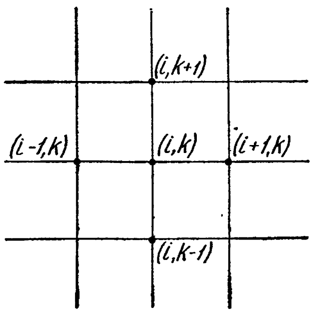

Рис. 2.20

$$\frac{\partial u}{\partial x} - \frac{d^2 u}{\partial y^2} = 0$$

требует такого набора узлов, в котором имеется по крайней мере два узла в направлении оси Ox и три в направлении оси Oy.

После того, как набор узлов уже осуществлен (в дальнейшем для определенности мы будем всюду предполагать, что $N = 4$ и расположение узлов соответствует рис. 2.20), составляют линейную комбинацию

$$(L_h[u])_{i,k} = d_0 u_{i,k} + d_1 u_{i+1,k} + d_2 u_{i-1,k} + d_3 u_{i,k+1} + d_4 u_{i,k-1}$$

с неопределенными коэффициентами d_j. Теперь стараются коэффициенты подобрать так, чтобы оператор L_h аппроксимировал оператор L, т. е.

$$(L_h[u])_{i,k} = (L[u])_{i,k} + \text{остаточный член.} \tag{2.162}$$

Чтобы можно было сравнить операторы L_h и L, разлагают $u_{i+1,k}$, $u_{i-1,k}$, $u_{i,k+1}$ и $u_{i,k-1}$ по формуле Тейлора в окрестности точки (i, k). Подставляя найденные значения в выражение для $(L_h[u])_{i,k}$, подбирают коэффициенты d_j таким образом, чтобы полученное соотношение как можно меньше отличалось от дифференциального выражения $L[u]$ в точке (i, k). Для этого требуют, чтобы коэффициенты при производных в дифференциальном уравнении совпадали с коэффициентами при соответствующих производных в левой части равенства (2.162). В результате получают систему алгебраических уравнений относительно d_j. Если эта система уравнений имеет решение, то задача построения конечно-разностного уравнения, аппроксимирующего заданное дифференциальное, может считаться решенной: получено конечно-разностное уравнение

$$(L[U])_{i,k} \equiv d_0 U_{i,k} + d_1 U_{i+1,k} + d_2 U_{i-1,k} + d_3 U_{i,k+1} + d_4 U_{i,k-1} = f_{i,k},$$

аппроксимирующее дифференциальное (2.154) с наперед заданной точностью.

Описанный способ построения конечно-разностного уравнения имеет ряд преимуществ перед другими способами, о которых речь шла выше. В первую очередь необходимо отметить, что при таком подходе можно рассматривать не только прямоугольную сетку, но и другие сетки (треугольную сетку, сетку параллелограммов и др.).

На этом мы хотим закончить рассмотрение второго этапа метода сеток.

Третий этап метода сеток состоит в решении полученного конечно-разностного уравнения. Оно представляет собой систему линейных алгебраических уравнений относительно значений $U_{i,k}$ во внутренних узлах сетки, где число уравнений равно числу неизвестных. Разрешая эту систему уравнений, получим приближенные значения искомого решения на конечном множестве точек, являющихся внутренними узлами сеточной области W_h. При этом мы считаем, что точное решение разностной задачи $U_{i,k}$ будет, вообще говоря, близким к искомому решению $u_{i,k}$ исходной дифференциальной задачи.

В заключение сделаем несколько замечаний.

Одной из важных проблем метода сеток является проблема сходимости полученного решения разностного уравнения $U_{i,k}$ к точному решению $u_{i,k}$ дифференциального. Эта проблема состоит в том, чтобы найти условия, при которых погрешность в решении

$$\varepsilon_{i,k} = u_{i,k} - U_{i,k}$$

при неограниченном измельчании сетки, т. е. при $h \to 0$ и $l \to 0$ для всех $i = 1, 2, \ldots, n-1$ и $k = 1, 2, \ldots, r-1$ стремилась бы к нулю.

Очевидно (мы уже об этом упоминали), только сходящиеся сеточные методы имеют практическую ценность.

Не менее важной проблемой метода сеток является проблема устойчивости конечно-разностного уравнения. Если через $U_{i,k}$ обозначить точное решение конечно-разностного уравнения, а через $\tilde{U}_{i,k}$ фактическое его решение, то разность $U_{i,k} - \tilde{U}_{i,k}$ можно объяснить ошибками округлений, которые всегда имеют место в процессе вычислений. Тогда проблема устойчивости конечно-разностного уравнения состоит в том, чтобы найти условия, при которых численная погрешность

$$\rho_{i,k} = U_{i,k} - \tilde{U}_{i,k}$$

при возрастании k равномерно по всем i, $0 \leqslant i \leqslant n$ стремится к нулю или остается ограниченной. Характер устойчивости конечно-разностного уравнения зависит от многих факторов. Как мы увидим в дальнейшем, он зависит, во-первых, от соотношения между шагами сетки h и l, во-вторых, от самого вида сеточного уравнения, его структуры. Так, например, сеточное уравнение (2.157), в рассматриваемом здесь смысле, предпочтительнее сеточного уравнения (2.160), хотя последнее, как было выяснено выше, лучше аппроксимирует исходное дифференциальное уравнение. Оказывается, сеточное уравнение (2.160) «предрасположено» к неблагоприятному распределению погрешностей округлений, что практически делает его непригодным для использования. Таким образом, характер устойчивости конечно-разностного уравнения является его внутренним свойством. Известный советский специалист по численным методам В. К. Саульев сравнивал устойчивое сеточное уравнение с хорошим радиоприемником, у которого шумы и разного рода помехи (они практически всегда имеют место по тем или иным причинам) человеческое ухо не улавливает. Хороший радиоприемник можно всегда настроить так, чтобы не был слышен этот неприятный фон. Точно так же некоторые сеточные уравнения можно «настроить» (т. е. выбрать соответствующим образом некоторые параметры, входящие в исследуемое уравнение) так, чтобы нежелательные помехи — погрешности округления не искажали основную «мелодию» — решение $U_{i,k}$ разностной задачи.

2. 5. 2. Решение задачи Дирихле для уравнения Лапласа

В этом пункте будет продемонстрирован метод сеток для решения уравнения эллиптического типа — уравнения Лапласа. На примере решения задачи Дирихле особое внимание будет уделено ряду новых для читателя деталей, с которыми он еще не встречался в тексте параграфа.

Пусть требуется решить задачу:

Найти функцию $u(x, y)$, гармоническую в конечной области D, т. е. удовлетворяющую уравнению Лапласа

$$L[u] \equiv \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial x^2} = 0 \qquad (2.163)$$

и принимающую на границе S этой области заданные значения

$$u(x, y)|_S = \varphi(x, y), \qquad (2.164)$$

где $\varphi(x, y)$ — непрерывная на S функция.

Мы будем предполагать, что граница S области D, вообще говоря, не состоит из прямолинейных отрезков, параллельных осям координат. Другими словами, область D ограничена контуром S общего вида (рис. 2.21).

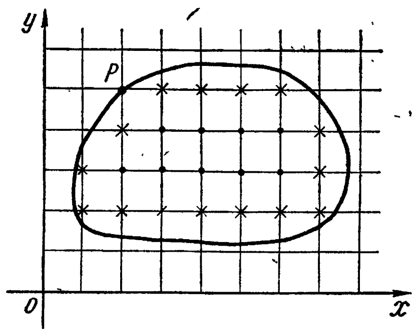

Рис. 2.21

Применение метода сеток для нахождения приближенного решения внутренней задачи Дирихле будет осуществляться в той последовательности как это указано в п. 2.5.1.

Сначала выберем сетку.

Самой простой сеткой на плоскости, очевидно, является квадратная сетка. Квадратной мы назовем сетку, когда точки деления x_i и y_k по осям координат находятся на одном и том же расстоянии h друг от друга. Итак, пусть плоскость, в которой расположена область D с границей S разделена двумя семействами параллельных прямых

$$x = ih; \; y = kh, \; i, \; k = 0, \; \pm 1, \; \pm 2, \; \ldots,$$

образующими квадратную сетку (см. рис. 2.21).

Нас в первую очередь должны интересовать те узлы (x_i, y_k) сетки, которые находятся или внутри области D или на ее границе S: во внутренних узлах сетки нужно определить приближенные значения искомого решения задачи, зная значения решения на границе S. Обратите внимание на то, что в рассматриваемом случае лишь некоторые из узлов попадают на границу S заданной области. На рис. 2.21 таким узлом будет точка P; остальные точки границы не совпадают с узлами выбранной сетки.

Второй этап метода сеток заключается в аппроксимации дифференциального уравнения и краевых условий.

Для аппроксимации частных производных $\frac{\partial^2 u}{\partial x^2}$ и $\frac{\partial^2 u}{\partial y^2}$ воспользуемся соотношениями (2.156) и (2.156′), считая $l = h$. Тогда получим

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \approx \frac{u_{i+1,k} - 2u_{i,k} + u_{i-1,k}}{h^2} + \frac{u_{i,k+1} - 2u_{i,k} + u_{i,k-1}}{h^2} =$$

$$= \frac{1}{h^2}\left[u_{i+1,k} + u_{i,k+1} + u_{i-1,k} + u_{i,k-1} - 4u_{i,k}\right].$$

Таким образом, конечно-разностное уравнение, аппроксимирующее дифференциальное (2.163), будет иметь вид

$$(L_h[U])_{i,k} \equiv \frac{1}{h^2}[U_{i+1,k} + U_{i,k+1} + U_{i-1,k} + U_{i,k-1} - 4U_{i,k}] = 0$$

или

$$U_{i,k} = \frac{1}{4}[U_{i+1,k} + U_{i,k+1} + U_{i,k-1} + U_{i-i,k}], \qquad (2.165)$$

т. е. значение U в какой-либо точке сетки равно среднему арифметическому ее значений в четырех соседних точках. Мы уже указывали, что рассматриваемая разностная схема пятиточечная, что и отражено на рис. 2.20. Каждой внутренней точке (i, k) должно соответствовать одно конечно-разностное уравнение, связывающее значения искомой функции U в четырех соседних точках. К сожалению, для внутренних точек области D, отмеченных на рис. 2.21 крестиками, выписать уравнение типа (2.165) не удастся: не все узлы соседние с фиксированным центральным узлом (i, k) попадают в рассматриваемую область. Итак, конечно-разностное уравнение (2.165) составлено для всех внутренних точек области D, исключая лишь точки, которые на рис. 2.21 обозначены крестиками. Будем эти точки называть граничными точками S_h сеточной области D_h, а все остальные точки — внутренними точками области D_h. Но чтобы можно было пользоваться уравнением (2.165), необходимо знать значения искомой функции U в граничных точках S_h сеточной области D_h. Тогда поступают следующим образом. Выполняют аппроксимацию заданных граничных условий (2.164) таким образом, чтобы стали известны значения искомой функции U в граничных точках S_h сеточной области D_h. Для каждого узла $A \in S_h$ (см. рис. 2.22) выбирают значение $U(A)$, равное $u(B)$, где $B \in S$ — ближайшая точка к A, лежащая на S. После того, как для всех точек S_h подобная процедура выполнена, можно считать, что аппроксимация граничных условий (2.164) проведена. В результате получают аппроксимирующее уравнение, которое мы обозначим следующим образом

$$U|_{S_h} = \varphi_A. \tag{2.166}$$

Аппроксимация граничных условий (2.164) описываемым выше способом называется сносом граничных условий, заданных на S, на граничные условия, заданные на S_h. При такой аппроксимации граничных условий вносится погрешность, величина которой зависит от близости S и S_h. Только в том случае, когда все граничные узлы попадут на S (см. п. 2.5.1), перенос граничных условий делать не следует. Отсюда вытекает, что сетку рекомендуется выбирать в зависимости от вида области D, где отыскивается решение задачи. Во всяком случае, иногда следует отказаться от квадратной сетки и рассматривать прямоугольную или треугольную сетку только потому, что в последнем случае граница S_h сеточной области D_h значительно ближе к границе S заданной области D.

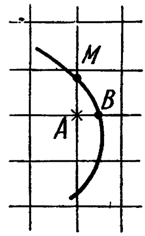

Рис. 2.22

Отметим также, что аппроксимация заданных граничных условий (2.164) может быть выполнена другим путем, отличным от описываемого здесь способа. Подробное обсуждение этого вопроса выходит за рамки нашего пособия.

Таким образом, в результате аппроксимации дифференциального уравнения (2.163) и граничных условий (2.164) описываемым здесь способом, мы получим следующую разностную краевую задачу

$$\left.\begin{aligned} &U_{i,k} = \frac{1}{4}\left[U_{i+1,k} + U_{i,k+1} + U_{i,k-1} + U_{i-1,k}\right]; \\ &U|_{S_h} = \varphi_A. \end{aligned}\right\} \tag{2.167}$$

Переходим к третьему, заключительному этапу метода сеток — решению полученной разностной краевой задачи.

Разностная краевая задача (2.167) представляет собой систему N линейных алгебраических уравнений с N неизвестными, если число узлов сетки, принадлежащих D, равно N: каждой точке (i, k) соответствует одно алгебраическое уравнение, связывающее значения U искомого решения в узлах сетки.

Разрешима ли система (2.167)? Если разрешима, то сколько она имеет решений?

Известно, что неоднородная система линейных алгебраических уравнений, каковой и является система (2.167), имеет единственное решение, если определитель этой системы отличен от нуля. Но этот определитель будет отличен от нуля, если соответствующая однородная система

$$\left.\begin{aligned} L_h[U] &= 0; \\ U|_{S_h} &= 0 \end{aligned}\right\} \tag{2.168}$$

имеет только тривиальное (нулевое) решение.

Теорема. *Разностная краевая задача (2.168) всегда имеет решение и притом единственное.*

Доказательство. Справедливость этой теоремы вытекает из замечательного свойства разностного оператора $L_h[U]$, названного принципом максимума. (Ср. со свойством 4 гармонических функций § 2.3.) Оказывается, если $L_h[U] = 0$ и $U \not\equiv \text{const}$, то U не принимает во внутренних точках сеточной области D_h ни своего наибольшего, ни своего наименьшего значения. Действительно, если предположить, что U принимает в некоторой точке (i, k) свое наибольшее (наименьшее) значение $U_{i,k}$, то это будет означать, что в соседних точках области D_h значения U будут меньше (больше) числа $U_{i,k}$. Тогда и среднее арифметическое этих значений будет, естественно, меньше (больше) этого наибольшего значения, т. е.

$$U_{i,k} > \frac{1}{4}[U_{i+1,k} + U_{i,k+1} + U_{i,k-1} + U_{i-1,k}]$$

$$\left(U_{i,k} < \frac{1}{4}[U_{i-1,k} + U_{i,k+1} + U_{i,k-1} + U_{i-1,k}]\right).$$

Но полученное равенство противоречит условию $L_h[U] = 0$, что и доказывает принцип максимума рассматриваемого разностного оператора. Теперь уже читатель без особого труда убедится в том, что система (2.168) имеет только нулевое решение. Ведь если система (2.168), кроме тривиального решения, имеет еще и отличное от нуля решение, то по принципу максимума оно должно достигать своего наибольшего (наименьшего) значения на S_h. Но на S_h значения U будут равны нулю (второе уравнение системы (2.168)), что и доказывает теорему.

Уместно отметить, что разрешимость соответствующей разностной краевой задачи может быть доказана и при использовании других разностных схем. Наибольшее распространение получила девятиточечная разностная схема (рис. 2.23), дающая аппроксимацию дифференциального оператора $L[u]$ с высокой степенью точности.

Подробные исследования показали, что полученное по методу сеток приближенное решение внутренней задачи Дирихле сходится к точному решению этой задачи. Мало того. Было установлено, что ошибки, возникающие при решении разностной краевой задачи, имеют тенденцию уменьшаться, т. е. рассматриваемое конечно-разностное уравнение устойчиво. Отсюда следует, что внутренняя задача Дирихле для уравнения Лапласа может быть приближенно решена по методу сеток со сколь угодно большой точностью.

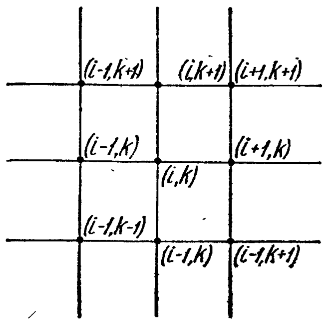

Рис. 2.23

2.5.3. Решение краевой задачи для уравнения теплопроводности

Решение нестационарных краевых задач методом сеток во многом отличается от решения задач стационарных. В задачах эллиптического типа разностная краевая задача, как мы видели, сводится к решению системы линейных алгебраических уравнений и основным вопросом является вопрос фактического решения этой системы уравнений. Ни вопросы сходимости найденного приближенного решения к точному решению, ни вопросы устойчивости конечно-разностного уравнения не волнуют исследователя по той причине, что наиболее употребительные разностные схемы решают эти вопросы в положительном смысле: решение сходится, соответствующее уравнение устойчиво.

Совершенно другая картина нас ждет при решении задач параболического и гиперболического типов. Начиная с выбора сетки (выбор квадратной сетки возможен здесь в исключительных случаях) и кончая исследованием полученного конечно-разностного уравнения на устойчивость, мы попадаем в более сложную ситуацию по сравнению с эллиптическим случаем. Теперь уже вопрос фактического решения конечно-разностного уравнения становится тривиальным и осуществляется «по шагам» (см. ч. 1, § 3. 11). Зато на первое место выступают вопросы сходимости и устойчивости, которые для нестационарных задач имеют первостепенное значение.

На примере решения простейшей краевой задачи для уравнения теплопроводности проиллюстрируем в очень краткой форме характерные особенности применения метода сеток к нестационарным задачам.

Итак, сформулируем задачу.

Найти решение $u(x, t)$ однородного уравнения теплопроводности

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < 1; \quad 0 \leqslant t \leqslant t_0, \tag{2.169}$$

удовлетворяющее начальному условию

$$u(x, t)|_{t=0} = u(x, 0) = f(x)$$

и однородным краевым условиям

$$u(x, t)|_{x=0} = u(0, t) = 0; \quad u(x, t)|_{x=1} = u(1, t) = 0.$$

Будем решать эту задачу методом конечных разностей. Сначала выберем сетку. Возьмем прямоугольную сетку

$$x = ih, \quad i = 0, 1, 2, \ldots, n,$$
$$y = kl, \quad k = 0, 1, 2, \ldots, r.$$

Здесь шаг по оси Ox равен $h = \frac{1}{n}$, а неравный ему шаг по оси Ot равен $l = \frac{t_0}{r}$.

Заметим, что к вопросу соотношения между шагами h и l мы еще вернемся. Дело в том, что в задачах подобного рода выбор шага по оси Ot существенно зависит от того, какой шаг выбран по оси Ox. Другими словами, соотношение между шагами выбранной сетки здесь нельзя считать произвольным.

Переходим к аппроксимации дифференциального уравнения и краевых условий.

Для аппроксимации частных производных $\frac{\partial^2 u}{\partial x^2}$ и $\frac{\partial u}{\partial t}$ воспользуемся уже известными нам разностными отношениями

$$\left(\frac{\partial u}{\partial t}\right)_{i,k} \approx \frac{u_{i,k+1} - u_{i,k}}{l};$$

$$\left(\frac{\partial^2 u}{\partial x^2}\right)_{i,k} \approx \frac{u_{i+1,k} - 2u_{i,k} + u_{i-1,k}}{h^2}.$$

Тогда конечно-разностное уравнение, аппроксимирующее заданное дифференциальное, будет иметь вид

$$\frac{U_{i,\,k+1}-U_{i,\,k}}{l}=\frac{U_{i+1,\,k}-2U_{i,\,k}+U_{i-1,\,k}}{h^2}.$$

Аппроксимация краевых условий дает

$$U_{i,\,0}=f(ih),\ i=1,\ 2,\ \ldots,\ n-1;$$
$$U_{0,\,k}=U_{n,\,k}=0,\ k=0,\ 1,\ 2,\ \ldots,\ r.$$

После некоторых преобразований и приведения подобных членов, мы получим следующую разностную краевую задачу

$$U_{i,\,k+1}=\frac{l}{h^2}\left(U_{i-1,\,k}+U_{i+1,\,k}\right)+\left(1-\frac{2l}{h^2}\right)U_{i,\,k}, \tag{2.170}$$

$$i=1,\ 2,\ \ldots,\ n-1;\ k=0,\ 1,\ 2,\ \ldots,\ r;$$

$$U_{i,\,0}=f(ih),\ i=1,\ 2,\ \ldots,\ n-1; \tag{2.171}$$

$$U_{0,\,k}=U_{n,\,k}=0;\ k=0,\ 1,\ \ldots,\ r. \tag{2.172}$$

Очевидно, уравнение (2.170) является самым простым из всех конечно-разностных уравнений, аппроксимирующих уравнение (2.169). Оно является четырехточечным уравнением, что соответствует рис. 2.24.

Прежде чем переходить к непосредственному решению разностной краевой задачи, внимательно еще раз вдумайтесь в ее содержание, обратите внимание на характер индексов в самом уравнении и в краевых условиях. Что здесь предполагается уже известным, что нужно отыскать? Почему пределы изменения индекса i одни, а индекса k — другие? С чем это связано?

Остановимся несколько подробнее на условиях (2.171) и (2.172). Условия (2.171) говорят о том, что значения искомого решения U известны на так называемом нулевом слое (k-м слоем мы будем называть совокупность узлов $x_i=ih$, $i=0$, $1,\ 2,\ \ldots,\ n$, $t_k=kl$), за исключением двух узлов (0, 0) и (n, 0). Но в этих узлах, как говорят условия (2.172), значения искомого решения равны нулю. Значит, известны все значения U на нулевом слое. Это обстоятельство играет существенную роль в решении полученной разностной задачи. Действительно, обратите внимание на правую часть уравнения (2.170). Второй индекс у всех слагаемых правой части равен k. Если k положить равным нулю, то правая часть рассматриваемого уравнения в силу предыдущего (ведь на нулевом слое значения U известны !) может быть вычислена. А это будет означать, что все значения искомого решения U будут найдены на первом слое, ибо в левой части уравнения (2.170) второй индекс равен $k+1$. Вот и наметился план решения полученной разностной краевой задачи (рис. 2.25).

Первый шаг при решении уравнения (2. 170) состоит в вычислении всех $n-1$ значений $U_{i,1}$, $i=1, 2, \ldots, n-1$, по известным начальным значениям $U_{i,0}$, заданным соотношениями (2.171) и (2.172).

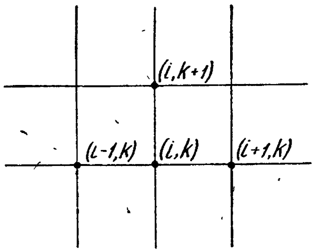

Рис. 2.24

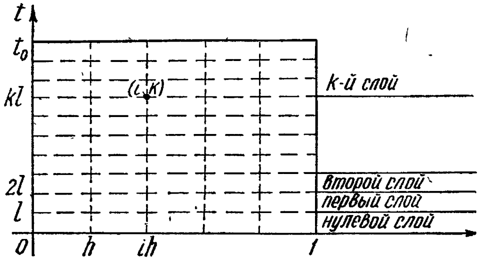

Рис. 2.25

Второй шаг состоит в вычислении всех $n-1$ значений $U_{i,2}$, $i=1, 2, \ldots, n-1$, по только что найденным значениям $U_{i,1}$ и краевым условиям (2.172), где $U_{0,1}=U_{n,1}=0$. Аналогично находятся значения $U_{i,3}$, $U_{i,4}$ и т. д.

Этот процесс продолжается до тех пор, пока не будут просчитаны все значения искомого решения на всех слоях выбранной сетки.

Описываемый процесс нахождения решения уравнения (2.170) носит название решения «по шагам». Простота, с которой раз за разом удается вычислить все значения искомого решения U, таит в себе целый ряд опасностей, связанных, в основном, с характером полученного таким способом решения разностной краевой задачи: оно очень чувствительно к фактору устойчивости.

Исследование конечно-разностного уравнения на устойчивость представляет проблему чрезвычайно трудную. Только в простейших случаях удается решить задачу на устойчивость элементарными средствами. Поискам эффективных критериев устойчивости разностных схем посвящены многие работы советских и зарубежных ученых.

Условие устойчивости для рассматриваемого здесь конечно-разностного уравнения (2.170) может быть получено следующим образом.

Коэффициенты правой части уравнения (2.170) содержат выражение вида $\frac{l}{h^2}$. Поэтому удобно ввести следующее обозначение

$$\tau = \frac{l}{h^2}.$$

Тогда исследуемое уравнение может быть переписано в виде

$$U_{i,\,k+1} = \tau U_{i-1,\,k} + (1 - 2\tau) U_{i,\,k} + \tau U_{i+1,\,k}. \tag{2.173}$$

Как известно (см. п. 2.5.1) проблема устойчивости конечно-разностного уравнения (2.173) состоит в том, чтобы найти условия, при которых численная погрешность

$$\rho_{i,\,k} = U_{i,\,k} - \tilde{U}_{i,\,k}$$

при возрастании k равномерно по всем i стремится к нулю или остается ограниченной. Здесь $U_{i,\,k}$ обозначает точное решение уравнения (2. 173), а $\tilde{U}_{i,\,k}$ — его фактическое решение.

Предположим, что выполнено округление всех значений U на нулевом слое сетки. Очевидно, погрешность, полученная в результате такого округления, может быть выражена разностью

$$\rho_{i,\,0} = f(ih) - \tilde{f}(i\,h),$$

а уравнение, связывающее значения $\tilde{U}_{i,\,k}$, будет иметь вид

$$\tilde{U}_{i,\,k+1} = \tau \tilde{U}_{i-1,\,k} + (1 - 2\tau) \tilde{U}_{i,\,k} + \tau \tilde{U}_{i+1,\,k}. \tag{2.174}$$

Покажем, что при $\tau \leqslant \frac{1}{2}$ уравнение (2.173) будет устойчивым, т. е. погрешность $\rho_{i,\,k}$ будет ограниченной при возрастании k.

Вычтем уравнение (2.174) из уравнения (2.173). Тогда для погрешности $\rho_{i,\,k}$ получим следущее соотношение

$$\rho_{i,\,k+1} = \tau\rho_{i-1,\,k} + (1 - 2\tau)\rho_{i,\,k} + \tau\rho_{i+1,\,k}. \tag{2.175}$$

По предположению $\tau \leqslant \frac{1}{2}$. Значит, правая часть уравнения (2.175) является линейной комбинацией значений ρ с неотрицательными коэффициентами, сумма которых равна единице. Тогда получим следующую оценку правой части уравнения (2.175)

$$|\tau\rho_{i+1,\,k} + (1 - 2\tau)\rho_{i,\,k} + \tau\rho_{i-1,\,k}| \leqslant [\tau + (1 - 2\tau) +$$

$$+\tau]\max(|\rho_{i-1,\,k}|,\ |\rho_{i,\,k}|,\ |\rho_{i+1,\,k}|)=\max(|\rho_{i-1,\,k}|,$$
$$|\rho_{i,\,k}|,\ |\rho_{i+1,\,k}|)\leqslant\max_{i=1,\,2,\,\dots,\,n-1}|\rho_{i,\,k}|.$$

Используя эту оценку, из уравнения (2.175) выводим неравенство

$$|\rho_{i,\,k+1}|\leqslant\max_i|\rho_{i,\,k}|. \tag{2.176}$$

В связи с тем, что правая часть неравенства (2.176) не зависит от i, то в левой части вместо $|\rho_{i,\,k+1}|$ можно написать $\max\limits_i|\rho_{i,\,k+1}|$. Окончательно

$$\max_i|\rho_{i,\,k+1}|\leqslant\max_i|\rho_{i,\,k}|.$$

Полученное неравенство известно под названием принципа максимума для рассматриваемого конечно-разностного уравнения. Оно говорит о том, что $\max\limits_i|\rho_{i,\,k+1}|$ не возрастает с ростом k.

Используя последнее неравенство, получаем цепочку соотношений

$$\max_i|\rho_{i,\,k}|\leqslant\max_i|\rho_{i,\,k-1}|;$$
$$\max_i|\rho_{i,\,k-1}|\leqslant\max_i|\rho_{i,\,k-2}|;$$
$$\cdots\cdots\cdots\cdots$$
$$\max|\rho_{i,\,1}|\leqslant\max|\rho_{i,\,0}|.$$

Складывая все эти неравенства почленно и упрощая, найдем

$$\max_i|\rho_{i,\,k+1}|\leqslant\max_i|\rho_{i,\,0}|.$$

Последнее неравенство будет иметь место для всех k, что и доказывает устойчивость исследуемой разностной схемы. Действительно, погрешность $\rho_{i,\,0}$ в начальных данных (на нулевом слое) при возрастании k не увеличивается.

Приведенное доказательство можно считать достаточно общим, хотя оно было приведено в предположениях, суживающих диапазон возможных здесь случаев: исследована погрешность лишь на начальном слое, хотя известно, что на практике погрешности округления вносятся на каждом слое сетки. Читатель без особого труда может перенести приведенные выкладки на случай любого m-го слоя сетки; а затем, используя принцип суперпозиции, справедливый для линейного конечно-разностного уравнения, доказать устойчивость исследуемого уравнения в самом общем случае.

Итак, если

$$\tau = \frac{l}{h^2} \leqslant \frac{1}{2}, \tag{2.177}$$

то конечно-разностное уравнение (2.170) устойчиво. Отметим, что полученное условие (2.177) является не только достаточным, но и необходимым условием устойчивости уравнения (2.170).

Можно показать, что при

$$\frac{l}{h^2} > \frac{1}{2}$$

уравнение (2.170) будет уже неустойчивым. Доказательство этого факта мы проводить не будем, так как оно далеко не тривиально (предыдущие рассуждения, проведенные для доказательства устойчивости, здесь не имеют силы). Вместо этого мы проиллюстрируем случай неустойчивости элементарным примером, который еще раз подчеркнет характер решения рассматриваемой разностной краевой задачи.

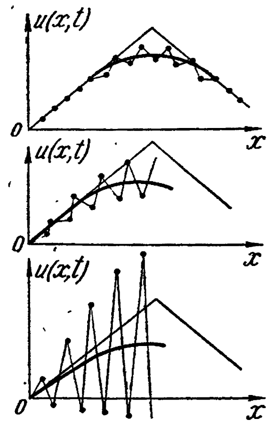

Рис. 2.26

Решается задача с начальной функцией $f(x)$, имеющей вид

$$f(x) = \begin{cases} x & \text{при} \quad 0 \leqslant x \leqslant \frac{1}{2}; \\ 1 - x & \text{при} \quad \frac{1}{2} < x \leqslant 1. \end{cases}$$

Результаты вычислений показаны на рис. 2.26 при соотношении шагов l и h, связанных неравенством $\frac{l}{h^2} = \frac{5}{9} > \frac{1}{2}$. Сплошными линиями изображено точное решение дифференциальной краевой задачи, а точками — решение соответствующей разностной краевой задачи. На рисунке заметно, как быстро накапливается погрешность в решении разностной задачи: после сравнительно малого числа шагов она увеличивается до неприемлемой величины.

Проведенное нами исследование показало, что поведение решения конечно-разностной краевой задачи существенно зависит от размеров сетки, а точнее, от соотношения между шагами сетки вдоль оси Ox и вдоль оси Ot. Если, например, шаг вдоль оси Ox уже выбран, то шаг вдоль временной оси следует выбирать с учетом характера устойчивости соответствующей разностной схемы.

Интересно отметить, что, если условие (2.177) выполняется, то, как правило, решение разностной краевой задачи сходится к решению данной дифференциальной задачи, т. е., грубо говоря, из устойчивости уравнения следует сходимость его решения.

В заключение сделаем ряд замечаний.

Читателю может показаться, что при применении метода сеток используется лишь разобранная здесь разностная схема. На самом деле существуют и используются многие другие разностные схемы, которые в некоторых отношениях даже лучше приведенной решают поставленную задачу. Подробное знакомство с этими разностными схемами выходит за рамки нашего пособия. Заметим лишь, что существуют также разностные уравнения, которые являются устойчивыми при любом значении $\frac{l}{h^2}$. К сожалению, для их решения приходится прибегать к системе линейных алгебраических уравнений как в эллиптическом случае, т. е. эти разностные уравнения являются неявными, а поэтому неудобны при практическом применении. Желательными для практики будут такие разностные схемы, которые занимают в известном смысле промежуточное положение между описываемыми здесь крайними типами уравнений. Они должны сочетать более слабые ограничения устойчивости со сравнительно несложным проведением вычислений. Поискам таких «оптимальных» разностных схем посвящены многие работы отечественных и зарубежных ученых

Вопросы для повторения

1. Расскажите о системах с сосредоточенными и распределенными параметрами. Чем они принципиально отличаются друг от друга?

2. Какие задачи теории поля описываются дифференциальными уравнениями в частных производных?

3. Сформулируйте определение обыкновенного дифференциального уравнения n-го порядка.

4. Перечислите основные свойства решений обыкновенного дифференциального уравнения n-го порядка.

5. Выпишите дифференциальное уравнение в частных производных n-го порядка.

6. Расскажите о структуре общего решения дифференциального уравнения в частных производных. Приведите примеры.

7. Дайте геометрическую интерпретацию решения обыкновенного дифференциального уравнения и решения уравнения в частных производных.

8. Какое уравнение называется линейным дифференциальным уравне-

нием в частных производных второго порядка? Приведите примеры таких уравнений.

9. Перечислите свойства решений линейного уравнения в частных производных(принцип суперпозиции).

10. Выведите уравнение колебания струны.

11. Выпишите уравнение движения среды в дифференциальной форме (трехмерное волновое уравнение).

12. Выведите уравнение теплопроводности.

13. Получите уравнение установившихся процессов (уравнения Лапласа и Пуассона).

14. Расскажите о специфике постановки задач математической физики.

15. Сформулируйте краевые задачи для волнового уравнения.

16. Сформулируйте краевые задачи для уравнения теплопроводности.

17. Сформулируйте краевые задачи для уравнения Лапласа. Задача Дирихле и задача Неймана.

18. Расскажите о корректности постановки задач математической физики.

19. В чем состоит идея приведения уравнения второго порядка к каноническому виду?

20. Запишите характеристическое уравнение. Что вы называете характеристиками дифференциального уравнения?

21. Какое уравнение называется уравнением гиперболического типа? Приведите примеры.

22. Какое уравнение называется уравнением параболического типа? Прибедите примеры.

23. Какое уравнение называется уравнением эллиптического типа? Приведите примеры.

24. Получите решение задачи Коши уравнения колебания струны методом характеристик.

25. Дайте физическое истолкование решения Даламбера уравнения колебания струны.

26. Выведите телеграфное уравнение.

27. Сформулируйте задачу Коши для телеграфного уравнения.

28. Дайте решение телеграфного уравнения для линии без потерь.

29. Дайте решение телеграфного уравнения для линии без искажений.

30. В чем заключается сущность метода Фурье (метода разделения переменных)?

31. Получите решение волнового уравнения, описывающего свободные колебания закрепленной струны методом Фурье.

32. Дайте физическую интерпретацию решения уравнения, описывающего свободные колебания закрепленной струны. Получите решение краевой задачи методом Фурье.

34. Сформулируйте задачу Дирихле для круга. Решите эту задачу методом Фурье.

35. Приведите общую схему метода Фурье.

36. В чем заключается сущность метода функции Грина?

37. Приведите примеры гармонических функций.

38. Перечислите свойства гармонических функций.

39. Какая функция называется функцией Грина задачи Дирихле для уравнения Лапласа?

40. Дайте решение задачи Дирихле методом функции Грина.

41. Получите решение задачи Дирихле для шара методом функции Грина.

42. Выпишите формулы прямого и обратного преобразований Фурье (Лапласа).

43. В чем состоит общая схема применения метода интегральных пре-

образований при решении дифференциальных уравнений в частных производных?

44. Решите задачу о распространении тепла в неограниченном (полуограниченном) стержне методом интегральных преобразований.

45. Расскажите о конечных интегральных преобразованиях и их применении.

46. В чем состоит суть приближенных методов решения задач математической физики?

47. Расскажите о содержании (идее) метода сеток (метода конечных разностей).

48. Приведите схему метода сеток.

49. Какие вопросы требуют дополнительного исследования при применении метода сеток?

50. Как вы понимаете сходимость приближенного решения к точному решению краевой задачи?

51. Что следует понимать под устойчивостью конечно-разностного уравнения? Какая существует связь между сходимостью решения и устойчивостью?

52. Приведите примеры наиболее распространенных разностных схем.

53. Расскажите о точности аппроксимации разностными отношениями частных производных. Приведите примеры.

54. Приведите два способа построения конечно-разностных уравнений.

55. Сформулируйте разностную краевую задачу Дирихле для уравнения Лапласа.

56. Получите конечно-разностное уравнение, аппроксимирующее уравнение Лапласа.

57. Как осуществляется аппроксимация граничных условий в случае задачи Дирихле?

58. Докажите теорему о существовании и единственности решения краевой задачи Дирихле.

59. Как решаются вопросы сходимости и устойчивости для стационарных задач (уравнений эллиптического типа)?

60. Расскажите о применении метода сеток к решению нестационарных задач.

61. Получите конечно-разностное уравнение для уравнения теплопроводности.

62. Расскажите о методе решения «по шагам» разностного уравнения теплопроводности.

63. Получите условия устойчивости разностного уравнения теплопроводности.

Упражнения к разделу I

I. Привести к каноническому виду уравнения:

1. $$3\frac{\partial^2 u}{\partial x^2}+2\frac{\partial^2 u}{\partial x\,\partial y}-\frac{\partial^2 u}{\partial y^2}+2\frac{\partial u}{\partial x}+3\frac{\partial u}{\partial y}=0.$$

2. $$\frac{\partial^2 u}{\partial x^2}+4\frac{\partial^2 u}{\partial x\,\partial y}+5\frac{\partial^2 u}{\partial y^2}+\frac{\partial u}{\partial x}+2\frac{\partial u}{\partial y}=0.$$

3. $$\frac{\partial^2 u}{\partial x^2}-2\frac{\partial^2 u}{\partial x\,\partial y}+\frac{\partial^2 u}{\partial y^2}+2\frac{\partial u}{\partial x}+\frac{\partial u}{\partial y}+4u=0.$$

4. $x^2 \dfrac{\partial^2 u}{\partial x^2} + 2xy \dfrac{\partial^2 u}{\partial x\, \partial y} + y^2 \dfrac{\partial^2 u}{\partial y^2} = 0.$

5. $4y^2 \dfrac{\partial^2 u}{\partial x^2} - e^{2x} \dfrac{\partial^2 u}{\partial y^2} - 4y^2 \dfrac{\partial u}{\partial x} = 0.$

6. $y^2 \dfrac{\partial^2 u}{\partial x^2} + 2xy \dfrac{\partial^2 u}{\partial x\, \partial y} + 2x^2 \dfrac{\partial^2 u}{\partial y^2} + y \dfrac{\partial u}{\partial y} = 0.$

II. Найти общее решение уравнения:

7. $\dfrac{\partial^2 u}{\partial x^2} - 2 \sin x \dfrac{\partial^2 u}{\partial x\, \partial y} - \cos^2 x \dfrac{\partial^2 u}{\partial y^2} - \cos x \dfrac{\partial u}{\partial y} = 0.$

8. $x^2 \dfrac{\partial^2 u}{\partial x^2} - y^2 \dfrac{\partial^2 u}{\partial y^2} - 2y \dfrac{\partial u}{\partial y} = 0.$

9. $x^2 \dfrac{\partial^2 u}{\partial x^2} - 2xy \dfrac{\partial^2 u}{\partial x\, \partial y} + y^2 \dfrac{\partial^2 u}{\partial y^2} + x \dfrac{\partial u}{\partial x} + y \dfrac{\partial u}{\partial y} = 0.$

10. Уравнений 1, 5.

III. Методом характеристик решить краевые задачи:

11. Найти решение уравнения

$$\frac{\partial^2 u}{\partial x^2} + 2 \frac{\partial^2 u}{\partial x\, \partial t} - 3 \frac{\partial^2 u}{\partial t^2} = 0, \quad -\infty < x < +\infty, \quad t > 0,$$

удовлетворяющее начальным условиям

$$u\Big|_{t=0} = 3x^2; \quad \frac{\partial u}{\partial t}\Big|_{t=0} = 0$$

(задача Коши).

12. Найти решение уравнения

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < +\infty, \quad t > 0,$$

удовлетворяющее начальным условиям

$$u\Big|_{t=0} = \sin x; \quad \frac{\partial u}{\partial t}\Big|_{t=0} = 0$$

(задача Коши).

13. Найти решение уравнения

$$\frac{\partial^2 u}{\partial t^2} = 4 \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < +\infty, \quad t > 0,$$

удовлетворяющее условиям

$$u\Big|_{t=0} = 0; \quad \frac{\partial u}{\partial t}\Big|_{t=0} = A \sin x$$

(задача Коши).

14. Бесконечной струне сообщена только на отрезке $-c \leqslant x \leqslant c$ поперечная начальная скорость $v_0 = \text{const}$. Решить задачу о колебании этой струны. Построить профиль струны для моментов времени $t_k = \dfrac{c}{2a} k$, $k = 1, 2, 3$.

15. Полуограниченная однородная струна $0 \leqslant x < \infty$ с закрепленным концом $x = 0$ возбуждена начальным отклонением

$$u(x, 0) = \begin{cases} 0 & \text{при } 0 \leqslant x \leqslant l, \\ -\sin \dfrac{\pi x}{l} & \text{при } l < x \leqslant 2l, \\ 0 & \text{при } 2l < x < \infty. \end{cases}$$

Решить задачу о колебании этой струны, предполагая, что начальные скорости отсутствуют.

IV. Методом разделения переменных (методом Фурье) решить краевые задачи:

16. Найти закон колебания струны длиной l, расположенной на отрезке $[0, l]$ и закрепленной на концах, если в начальный момент струне придана форма кривой $f(x) = \dfrac{x(l-x)}{8l}$, а затем струна отпущена без начальной скорости.

17. Однородная струна длиной l натянута между точками $x = 0$ и $x = l$. В точке $x = c$ струна оттягивается на небольшое расстояние h от положения равновесия и в момент $t = 0$ отпускается без начальной скорости. Определить отклонение $u(x, t)$ струны для любого момента времени.

18. Решить задачу о колебании струны с жестко закрепленными концами под действием импульса P, сообщенного струне в момент времени $t = 0$ в точке $x = x_0$.

19. Найти решение уравнения

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < l, \quad t > 0,$$

удовлетворяющее краевым условиям

$$u(0, t) = u(l, t) = 0$$

и начальным условиям

$$u(x, 0) = \begin{cases} x & \text{при } 0 < x \leqslant \dfrac{l}{2}, \\ l - x & \text{при } \dfrac{l}{2} \leqslant x < l. \end{cases}$$

20. Дан тонкий однородный стержень длиной l, изолированный от внешнего пространства, начальная температура которого равна

$$f(x) = \frac{cx(l-x)}{l^2}.$$

Концы стержня поддерживаются при температуре, равной нулю. Определить температуру стержня в момент времени $t > 0$.

21. На струну длиной l постоянно действует внешняя возбуждающая сила, равная (в расчете на единицу массы струны) $\dfrac{a^2}{100} \sin \dfrac{a\pi t}{l}$; здесь a— постоянная, фигурирующая в уравнении колебания струны. Найти закон колебания струны, если начальное отклонение и начальная скорость равны нулю, а концы струны закреплены.

22. Решить уравнение

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} + bx(x - l)$$

при нулевых начальных и краевых условиях

$$u(0,\ t)=u(l,\ t)=0.$$

23. Дан тонкий однородный стержень, начальная температура которого равна 0°. На конце $x=l$ температура поддерживается равной 0°, а на конце $x=0$ она растет пропорционально протекшему времени: $u(0,\ t)=At$, где A — постоянная. Найти закон изменения температуры внутри стержня.

24. Найти решение уравнения Лапласа внутри кольца $1\leqslant R\leqslant 2$, удовлетворяющее краевым условиям

$$u\,|_{R=1}=0;\ u\,|_{R=2}=Ay.$$

25. Найти решение уравнения Лапласа в прямоугольнике R: $0\leqslant x\leqslant a$; $0\leqslant y\leqslant b$, удовлетворяющее краевым условиям:

$$u(0,\ y)=0;\ u(a,\ y)=0;\ u(x,\ 0)=A\left(1-\frac{x}{a}\right);\ u(x,\ 0)=0.$$

26. Однородная квадратная мембрана, имеющая в начальный момент времени $t=0$ форму

$$u(x,\ y,\ 0)=Axy(l-x)(l-y),\quad A=\mathrm{const};$$

начала колебаться без начальной скорости. Изучить свободные колебания мембраны, закрепленной по контуру.

V. Методом интегральных преобразований решить краевые задачи:

27. Найти распределение температуры в бесконечном стержне, если в начальный момент температура в стержне была распределена следующим образом:

$$u(x,\ t)\,|_{t=0}=\begin{cases}0 & \text{при}\quad |x|>h;\\ -T & \text{при}\quad -h<x<0;\\ T & \text{при}\quad 0<x<h.\end{cases}$$

28. Начальное распределение температуры внутри бесконечного стержня задается формулой

$$u(x,\ t)\,|_{t=0}=T_0e^{-\frac{x^2}{b^2}},$$

где T_0 и b — заданные постоянные величины. Найти закон распределения температуры внутри стержня в любой момент времени $t>0$.

29. Найти решение уравнения

$$\frac{\partial^2u}{\partial t^2}=a^2\frac{\partial^2u}{\partial x^2},\quad 0<x<\infty,\quad t>0,$$

удовлетворяющее нулевым начальным условиям и краевому условию

$$u(0,\ t)=t^2.$$

30. Найти решение уравнения Лапласа

$$\frac{\partial^2u}{\partial x^2}+\frac{\partial^2u}{\partial y^2}=0,\quad 0\leqslant x<\infty,\quad 0\leqslant y<\infty,$$

при краевых условиях

$$u\,\Big|_{y=0}=0;\quad \frac{\partial u}{\partial x}\,\Big|_{x=0}=f(y).$$

31. Найти решение уравнения

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < \infty, \quad t > 0,$$

удовлетворяющее условиям

$$\left.\frac{\partial u}{\partial x}\right|_{x=0} = 0; \quad \left. u(x, y)\right|_{y=0} = f(x).$$

32. Плоская граница полубесконечного тела, начиная с момента $t = 0$, поддерживается при температуре, изменяющейся по закону $u|_{x=0} = f(t)$. Найти значение температуры в каждой точке тела, принимая его начальную температуру равной нулю.

Рассмотреть частные случаи, соответствующие следующим законам изменения температуры на границе: а) $f(t) = T_0$; б) $f(t) = At$; в) $f(t) = T_0 \sin \omega t$.

33. Найти решение уравнения

$$\frac{\partial u}{\partial t} = a^2 \frac{\partial^2 u}{\partial x^2}, \quad 0 < x < \pi, \quad t > 0,$$

при начальном условии

$$u(x, 0) = x^2$$

и краевых условиях

$$\frac{\partial u(0, t)}{\partial x} = \frac{\partial u(\pi, t)}{\partial x} = 0.$$

Литература к разделу I

1. Араманович И. Г. и Левин В. И. Уравнения математической физики. М., 1964.
2. Арсенин В. Я. Математическая физика. М., 1966.
3. Берс Л., Джон Ф., Шехтер М. Уравнения с частными производными. М., 1966.
4. Березкин И. С. и Жидков Н. П. Методы вычислений, т. 1,2. М., 1960.
5. Будак Б. М., Самарский А. А., Тихонов А. И. Сборник задач по математической физике. М., 1956.
6. Вазов В. и Форсайт Дж. Разностные методы решения дифференциальных уравнений в частных производных. М., 1934.
7. Вебстер А. и Сеге Г. Дифференциальные уравнения в частных производных математической физики. М.—Л., 1934.
8. Гринберг Г. А. Избранные вопросы математической теории электрических и магнитных явлений. М., 1948.
9. Годунов С. К. и Рябенький В. С. Введение в теорию разностных схем. М., 1962.
10. Деч Г. Руководство к практическому применению преобразования Лапласа. М., 1965.
11. Зоммерфельд А. Дифференциальные уравнения в частных производных физики. М., 1950.
12. Канторович Л. В. и Крылов В. И. Приближенные методы высшего анализа. М.—Л., 1962.
13. Карплюс У. Моделирующие устройства для решения задач теории поля. М., 1962.

14. Коллатц Л. Численные методы решения дифференциальных уравнений. М., 1953.
15. Колобов А. М. Избранные главы высшей математики, ч. I. Мн., 1965.
16. Колобов А. М. и Черенкова Л. П. Избранные главы высшей математики, ч. 2, Мн., 1967.
17. Кошляков Н. С., Глинер Э. Б., Смирнов М. М. Основные дифференциальные уравнения математической физики. М., 1962.
18. Курант Р. и Гильберт Д. Методы математической физики, т. I, II. М.—Л., 1951.
19. Лебедев Н. Н., Скальская И. П., Уфлянд Я. С. Сборник задач по математической физике. М., 1955.
20. Лыков А. В. Теория теплопроводности. М., 1952.
21. Математика, ее содержание, методы и значение, т. II. М., 1956.
22. Милн В. Э. Численное решение дифференциальных уравнений. М., 1955.
23. Мисюркеев И. В. Сборник задач по методам математической физики. М., 1964.
24. Михлин С. Г. Курс математической физики. М., 1968.
25. Очан Ю. С. Методы математической физики. М., 1965.
26. Очан Ю. С. Сборник задач по методам математической физики. М., 1967.
27. Петровский И. Г. Лекции об уравнениях с частными производными. М.—Л., 1950.
28. Положий Г. Н. и др. Математический практикум. М., 1960.
29. Рихтмайер Р. Д. Разностные методы решения краевых задач. М., 1960.
30. Саульев В. К. Интегрирование уравнений параболического типа методом сеток. М., 1960.
31. Смирнов В. И. Курс высшей математики. М., 1953.
32. Смирнов М. М. Задачи по уравнениям математической физики. М., 1968.
33. Соболев С. Л. Уравнения математической физики. М., 1954.
34. Тиман А. Ф. и Трофимов В. Н. Введение в теорию гармонических функций. М., 1968.
35. Тихонов А. Н. и Самарский А. А. Уравнения математической физики. М., 1966.
36. Трантер К. Дж. Интегральные преобразования в математической физике. М., 1956.
37. Трикоми Ф. Лекции по уравнениям в частных производных. М., 1957.
38. Франк Т. и Мизес Р. Дифференциальные и интегральные уравнения математической физики. М., 1937.
39. Шварц Л. Математические методы для физических наук. М. 1965.

Раздел II

ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ ЛОГИКИ

Во второй половине прошлого столетия в математике, в связи с множеством новых фундаментальных открытий, сложилась обстановка, вынуждающая дать строгие, научно обоснованные доказательства истинности отправных положений и средств вывода математики.

Это потребовало разработки специальной теории со своим вычислительным аппаратом. Объектами изучения в ней должны были быть высказанные человеком мысли, системы умозаключений, методы доказательств, способы установления истинности или ложности утверждений.

Такой теорией, обслуживающей внутренние потребности математики, и явилась математическая логика.

В 30-х годах текущего столетия было обнаружено, что созданный в математической логике вычислительный аппарат, так называемую алгебру высказываний, можно интерпретировать на электрических цепях. После этого алгебра высказываний нашла широкое применение во многих прикладных науках. В настоящее время можно сказать, что математическая логика в целом стала одной из основных составляющих кибернетики со всеми многочисленными ее разветвлениями и приложениями.

Вследствие этого стало необходимым включение определенных разделов математической логики в программы университетов, пединститутов, технических вузов, где изучаются электроника, вычислительная техника, автоматика.

Таким образом, мы видим две области приложения математической логики: с одной стороны —основания математики, сугубо теоретические иселедования с наивысшей степенью абстракции; с другой—прикладные теории, решающие практические задачи производства и управления. В соответствии с этим существующую в настоящее время литературу по математической логике можно разделить на две группы.

К первой отнесем книги, в которых излагается сама математическая логика, ее вопросы и проблемы.

Ко второй — книги, которые используют в практических целях отдельные ее идеи, понятия и вычислительный аппарат.

Книги первой группы написаны в строгом стиле, специфическим языком и для неспециалиста труднодоступны.

Книги второй группы значительно проще. Чаще это главы или даже отдельные параграфы книг по тому или иному техническому вопросу. В них понятия математической логики и вычислительный аппарат даются упрощенно, без общей связи, только так, как это нужно для изучения вопросов, излагаемых в соответствующей книге.

В настоящем учебном пособии подбор материала и стиль изложения имеют целью прежде всего познакомить читателя с основными понятиями математической логики, ее вычислительным аппаратом, не привязывая его к каким-то конкретным техническим приложениям.

В соответствии с этим вначале дается представление о математической логике как о предмете. Далее, даются определения основных понятий математической логики и систематически излагается вычислительный аппарат — алгебра высказываний. Этих сведений из математической логики достаточно для решения практических вопросов во многих технических приложениях.

В последней главе изучаются вопросы интерпретации понятий математической логики на контактных и бесконтактных электрических цепях, использование алгебры высказываний для анализа и синтеза функциональных схем устройств вычислительной техники, рассматриваются вопросы минимизации формул и соответствующих им схем. Дан краткий обзор других областей, в которых математическая логика находит широкое и плодотворное приложение.

Глава 1. АЛГЕБРА ВЫСКАЗЫВАНИЙ

§ 1. 1. ПРЕДМЕТ МАТЕМАТИЧЕСКОЙ ЛОГИКИ

Возникновение математической логики уходит в глубокую древность. Предшественницей ее была формальная, или традиционная, логика, создателем которой считают Аристотеля.

Из истории известно, что в античном мире вопросы политики, проблемы естествознания решались на публичных дискуссиях. В полемике побеждал тот оратор, который мог доказать истинность своих утверждений путем строгих логических умозаключений, вытекающих из посылок, признанных истинными обеими спорящими сторонами.

Исследования и обобщения законов доказательств истинности логических умозаключений и явились предметом формальной логики. Аристотель сумел подметить основные закономерности логических умозаключений и сформулировал соответствующие законы: закон противоречия, закон исключенного третьего.

В дальнейшем были сформулированы еще закон тождества и закон достаточного основания.

В рамках формальной логики эти законы без существенных изменений сохранились до сих пор.

В средние века логика Аристотеля развития не получила.

Коренной сдвиг произошел в XVII в., когда в математику вторглась и заняла господствующее положение переменная, заменив постоянную. Этот революционный скачок, изменивший основы и методы решения практических задач, был подготовлен многовековым развитием математики постоянных и вызван к жизни материальными потребностями общества в эпоху Возрождения.

Благодаря переменной величине появились понятия бесконечно малой и бесконечно большой величин, предела, функции, ее непрерывности и ряд других фундаментальных понятий математического анализа.

Потребности производства ставили перед наукой, в том числе перед математикой, новые и новые проблемы. Созданные в дифференциальном и интегральном исчислении методы позволяли успешно справляться с решением большинства возникающих задач. Однако многие ученые и прежде всего Лейбниц и Ньютон заметили, что доказательства теорем и методы решения в математическом анализе не свободны от существенных логических недостатков. Более того, основы фундаментальных понятий опирались на положения, которые по существовавшим тогда представлениям были ложными. Например, предел переменной величины α понимался как ее «последнее» значение. Для бесконечно малой таким «последним» значением был нуль. Отсюда вытекало, что в математике переменных справедливо равенство $A+\alpha=A$, что с позиций математики постоянных неверно.

Еще большие сомнения вызывала основная формула дифференциального исчисления $\frac{dx}{dy}=\frac{0}{0}$, которая следовала из определения предела как «последнего» значения бесконечно малых приращений dy и dx.

Ньютон и Лейбниц пытались доказать истинность отправных положений математического анализа. Для доказательств они использовали существовавшие тогда понятия математики постоянных, а за образец логической строгости принимали аксиоматическое построение геометрии, описанное Евклидом в его «Началах».

Ньютон в своих доказательствах пытался использовать выводы механики и геометрии, истинность которых с такой очевидностью подтверждалась опытом.

Лейбниц пошел другим путем. Он пытался доказать истинность оснований высшей математики с помощью логических умозаключений. Для этого он обратился к аристотелевской формальной логике, значительно обогатив ее в направлении формализации и математической строгости.

Им была введена символика и осуществлена попытка распространить буквенное исчисление на логику. Получилось своеобразное «исчисление рассуждений», похожее на современное «исчисление высказываний» и позволявшее доказывать истинность некоторых логических следствий по известным значениям истинности исходных высказываний. Однако попытка решить основную задачу: логически доказать теоремы, лежащие в основе дифференциального и интегрального исчислений, Лейбницу не удалась. И вообще, созданное им «исчисление рассуждений» практического применения не нашло и должного развития не получило.

В последующие годы появляется ряд новых работ по логике. Буль (1854) впервые ввел в логические исследования операции математического типа. Дальнейшее развитие этого направления находим в работах Гроссмана

(1872), Шредера (1877), Джевонса (1873) и завершают этот этап труды казанского астронома и математика П. С. Порецкого (1887, 1888). Порецкий ввел удобную символику, обосновал и логически проанализировал все операции и полученные результаты. Он решил две основные задачи:

1) получение всех логических следствий из данных посылок;

2) определение тех посылок, из которых вытекает данное следствие.

Однако и эти довольно значительные результаты оказались недостаточными, чтобы формальная логика определилась как самостоятельная математическая теория и оказала бы заметное влияние на развитие естествознания. К этому времени еще недоставало количественных накоплений, необходимых для перехода формальной логики в новое качественное состояние. Произошло это только в конце XIX и начале XX вв.

Конец XVIII и начало XIX вв. вошли в историю технического прогресса как период бурного развития машинной техники, базировавшейся на использовании парового двигателя.

Технические достижения вызвали интенсивное и широкое развитие естествознания, в том числе математики.

Этот период бурного развития математики имеет примечательную особенность. Состоит она в том, что развивались и решались прикладные вопросы, выдвигаемые требованием растущего производства. В соответствии с этим создавались методы решения практических задач, и если таковые находились и они удовлетворяли практике, то задача считалась решенной. Вопросы обоснования математики отодвигались на второй план. Практика не ставила остро вопрос о разрешении противоречий в основах математических теорий. А между тем парадоксов обнаруживалось все больше и больше.

Не находили удовлетворительного объяснения основы теории чисел. Как доказать правило знаков $(-a)\ (-b)=ab$? Истинна ли пропорция $1:(-1)=(-1):1$? (Как большее, деленное на меньшее, может быть равно меньшему, деленному на большее.) Что такое мнимое и комплексное число? Вымысел ли это, или их прообразы существуют объективно?

Еще больше парадоксов обнаруживалось там, где рассматривались бесконечные процессы.

По-прежнему не находило объяснения равенство $A+\alpha=A$, где α — бесконечно малое.

Парадоксы возникали в теории рядов. Например, полагали, что сумма ряда

$$\ldots+\frac{1}{z^3}+\frac{1}{z^2}+\frac{1}{z}+1+z+z^2+z^3+\ldots$$

равна нулю, так как

$$1+\frac{1}{z}+\frac{1}{z^2}+\ldots=\frac{z}{z-1},$$

а

$$z+z^2+z^3+\ldots=\frac{z}{1-z},$$

хотя, очевидно, области сходимости каждого из этих рядов разные, поэтому такое сложение неправомерно.

Совершенно непонятными были вопросы определения суммы ряда с бесконечным числом членов, например в ряде

$$+1-1+1-1+\ldots,$$

с одной стороны, сумма равна нулю:

$$(+1-1)+(+1-1)+\ldots=0,$$

с другой — единице:

$$+1-(1-1)-(1-1)-\ldots=1.$$

Такая масса парадоксов, неумение понять и объяснить их, отсутствие обоснований высшей математики серьезно беспокоили ведущих математиков. Они проводили многогранные исследования, пытались построить логически безупречные основания высшей математики. Но все поиски не достигали цели, так как опирались на так называемые «вечные истины», созданные в период становления и развития математики постоянных, где нет движения, нет диалектики. Отмечая обреченность всех этих попыток, Ф. Энгельс писал: «Нет ничего комичнее, чем жалкие уловки, увертки и вынужденные приемы, к которым прибегают математики, чтобы разрешить это противоречие, примирить между собой высшую и низшую математику, уяснить себе, что то, что у них получилось в виде неоспоримого результата, не представляет собой чистой бессмыслицы, — и вообще рационально объяснить исходный пункт, метод и результаты математики бесконечного» (К. Маркс и Ф. Энгельс. Соч.,изд. 2-е, т. 20, стр. 519).

Так продолжалось до второй половины XIX в., до того времени, когда ученые поняли и оценили великие открытия Н. И. Лобачевского.

В XIX в. продолжается интенсивное развитие математики, причем области ее применения в естествознании далеко выходят за традиционные рамки астрономии, физики и механики. Это расширение заставило передовых ученых взглянуть на математический анализ как на науку, применимую во многих областях естествознания, т.е. как на науку, независимую от представлений геометрии и механики. Начался так называемый процесс арифметизации анализа. Суть его в том, что доказательства в различных разделах математического анализа должны опираться на понятия, принадлежащие этим разделам, на учение о числе и алгебру. Следовательно, в иерархическом построении теория чисел и алгебра должны предшествовать математическому анализу и быть с логической точки зрения построены безупречно.

В первой половине XIX в. большой вклад в дело арифметизации анализа внесли Больцано, Коши, Фурье, Лобачевский и ряд других математиков. В результате был подведен научно обоснованный фундамент под математический анализ Ньютона — Лейбница.

Этот факт дал некоторым математикам повод утверждать, что в первой половине XIX в. математический анализ получил окончательное и совершенно безупречное логическое обоснование. Но оказалось, что это не так. Коши в своих исследованиях рассматривал бесконечность как переменную, принимающую как угодно большое без всяких ограничений значение, или как малую величину, неограниченно приближающуюся к нулю. В современной терминологии такое толкование бесконечности называется потенциальной бесконечностью. Но еще в те времена было замечено, что для доказательства некоторых теорем (например, в интегральном исчислении) понятия потенциальной бесконечности недостаточно.

Понадобилась иная точка зрения: представление бесконечности как множества, заданного всеми своими элементами. Зарождалось новое понятие: так называемая актуальная бесконечность. Потребовалось еще немало усилий, чтобы подвести логически строгий фундамент под математический анализ.

В XIX в. продолжалось интенсивное и успешное развитие геометрии. Не прекращались многовековые попытки доказать пятый постулат Евклида. В 40-х годах увидели свет великие открытия Н. И. Лобачевского. Изучая и по достоинству оценивая труды предшественников, Лобачевский попытался встать на противоположную относительно традиционной точку зрения: исследовать, что получится, если в системе аксиом Евклида пятый постулат, который невозможно доказать, заменить противоположным высказываением, т. е. допустить, что через точку вне прямой можно провести в их плоскости не одну, а множество прямых, параллельных данной. Результат этого допущения был поразителен! Оказывается, из системы аксиом, в которой пятый постулат заменен противоположным высказыванием, возникает новая непротиворечивая логически строгая геометрическая система.

Историческое значение открытий Лобачевского состояло не столько в создании им неевклидовой геометрии, сколько в совершенном им перевороте во взглядах на способы мышления, системы доказательств, построение логических систем. И в самом деле, Лобачевский показал, что если одну из аксиом некоторой системы аксиом заменить противоположной по смыслу, то получается новая, неизвестная науке, но логически непротиворечивая система.

Чем же в таком случае является система аксиом? Как их строить? Что такое логический вывод? Что такое строгое математическое доказательство? Такие вопросы требовали ответа. Поиски этих ответов создавали благодатнейшую почву для новых открытий, исследований, достижений. Открытия Лобачевского явились поворотным пунктом в способах мышления математиков (и вообще в естествознании) XIX в.

Появляются работы Римана и других геометров, значительно развивших неевклидову геометрию, внесших большой вклад в разработку новых методов в дифференциальной, проективной геометрии, создавших геометрию n-мерного пространства.

В математическом анализе выясняется, что для обоснования арифметики действительных чисел, для доказательства фундаментальных теорем анализа необходимо понятие бесконечного множества. В 70-х годах возникает и быстро проникает почти во все математические науки теория множеств.

В алгебре фундаментальные результаты связаны с именем Галуа. Исследуя проблемы о разрешимости в радикалах алгебраических уравнений степени выше четвертой, Галуа построил специальный аппарат, названный впоследствии теорией групп. Этот аппарат оказался мощным инструментом исследования не только в алгебре, но и в ряде других наук. И более того, оказалось, что понятие группы наиболее ярко проявляет свои возможности, когда теория метода описывается независимо от природы объектов, к которым он будет применен. Иначе говоря, одна и та же математическая теория давала возможность одинаково успешно решать задачи, когда объектами исследования были понятия из различных областей: из алгебры и анализа, из геометрии, или элементы нематематической природы. Вообще аппарат математики во второй половине XIX в. значительно обогатился новыми теориями, понятиями, методами решения практических задач.

Совершенствование аксиоматического и теоретико-множественного методов, проникновение их в различные разделы математики плодотворно сказывались на их развитии и в то же время нагромождали новые проблемы и затруднения. Возникшие трудности были связаны, с одной стороны, с обилием новых теорий и методов, с исследованием их взаимозависимости и взаимного проникновения методов, с другой стороны, с изучением внутренней структуры каждой новой и «обновленной» теории.

Своеобразной чертой развития математики этого периода явилось то, что многие из возникших и нерешенных вопросов носили логический характер. Это вопросы о выводимости или невыводимости тех или иных положений из данных посылок, о разрешимости задач данными средствами, о логических связях между отдельными понятиями и целыми теориями.

Несмотря на особенности в каждом частном приложении, весь круг новых затруднений сводился к рассмотрению трех проблем:

1. Исследование возможности применения законов формальной логики к бесконечным множествам (в частности, закона исключенного третьего).

2. Обнаружение и объяснение парадоксов теории множеств.

3. Изучение структур и сущности математических доказательств.

Отсутствие достаточной ясности в этих вопросах стало серьезным тормозом развития математики. Если в XVIII в. можно было уйти от логически строгих обоснований дифференциального и интегрального исчисления, довольствуясь подтверждениями теоретических результатов на практике, то на рубеже XX в. аналогичного явления быть не могло. Человечество стояло на пороге великих открытий: строения материи, радио, воздухоплавания, космонавтики.

Практика ставила перед математикой такие задачи, правильность решения которых непосредственным опытом проверить было нельзя. В силу этого в исследовании проблем естествознания было необходимо опираться на достоверность выводов, полученных путем математических исследований.

Насколько серьезны были трудности видно из следующего. Закон исключенного третьего читается так: «Из двух противоположных суждений *A* и не *A* по крайней мере одно истинное».

В математике на основе этого закона построены доказательства методом «от противного»: если трудно непосредственно доказать истинность утверждения, то доказывают ложность противоположного и тогда, опираясь на логический закон исключенного третьего, утверждают истинность исходного предложения.

Если этот закон для бесконечных множеств не имеет места, то подвергаются сомнению все теоремы, доказанные методом «от противного».

Еще большее беспокойство вызывали парадоксы теории множеств. Методы теории множеств в совокупности с аксиоматическим методом проникли во многие разделы математики. С их помощью были получены замечательные результаты, позволившие свести доказательства истинности оснований математических теорий к доказательству истинности законов натурального ряда чисел.

Казалось, нужно было сделать так мало: исследовать давно знакомый и такой простой натуральный ряд. Но ведь этот ряд бесконечен, т.е. вопрос по-прежнему уперся в исследование бесконечности. Теория множеств была весьма эффективным инструментом исследования вообще и в том числе бесконечных процессов, но сама она, к сожалению, была не свободна от логических недостатков, и в ней появлялись пародоксы. Опять возникали вопросы: каким доказательствам можно доверять и каким нет? в каких логических структурах могут возникать ложные положения? какова сущность математических доказательств?

Возникновение вопроса о сущности математических доказательств вызвано традиционным убеждением в математике, что если отправные посылки верны и если доказательства вести в соответствии с законами формальной логики, то полученный результат будет также верен. Но оказалось, что не всегда можно применять законы формальной логики. И с другой стороны, если при доказательствах используются методы теории множеств, то результат может оказаться неверным вследствие наличия неразгаданных парадоксов. В связи с этим встал вопрос: какова должна быть структура математического доказательства, как его вести, какими средствами можно пользоваться?

Таким образом, сомнению подвергались не только основания, но и методы доказательств и полученные результаты. Нашлись даже математики-идеалисты, которые, пользуясь действительными трудностями в основаниях математики, заявили о неправомерности теории множеств, математического анализа и даже элементарной геометрии Евклида. Но большинство математиков не восприняли этих идеалистических воплей и продолжали исследования в разрешении возникших проблем.

Казалось, наиболее коротким путем разрешения всех противоречий был путь «исправления» теории множеств так, чтобы она была непротиворечивой и достаточной для обоснования всей математики. Простая на первый взгляд задача фактически оказалась чрезвычайно трудной. Дело в том, что ни в математике, ни в формальной логике не было методов, инструмента познания логических структур самой математики, закономерностей ее построений, выводов, следствий.

На этой почве в конце XIX и в начале XX вв. получила мощные стимулы развития новая ветвь математики — математическая логика. Ею начала заниматься одновременно и независимо друг от друга большая группа ведущих математиков. Этому посвящены знаменитые «Principia Mathematica» Рассела и Уайтхеда. Этим же занимались Гильберт, Бернайс и Аккерман. Наши соотечественники во главе с П. С. Новиковым получили замечательные результаты в доказательствах непротиворечивости арифметики. Изучением сложных логических структур занимались Гёдель, Куайн и многие другие ученые. В результате математическая логика приобрела определенную очерченность и самостоятельность. По предмету своему она яляется логикой, так как изучает законы мышления, а по методу — математикой, так как является дедуктивной теорией. Теория, предметом изучения которой является какая-то другая теория, называется метатеорией. Таким образом математическая логика является метатеорией математики или короче — метаматематикой. Существенный вклад в становление и развитие метаматематики внес Гильберт. Для изучения проблем, возникающих в основаниях геометрии, Гильберт разработал специальный метод исследования логических умозаключений. Для этого он использовал символический аппарат логических исчислений, применяемый еще в середине XIX в., и построил метаматематику как дедуктивную теорию, покоящуюся на своей системе аксиом. Пользуясь ею, он получил замечательные результаты.

Некоторым даже казалось, что наконец-то преодолены все загадки бесконечного и восстановлено доброе имя всех математических теорий и их оснований. Как показало дальнейшее развитие, такие выводы были неверны.

Историки считают, что после выхода в свет трудов Гильберта математическая логика начинает свой второй период развития. Как следует из законов диалектики, после преодоления одних противоречий возникают другие на более высоком уровне. Такое явление произошло в развитии математической логики.

В 1931 г. Гёдель опубликовал работу «О формально неразрешимых предложениях «Principia Mathematica» и родственных систем». В ней доказывается названная по имени автора теорема Гёделя о неполноте, говорящая о том, что в любой дедуктивной теории, покоящейся на некоторой системе аксиом, обязательно найдется предложение, доказать или опровергнуть истинность которого средствами данной теории нельзя. Для доказательства такого предложения необходима новая, более общая теория. В ней снова будет недоказуемое предложение и т.д.

Нетрудно заметить, что теорема Гёделя есть математический аналог законов диалектического развития, сформулированных классиками марксизма-ленинизма.

После этих сенсационных открытий развитие математической логики идет еще быстрее. В период с 30-х годов до настоящего времени особый вклад в эту науку внесли советские математики. Они явились создателями нового направления математической логики, получившего наименование «конструктивизм», в котором при логических построениях не используется закон исключенного третьего. Эти идеи оказались чрезвычайно плодотворными, нашли приложение во многих разделах математики и оказали существенное влияние на ее развитие, особенно на разделы, связанные с вычислительной техникой. А. А. Марков, П. С. Новиков и их последователи создают теорию алгоритмов, с помощью которой по-новому ставятся и раскрываются вопросы оснований математики, и которая сыграла и продолжает до сих пор играть важнейшую роль не только внутри математики, но и в большом круге прикладных проблем, и особенно в проникновении математических методов в различные нематематические науки: биологию, лингвистику, экономику. В 30-х годах советский ученый В. И. Шестаков (и независимо от него Шеннон в США) указал, что вычислительный аппарат математической логики, так называемая алгебра высказываний (булева алгебра), может быть интерпретирован на законах электрических цепей. Результатом этого открытия является выход математической логики за рамки обслуживания внутренних потребностей математики на широкий простор технических приложений.

В настоящее время идеи и методы математической логики имеют широчайшее приложение в кибернетике, а вычислительный аппарат — алгебра высказываний — стал «рабочим инструментом» в большом ряде технических наук: в теории конечных автоматов, в теории автоматического регулирования, синтеза релейно-контактных и бесконтактных схем, в вычислительной технике и программировании и т. д.

Настоящее пособие имеет относительно узкую цель: познакомить читателя с алгеброй высказываний и показать, каким образом она используется при решении некоторых технических вопросов. Аксиоматический метод и вопросы теории предикатов даются в виде ознакомительного очерка.

§ 1. 2. АКСИОМАТИЧЕСКИЙ МЕТОД

В рамках настоящего пособия ограничимся только лишь ознакомлением с идеями аксиоматического метода в математике и приложением его к математической логике. Сущность аксиоматического метода состоит в формальном способе определения совокупности математических объектов и отношения между ними и формальном способе получения новых истинных утверждений относительно объектов выделенной совокупности.

Пусть имеются какие-то объекты, свойства и закономерности отношений между которыми нам надо изучить. Обозначим их символами, например буквами, и не будем приписывать им ни содержания, ни определенного смысла. Это просто символы! Далее, укажем какие-то отношения, в которых объекты могут выступать, и также относительно их не будем утверждать, что они заведомо истинные. Это просто какая-то начальная позиция, которая может меняться, если будут указаны правила преобразований. Такие отношения, изображенные средствами обычной разговорной речи или с помощью известных значков (больше, меньше, равно и т. д.), называют *аксиомой*, а совокупность аксиом — *системой аксиом*.

Таким образом, в первоначальном виде система аксиом — это совокупность символов без конкретного содержания и отношений, в которых могут выступать эти символы. Если можно указать реально существующие объекты (например, параллельно-последовательные цепи с контактами) или абстрактные понятия математики (точки, прямые, числа), для которых выполняются отношения, сформулированные в аксиомах, то вся система аксиом приобретает содержательный смысл, и об отношениях между объектами можно говорить, истинны они или ложны.

Совокупность объектов, которые дают аксиомам реальное истолкование, называют *интерпретацией данной системы аксиом.*

Рассмотрим пример [18]. Пусть дана совокупность каких-то объектов и между ними установлено отношение, выраженное словом «предшествует». Не определяя ни объекты, ни смысл отношения «предшествует», выскажем два утверждения:

1) никакой объект не предшествует сам себе;

2) если x предшествует y, а y предшествует z, то x предшествует z.

Нетрудно видеть, что существуют системы объектов с таким отношением между ними, что если под термином «x предшествует y» понимать данное отношение, то утверждения 1) и 2) окажутся истинными.

Пусть, например, x, y, z — люди, а отношение «предшествует» принимает смысл слова «старше». В таком случае, если «x предшествует y» означает «x старше y», то утверждения 1) и 2) истинны.

Если под объектами совокупности понимать действительные числа, а под отношением «предшествует» понимать «x меньше y», тогда утверждения 1) и 2) также будут выполняться. Они определяют упорядоченную последовательность действительных чисел.

Из построения вытекает, что принадлежать совокупности могут только такие объекты, которые обладают свойствами, описанными в утверждениях 1) и 2). Тогда сами утверждения 1) и 2) являются определениями этой совокупности.

Утверждения 1) и 2), с помощью которых выделяются классы объектов, играют здесь роль аксиом, а совокупность объектов, удовлетворяющих этим аксиомам (в нашем примере люди и действительные числа), является интерпретацией данной системы аксиом.

Аксиоматический метод характеризуется также четко сформулированными так называемыми *правилами вывода:* должен быть указан способ действий, который гарантировал бы получение истинных следствий из заведомо истинных исходных полжений. Иначе говоря, если аксиомы истинны, то с помощью правил вывода можно строить новые предложения — теоремы, которые будут так же истинны.

Система аксиом должна обладать непременным свойством — быть *непротиворечивой (совместной).* Это означает, что не может быть явления, когда для объектов, определяемых данной системой аксиом, с помощью правил вывода были бы получены противоположные результаты, т. е. не может быть одинаково успешно доказана истинность высказывания A и его отрицания не A.

Если окажется, что такие доказательства можно провести, то система аксиом будет не совместна, т. е. она не будет отображать какой-либо класс объектов, не будет описывать чего-либо, превратится в пустой набор слов.

Система аксиом характеризуется еще свойством *независимости.* Оно не является категорически обязательным и состоит в том, что ни одна из аксиом системы не может быть выведена как следствие из остальных аксиом этой же системы. Доказать независимость аксиомы «a» от других аксиом непротиворечивой системы «A» — значит доказать непротиворечивость другой системы «B», в которой аксиома «a» заменена противоположным высказыванием «не a». Такая замена приведет к тому, что система аксиом будет определять собой объекты с другими качествами и иметь иную интерпретацию. Известно, что попытка доказать пятый постулат с помощью аксиом геометрии Евклида не увенчалась успехом. Допущение противоположного предложения создало новую непротиворечивую систему — неевклидову геометрию.

Отметим еще одно свойство — так называемую *равносильность* систем аксиом. Суть его состоит в том, что, оказывается, одну из аксиом некоторой непротиворечивой системы можно заменить предложением, которое в этой системе было получено как следствие (именовалось теоремой). В результате замены возникает новая система аксиом, в которой изъятая аксиома может быть доказана как теорема. Например, если в системе аксиом геометрии Евклида аксиому параллельности заменить предложением «Сумма углов треугольника равна 180°», то аксиома параллельности будет доказана как теорема.

Систему аксиом характеризуют еще свойством полноты. Система аксиом является *полной*, если любые свойства объектов каждой из ее интерпретаций могут быть доказаны при помощи аксиом этой системы.

Если свойство непротиворечивости гарантирует невозможность получения взаимоисключающих выводов A и не A, то свойство полноты характеризует обязательное получение одного из них. Математические исследования показывают, что достигнуть полноты системы аксиом невозможно.

В соответствии с общими принципами аксиоматического метода построена аксиоматика математической логики.

В математической логике предметом изучения являются не абстракции образов реального мира, а категории мышления, способы умозаключений, закономерности логического вывода.

В соответствии с этим объектами изучения являются отдельные высказывания, связи между ними, способы определения истинности различных сложных конструкций, составленных из отдельных высказываний. Под *высказыванием* в математической логике понимают законченную мысль, выраженную словами, о которой можно однозначно сказать, истинна она или ложна. Например, высказываниями будут: «Семь — простое число» — истинное высказывание; «Перпендикуляр длиннее наклонной» — ложное; «Москва — столица СССР» — истинное; «Снег — черный» — ложное.

Двумя последними примерами мы специально подчеркнули, что объектами в математической логике могут быть любые высказывания, несущие информацию из любых областей знаний, а не только из математики. Основное требование к высказыванию — это однозначная определенность ответа о значении истинности: истинно оно или ложно. Понятие «высказывание» является первичным, поэтому точного определения ему не дается. Но тем не менее из речевого общения очевидно, что такие фразы, как «Да здравствует дружба между народами», «Поедешь ли ты в командировку» и вообще вопросительные и восклицательные предложения, не являются высказываниями в смысле математической логики. Очевидно, не следует относить к высказываниям фразы, выражающие субъективные чувства, например, «Люблю кататься на лыжах», «Этот кинофильм интересный» и т. д.

Элементарные высказывания обозначают большими буквами латинского алфавита — $A, B, C, \dots$

В обычной разговорной речи из отдельных выражений строятся сложные предложения с помощью союзов «и», «или», отрицания «не», отношения следования «если..., то...» и т. д.

Эти связи получили в математической логике строгие определения и символические обозначения и являются в ней теми логическими связями, с помощью которых из элементарных высказываний строятся сложные и исследуются значения истинности сложных высказываний.

Союз **и** называют *конъюнкцией* и обозначают символом &. Союз **или** называют *дизъюнкцией* и обозначают символом $\vee$. Отрицание высказывания **не** обозначают черточкой над символом высказывания. Отношение следования **если... , то...** называют *импликацией* и обозначают стрелкой между символами $A \to B$.

Ниже приведена одна из возможных систем аксиом исчисления высказываний — первого и основного раздела математической логики. Она состоит из одиннадцати аксиом, сведенных в четыре группы.

Первая группа содержит две аксиомы и выражает действие импликации, остальные — содержат по три аксиомы и выражают содержание соответственно действия конъюнкции, дизъюнкции и отрицания. Вот эти аксиомы:

Первая группа

1. $A \to (B \to A)$
2. $(A \to (B \to C)) \to ((A \to B) \to (A \to C))$

Вторая группа

1. $A \& B \to A$
2. $A \& B \to B$
3. $(A \to B) \to ((A \to C) \to (A \to B \& C))$

Третья группа

1. $A \to A \vee B$
2. $B \to A \vee B$
3. $(A \to C) \to ((B \to C) \to (A \vee B \to C))$

Четвертая группа

1. $(A \to B) \to (\overline{B} \to \overline{A})$
2. $A \to \overline{\overline{A}}$
3. $\overline{\overline{A}} \to A$

Логические преобразования производятся формальным путем по правилам вывода: правилу подстановки и правилу заключений.

Правило подстановки. *Пусть W — истинная формула, содержащая букву A. Заменив всюду в формуле W букву A другой формулой U, получим новую, также истинную, формулу.*

Практический смысл этого правила состоит в том, что если в формуле W мы нашли какую-то часть, которая может быть заменена другим выражением, то мы можем сделать такую замену, сохраняя при этом истинное значение исходной формулы.

Правило заключений. *Если W и $W \to U$ — истинные формулы, то и U — истинная формула.*

Здесь сформулирована основная идея логического умозаключения. Смысл ее состоит в том, что, имея одну истинную формулу W (например, аксиому), мы логически доказываем истинность формулы $W \to U$. Отсюда как следствие вытекает истинность новой формулы U. Принято называть $W \to U$ *большой посылкой*, W *малой посылкой* и U *следствием*.

Для иллюстрации рассмотрим пример доказательства переместительного закона математической логики.

Пример 1. Пусть есть два высказывания A и B. Построим из них сложное высказывание $A \& B$ и покажем, что если $A \& B$ истинно, то, переставив местами A и B, получим также истинное высказывание. Сложное высказывание, которое мы хотим доказать в символах, изобразится так:

$$(A \& B) \to (B \& A).$$

Используя аксиому 3 второй группы и заменив в ней A на $A\&B$, получим в соответствии с правилом подстановки заведомо истинную формулу

$$[(A\&B)\to B]\to[((A\&B)\to C)\to[(A\&B)\to(B\&C)]]. \qquad (1.1)$$

Малая посылка этой формулы $(A\&B)\to B$ полностью совпадает с аксиомой 2 второй группы, потому истинна. Формула (1.1) истинна по построению. Используя правило заключений, получим

$$[(A\&B)\to C]\to[(A\&B)\to(B\&C)]. \qquad (1.2)$$

Формула (1.2) истинна. Заменив в ней всюду C на A, получим истинную формулу

$$[(A\&B)\to A]\to[(A\&B)\to(B\&A)]. \qquad (1.3)$$

В ней малая посылка полностью совпадает с аксиомой 1 второй группы.

Применяя правило заключений к выражению (1.3), получим истинную формулу $(A\&B)\to(B\&A)$.

Доказать это мы и ставили целью.

Формальный аппарат преобразований математической логики, получив на основе системы аксиом и правил вывода значительное развитие, в настоящее время представляет собой особую алгебру. Она имеет много сходного с обычной алгеброй, но в то же время принципиально отлична от нее. Ее называют *алгеброй высказываний* или *булевой алгеброй* (по имени одного из основателей этого вычислительного аппарата).

В общем построении математической логики алгебра высказываний выполняет служебную роль. Ею пользуются во всех разделах математической логики: в исчислении высказываний, в исчислении предикатов, в теории доказательств и др.

В технических приложениях допустим другой метод доказательства истинности логических формул — с помощью так называемых таблиц истинности. Он не обладает общностью, но в технических вопросах удобен. Дальше мы будем пользоваться только им.

§ 1.3. ПОНЯТИЕ ДВОИЧНОЙ ФУНКЦИИ

В предыдущем параграфе мы познакомились с одним из основных понятий математической логики — элементарным высказыванием.

С помощью элементарных высказываний могут быть выражены образы и понятия различных наук и вообще это может быть любая законченная мысль, выраженная словами.

К элементарному высказыванию предъявляется единственное требование: оно должно иметь вполне определённое значение истинности.

Простой на первый взгляд вопрос однозначной определенности элементарного высказывания при глубоком логическом анализе оказался достаточно сложным, подчиненным определенной закономерности. Ее подметил еще Аристотель и сформулировал в законах классической логики.

Для нас практический интерес представляет один из них: закон исключенного третьего.

Он гласит: «Между утверждением чего-либо и отрицанием того же самого нет ничего третьего или среднего; одно из них (утверждение или отрицание) истинно, другое ложно».

Практический смысл этого закона состоит в том, что он создает возможность разделять различные высказывания на две группы: удовлетворяющие закону исключенного третьего и не удовлетворяющие ему.

С высказываниями первой группы в логических построениях используются одни, соответствующие им, приемы умозаключений, с высказываниями второй группы — другие. Например, прямая a и плоскость ρ в евклидовом пространстве могут быть только параллельны или непараллельны. Вследствие этого и высказывание «Прямая a пересекает плоскость ρ» истинно, если a непараллельна ρ, и ложно в противном случае. Однозначная определенность такого высказывания позволяет отнести его к высказываниям, удовлетворяющим закону исключенного третьего, и в соответствии с этим строить умозаключения при доказательствах теорем.

Другое высказывание, например «Поезд идет быстро», не имеет определенного значения истинности. В нем не ясно, что значит «быстро», с чем сравнивать эту скорость. Такая оценка субъективна и строгих умозаключений с ее помощью получить нельзя.

Закон исключенного третьего значительно сокращает количество высказываний, выделяя из них те, для которых существует однозначный ответ об их истинности. Вместе с тем сокращается область, в которой строятся строгие умозаключения. Для расширения ее на практике при построении элементарных высказываний для того, чтобы о них можно было однозначно сказать истинны они или ложны, из совокупности свойств и характеристик объекта выделяют только необходимые. Уточняют и конкретизируют их и только относительно этих выделенных характеристик строят высказывания. Например, при исследовании электрических цепей, можно однозначно говорить об истинности высказываний «Цепь замкнута», «Цепь разомкнута» только тогда, когда рассматриваются установившиеся состояния, а переходные процессы, происходящие при замыкании и размыкании цепи, не учитываются.

Элементарные высказывания могут быть представлены в различных формах: обычная разговорная речь, символы и термины технических наук, математические формулы. Важно лишь одно, чтобы смысл высказывания был ясен и о нем можно было бы однозначно сказать истинно оно или ложно.

Для формального изображения двух значений истинности элементарных высказываний используют несколько пар обозначений (в зависимости от приложений): истинно, ложно или короче И, Л; +, —; 1, 0. В блок-схемах алгоритмов используют «да», «нет».

В практике часто элементарные высказывания представляются в функциональной форме. Смысл такой записи рассмотрим на примере. Пусть высказывание A есть «5 — простое число». Это высказывание удовлетворяет закону исключенного третьего и истинно. Высказывание B — «6 — простое число» — ложно. О значении истинности высказывания C — «x — простое число» — ничего определенного сказать нельзя. Это так называемая *пропозиционная форма*, которая принимает определенные значения истинности только тогда, когда вместо x будет подставлен конкретный объект. Ясно, что постановка вопроса об истинности пропозиционной формы не имеет смысла.

В математической логике в разделе исчисления высказываний изучаются связи и отношения между элементарными высказываниями, представленными своими пропозиционными формами.

Сложные высказывания строятся путем объединения элементарных с помощью логических связей **и**, **или**, **если** ..., **то** ..., отрицания **не**. Например, из высказываний: A — «Фигура x — плоская», B — «Фигура x — прямолинейная», C — «Фигура x — равносторонняя», D — «Фигура x — симметричная» можно построить сложное высказывание

$$W = A \& B \& \bar{C} \& D \vee A \& D,$$

которое читается так: «Фигура x — плоская **и** прямолинейная **и** неравносторонняя **и** симметричная **или** плоская **и** симметричная». Анализируя такое высказывание с содержательной точки зрения, замечаем, что оно в целом может быть либо истинным (например, для трапеции), либо ложным (например, для замкнутой пространственной кривой). Значение истинности сложного высказывания будет зависеть от значений истинности составляющих элементарных высказываний и от вида логической связи между ними.

Все такие конструкции изучаются в исчислении высказываний. Сложное высказывание является функцией элементарных высказываний, что записывается так:

$$U = F(A, B, \ldots, N),$$

где $A, B, \ldots, N$ рассматриваются как аргументы.

В дальнейшем такие функции будем называть *логическими*.

Существенным отличием логических функций от функций, изучаемых в математике, является то, что и аргументы логических функций и сами логические функции могут принимать только два значения: истинно и ложно.

Логические функции $U = F(A, B, \dots, N)$ принято называть *двоичными функциями*.

Рассмотрим область существования и область изменения такой функции.

Согласно определению двоичная функция $U = F(A, B, \dots, N)$ может принимать только два значения: истинно, ложно. Следовательно, область изменения ее ограничивается этими двумя значениями. Область существования — значительно шире. Рассматривая какую-то двоичную функцию, мы имеем дело с вполне определенной областью знаний. В силу этого в качестве аргументов — элементарных высказываний — выступают всевозможные высказывания из этой же области. Из данного набора таких высказываний $A, B, \dots, N$ в конкретную двоичную функцию могут входить в качестве аргументов только некоторые,

Рассмотрим область изменения аргументов в случае, когда количество и состав их известен, например: A, C, K, L.

Здесь каждый аргумент — двоичный. Обозначим истинное значение аргумента символом A^+, ложное — A^-. Если конкретное значение истинности неизвестно, но нужно показать, что рассматриваются не оба значения, а какое-то одно фиксированное, то записывается это символом A^*.

Каждый из аргументов в условиях конкретного примера принимает одно фиксированное значение A^*, B^*, K^*, L^*, и таким образом создаются определенные комбинации фиксированных значений. Такой комбинации соответствует единственное фиксированное значение двоичной функции. Всего из четырех элементарных высказываний можно составить 16 различных комбинаций фиксированных значений аргументов. Из n аргументов можно составить 2^n различных комбинаций (это в дальнейшем будет доказано). Каждой фиксированной комбинации $A^*, B^*, \dots, N^*$ соответствует единственное значение истинности двоичной функции $U(A, B, \dots, N)$.

Таким образом, областью определения двоичной функции $U(A, B, \dots, N)$ являются 2^n комбинаций фиксированных значений аргументов: $(A^*, B^*, \dots, N^*)_j$, $j = 1,2, \dots, 2^n$.

Двоичная функция $U(A, B, \dots, N)$ устанавливает взаимно-однозначное соответствие между каждой комбинацией фиксированных значений аргументов $A^*, B^*, \dots, N^*$ и единственным значением истинности самой функции:

$$U(A,B,\ \ldots,\ N)=\begin{cases}\text{И, если функция } U(A^*, B^*,\ \ldots,\ N^*)\\ \text{истинна;}\\ \text{Л, если функция } U(A^*, B^*,\ \ldots,\ N^*)\\ \text{ложна.}\end{cases} \qquad (1.4)$$

В технических приложениях для обозначения элементарных высказываний принят символ x_i, а для обозначения «истинно» и «ложно» — 1 и 0 соответственно.

Тогда запись $F(x_1, x_2, \ldots, x_n)$ означает, что рассматривается двоичная функция, построенная на аргументах $x_1, x_2, \ldots, x_n$. Совокупность фиксированных значений истинности аргументов $\overset{*}{x}_1, \overset{*}{x}_2, \ldots, \overset{*}{x}_n$ в принятых обозначениях представится в виде последовательности нулей и единиц, например 11000010110. Такие последовательности называют *наборами* и говорят: функция $F(x_1, x_2, \ldots, x_n)$ рассматривается на наборе $(\overset{*}{x}_1, \overset{*}{x}_2, \ldots, \overset{*}{x}_n)_j$, где $\overset{*}{x}_i$ есть нуль или единица, а индекс j указывает, что из всевозможных наборов рассматривается именно j-й.

Формула (1.4) в принятых обозначениях будет иметь вид

$$F(x_1, x_2, \ldots, x_n)=\begin{cases}1, \text{ если функция } F(\overset{*}{x}_1, \overset{*}{x}_2, \ldots, \overset{*}{x}_n)\\ \text{истинна;}\\ 0, \text{ если } F(\overset{*}{x}_1, \overset{*}{x}_2, \ldots, \overset{*}{x}_n) \text{ ложна.}\end{cases}$$

Важно отметить, что набор $\overset{*}{x}_1, \overset{*}{x}_2, \ldots, \overset{*}{x}_n$, представленный нулями и единицами, может быть рассмотрен и как собственно набор фиксированных значений аргументов $x_1, x_2, \ldots, x_n$ и как число в двоичной системе счисления.

§ 1. 4. АЛГЕБРА ВЫСКАЗЫВАНИЙ

Существует несколько способов задания двоичных функций. Рассмотрим два из них: аналитический и табличный.

Функция, заданная *аналитическим способом*, представляется формулой в исчислении высказываний. Такие формулы состоят из символов элементарных высказываний $A, B, \ldots, N$, объединённых символами логических связей $\&, \vee, \overline{}, \rightarrow$. Для наглядного представления приведем одну из формул, построенную на элементарных высказываниях A, B, C:

$$U(A,\ B,\ C)=[(\bar{A}\rightarrow B)\ \&\ (\overline{\bar{A}\vee B\vee C})]\rightarrow[(\bar{B}\vee C)\rightarrow(A\vee\bar{C})].$$

Более употребительным, особенно в технических приложениях, является *табличный способ* задания двоичных функций. В этом случае каждой комбинации фиксированных значений истинности аргументов $A_1, A_2, \ldots, A_n$ ставится в соответствие значение истинности самой двоичной функции $F(A_1, A_2, \ldots, A_n)$. Осуществляется это следующим образом.

Задается двоичная функция $F(A_1, A_2, \ldots, A_n)$, зависящая от n аргументов. Тогда таблица задания содержит $n+1$ столбец: n для аргументов и один для функции. В каждой строчке таблицы записывается одна из комбинаций фиксированных значений аргументов, т. е. набор.

Так как наборы можно рассматривать как числа в двоичной системе счисления, то их удобно располагать в порядке возрастания двоичных чисел от 0 до 2^n.

Внешний вид таблицы задания для $n = 4$:

A	B	C	D	$F(ABCD)$
0	0	0	0	1
0	0	0	1	0
0	0	1	0	0
0	0	1	1	1
0	1	0	0	0
0	1	0	1	1
0	1	1	0	1
0	1	1	1	0
1	0	0	0	0
1	0	0	1	1
1	0	1	0	1
1	0	1	1	0
1	1	0	0	1
1	1	0	1	0
1	1	1	0	0
1	1	1	1	1

Читая по строкам, имеем, например: в пятой строке, на наборе 0100, функция $F(ABCD)$ ложна.

В таблице должны быть представлены все наборы, которые могут быть построены из n двоичных переменных. Таких наборов может быть 2^n. Докажем это. Если $n = 1$, то число строк равно 2 для $X = 0$ и $X = 1$. Допустим, наше утверждение справедливо для $n = k$, т. е. для k двоичных переменных $X_1, X_2, \ldots, X_k$ число строк в таблице будет 2^k. Докажем, что при добавлении $(k+1)$-й переменной X_{k+1} число строк удвоится, т. е. будет $2\cdot 2^k = 2^{k+1}$.

Если в таблице для k переменных есть 2^k строк и добавили еще одну переменную, то для того, чтобы получить все комбинации значений переменных, нужно с каждой из 2^k строк рассмореть комбинацию для $X_{k+1}=0$ и для $X_{k+1}=1$. В каждом случае потребуется по 2^k строк, а всего будет $2^k+2^k = 2\cdot 2^k = 2^{k+1}$ строк, что и требовалось доказать.

Воспользуемся таблицами задания функций для определения значений истинности основных логических связей.

Отрицанием высказывания A называется новое высказывание (читается: не A), которое истинно, когда A ложно, и ложно, когда A истинно.

Значение истинности A определяется таблицей:

A	$\bar{A}$
0	1
1	0

Отрицающие друг друга высказывания A и $\bar{A}$ называют *противоположными*.

Конъюнкцией двух высказываний A **и** B называется новое высказывание $A\,\&\,B$ (читается: A и B), которое истинно тогда и только тогда, когда истинно каждое из высказываний A и B.

Значение истинности $A\,\&\,B$ определяется таблицей:

A	B	$A\,\&\,B$
0	0	0
0	1	0
1	0	0
1	1	1

Конъюнкцию высказываний называют также *логическим произведением* и часто обозначают AB (как произведение в математике).

Дизъюнкцией двух высказываний A и B называется новое высказывание $A \vee B$ (читается: A или B), которое истинно тогда, когда истинно хотя бы одно из составляющих высказываний.

Значение истинности $A \vee B$ определяется таблицей:

A	B	$A \vee B$
0	0	0
0	1	1
1	0	1
1	1	1

Дизъюнкцию высказываний называют также *логической суммой* и иногда обозначают $A+B$.

Из определения дизъюнкции видно, что логическая связка **или** имеет неразделительный смысл: истинно A **или** B **или** оба.

Импликацией высказываний A и B называется новое высказывание $A \rightarrow B$ (читается: если A, то B; из A следует B), которое ложно тогда и только тогда, когда A истинно, а B ложно.

Значение истинности $A \rightarrow B$ определяется таблицей:

A	B	$A \rightarrow B$
0	0	1
0	1	1
1	0	0
1	1	1

В импликации $A \rightarrow B$ высказывание A называется *основанием* или *посылкой*, а высказывание B — *следствием*.

Вместо термина «импликация» иногда встречается термин «следование».

Равнозначностью (эквивалентностью) высказываний A и B называется новое высказывание $A \sim B$ (читается: A равнозначно B; A эквивалентно B), которое истинно тогда и только тогда, когда A и B одновременно истинны или одновременно ложны.

Значение истинности $A \sim B$ определяется таблицей:

A	B	$A \sim B$
0	0	1
0	1	0
1	0	0
1	1	1

С помощью введенных операций из элементарных высказываний $A_1, A_2, \ldots, A_n$ строятся сложные высказывания, например:

$$A_1 \& A_2;\ A_1 \vee A_2;\ A_1 \to A_2;\ \bar{A};\ A_1 \sim A_2.$$

Такие выражения называют *логическими формулами.*

Каждая из них, являясь двоичной, может принимать определенное значение истинности и, будучи обозначена одной буквой, может быть рассмотрена как элементарное высказывание. Объединяя такие формулы логическими связями, получим более крупные формулы. Например,

$$\{A_1 \to [(A_1 \& A_2) \to (A_3 \sim A_4)]\} \& \overline{(A_1 \vee A_2)}.$$

Скобки в таких выражениях указывают область действия логического знака. В нашем примере знак $\to$ внутри квадратной скобки связывает $(A_1 \& A_2)$ и $(A_3 \sim A_4)$. Знак $\to$ после A_1 связывает A_1 и квадратную скобку, но не распространяется на $\overline{A_1 \vee A_2}$. Знак отрицания, если он стоит над сложным выражением, относится ко всему этому выражению.

Логические формулы допускают различные тождественные преобразования с целью их упрощения или придания более удобного для практики вида.

Некоторые из основных законов преобразований алгебры логики имеют аналоги в обычной алгебре. Но есть и специфичные, присущие только алгебре высказываний. Наиболее употребительные сведем в таблицу (табл. 1.1).

Распределительные законы обобщаются на случай n элементарных высказываний. Делается это следующим образом: несколько элементов формулы, которые имеют между собой какую-то тесную связь, объединяют и обозначают одной буквой. Полученную формулу преобразуют, используя распределительный закон.

Для иллюстрации преобразуем выражение $(A \vee B) \& (C \vee D)$.

Таблица 1.1

Наименование закона	Формула в булевой алгебре	Формула в обычной алгебре
Переместительный закон дизъюнкции	$A \vee B = B \vee A$	$a + b = b + a$
Сочетательный закон дизъюнкции	$A \vee (B \vee C) = (A \vee B) \vee C$	$a + (b + c) = (a + b) + c$
Переместительный закон конъюнкции	$A \& B = B \& A$	$ab = ba$
Сочетательный закон конъюнкции	$(A \& B) \& C = A \& (B \& C)$	$(ab)\, c = a\, (bc)$
Первый распределительный закон (конъюнкции относительно дизъюнкции)	$A \& (B \vee C) = (A \& B) \vee (A \& C)$	$a\,(b + c) = ab + ac$
Второй распределительный закон (дизъюнкции относительно конъюнкции)	$A \vee (B \& C) = (A \vee B) \& (A \vee C)$	В алгебре аналога нет. По форме этот закон имел бы вид $a + bc = (a + b)\,(a + c)$

Обозначим $(A \vee B) = U$. Тогда

$$U \& (C \vee D) = U \& C \vee U \& D.$$

После этого подставим значение U:

$$U \& C \vee U \& D = (A \vee B) \& C \vee (A \vee B) \& D = \\ = A \& C \vee B \& C \vee A \& D \vee B \& D.$$

Аналогом таких преобразований в обычной алгебре будет

$$(a + b)\,(c + d) = ac + bc + ad + bd.$$

Обратим внимание на следующее. В преобразовании вида

$$U \& (C \vee D) = U \& C \vee U \& D$$

в правой части скобки опущены. Вследствие этого может создаться впечатление неоднозначности результата. В самом деле, указывая последовательность действий с помощью скобок, можно получить правильное решение

$$U \& (C \vee D) = (U \& C) \vee (U \& D),$$

но можно записать и так:

$$U \& (C \vee D) = U \& (C \vee U) \& D.$$

В обычной алгебре правильным решением является

$$a(b+c)=ab+ac,$$

но записать можно и так:

$$a(b+c)=a(b+a)c.$$

Однако никому не придет в голову пользоваться таким равенством. Происходит это потому, что в алгебре принято прежде выполнять действие умножения (деления), а затем сложения (вычитания). То же самое явление имеет место в алгебре высказываний: принята договоренность, что более тесную связь дает конъюнкция и за ней — дизъюнкция. Потому выражение $U \& C \vee U \& D$, в котором отсутствуют скобки, может быть прочитано только так: $(U \& C) \vee (U \& D)$. Выражение $U \& (C \vee U) \& D$ имеет уже другой смысл.

Для остальных логических операций принята следующая договоренность: знак отрицания связывает теснее знаков &, $\vee$, $\rightarrow$, $\sim$; знаки $\vee$ и & — теснее знаков $\rightarrow$ и $\sim$.

Доказательства формул, сведенных в табл. 1.1, можно провести в рамках аксиоматического метода. Пример такого доказательства для переместительного закона конъюнкции был рассмотрен в § 1.2.

В технической литературе для доказательств обычно используются таблицы истинности.

Так будем поступать и мы.

Таблицы истинности строятся следующим образом: слева отводится n столбцов для записи фиксированных значений аргументов. Столбцы озаглавливаются символами этих аргументов.

Затем следует группа столбцов, озаглавленных отдельными частями доказываемого равенства. Последний столбец озаглавливается всем доказываемым равенством.

В столбцах записываются такие выражения, чтобы постепенно, дополняя предыдущий столбец одной логической связью, строить полностью прежде левую, затем правую часть доказываемого равенства. В таблице 2^n строк в соответствии с числом всевозможных наборов фиксированных значений аргументов.

Первые n столбцов заполняются наборами, выраженными символами И и Л или 1 и 0 точно так же, как было описано в таблицах задания двоичных функций.

В последующих столбцах значения истинности определяются следующим путем.

Столбец озаглавлен частью логической формулы, входящей в доказываемое равенство.

Каждая из логических связей, участвующая в формуле.

озаглавливающей данный столбец, имеет однозначное значение истинности, соответствующее фиксированному набору значений аргументов.

Рассматривая строку таблицы, берем из первых n столбцов конкретный набор и определяем значение истинности формулы на этом наборе. Полученное значение проставляем на пересечении строки и столбца. Такая операция проделывается по всем строкам и столбцам таблицы, кроме последнего.

Для простановки символов значений истинности в последнем столбце отыскиваем среди столбцов таблицы такие, которые озаглавлены формулами, представляющими правую и левую части равенства. По строчкам в этих столбцах сравниваем значения истинности. Если они совпадают (в обоих столбцах нули или в обоих столбцах единицы), то в последнем столбце проставляется символ «истинно». Если по всем 2^n столбцам будет стоять символ истинности, то считаем равенство доказанным.

Пример 1. В качестве примера с помощью таблицы истинности докажем справедливость второго распределительного закона

$$A \vee (A \& C) = (A \vee B) \& (A \vee C).$$

A	B	C	$B\&C$	$A\vee(B\&C)$	$A\vee B$	$A\vee C$	$(A\vee B)\&(A\vee C)$	$A\vee(B\&C) = (A\vee B)\&B(A\vee C)$
0	0	0	0	0	0	0	0	1
0	0	1	0	0	0	1	0	1
0	1	0	0	0	1	0	0	1
0	1	1	1	1	1	1	1	1
1	0	0	0	1	1	1	1	1
1	0	1	0	1	1	1	1	1
1	1	0	0	1	1	1	1	1
1	1	1	1	1	1	1	1	1

На практике удобно заполнять таблицу по столбцам. Для примера опишем рассуждения при заполнении пятого столбца. Он озаглавлен $A \vee (B\&C)$. В отличие от предыдущего столбца ($B\&C$) в пятом столбце добавлено действие дизъюнкции. Для определения значения истинности берем фиксированное значение аргументов из первого и четвертого столбцов (заметим, что можно брать весь набор по трем первым столбцам и подставлять в пятый). Значения истинности в пятом столбце проставлены в соответствии с определениями действий дизъюнкции и конъюнкции.

Для доказательства равенства сравниваем значения истинности в пятом и восьмом столбцах, и так как по всем строкам эти значения совпадают, то в девятом столбце имеем символ «истинно». Эти и доказывает справедливость второго распределительного закона.

В исчислении высказываний имеет место ряд формул, специальных для математической логики, не имеющих аналогов в обычной алгебре. Приведем наиболее употребительные из них:

$$\overline{\overline{A}} = A \tag{1.5}$$

(т. е. второе отрицание ликвидирует действие первого);

$$A \& A \& \ldots \& A = A; \tag{1.6}$$

$$A \vee A \vee \ldots \vee A = A. \tag{1.7}$$

В обычной алгебре аналогичные формулы имеют вид $a\, a \ldots a = a^n$ и $a + a + \ldots + a = na$. Такое отличие обусловлено тем, что в алгебре суммируются и перемножаются количества, выраженные символами a. В математической логике символ A обозначает высказывание и оцениваются его качественные характеристики. От повторения одного и того же высказывания несколько раз с союзом **и** или **или** смысл высказывания A не меняется.

Особое место в исчислении высказываний занимают так называемые *тождественно истинные* и *тождественно ложные формулы*. Смысл их в том, что они на любых наборах значений аргументов сохраняют одно и то же значение истинности. Например,

$$A \vee \overline{A} = \text{И} \tag{1.8}$$

всегда истинна,

$$A \& \overline{A} = \text{Л} \tag{1.9}$$

всегда ложна,

$$\overline{A \& \overline{A}} = \text{И}$$

всегда истинна.

Эти формулы широко используются при логических преобразованиях. Докажем их с помощью таблиц истинности:

A	$\overline{A}$	$A \vee \overline{A}$
1	0	1
0	1	1

A	$\overline{A}$	$A \& \overline{A}$
1	0	0
0	1	0

A	$\overline{A}$	$A \& \overline{A}$	$\overline{A \& \overline{A}}$
1	0	0	1
0	1	0	1

$$A \,\&\, \text{И} = A. \tag{1.10}$$

Справедливость формулы (1.10) вытекает из рассуждения: если в конъюнкции один элемент заведомо истинный, то истинность или ложность всей конъюнкции зависит от истинности или ложности второго элементарного высказывания. Из формулы (1.10) вытекает, что в конъюнкции тождественно истинный элемент можно отбросить.

При одном тождественно ложном элементе конъюнкция будет ложной:

$$A \,\&\, \text{Л} = \text{Л}. \tag{1.11}$$

В дизъюнкции тождественно ложный элемент можно отбросить:

$$A \vee \text{Л} = A. \tag{1.12}$$

Дизъюнкция всегда истинна, если один элемент тождественно истинный:

$$A \vee \text{И} = \text{И}. \tag{1.13}$$

Аналогичные свойства имеют место для импликации и эквивалентности.

Значение истинности импликации при тождественно истинной посылке совпадает со значением следствия:

$$\text{И} \rightarrow A = A. \tag{1.14}$$

При тождественно ложной посылке импликация истинна при любых значениях следствия:

$$\text{Л} \rightarrow A = \text{И}. \tag{1.15}$$

По определению импликация ложна в единственном случае: когда посылка истинна, а следствие ложно. Формула (1.15) при любых значениях A окажется истинной.

Укажем два соотношения эквивалентности:

$$A \sim \text{И} = A; \tag{1.16}$$

$$A \sim \text{Л} = \overline{A}, \tag{1.17}$$

справедливость которых проверим с помощью таблиц:

A	И	$A \sim$ И	$A \sim$ И $= A$
1	1	1	1
0	1	0	1

A	$\overline{A}$	Л	$A \sim$ Л	$A \sim$ Л $= \overline{A}$
1	0	0	0	1
0	1	0	1	1

В логических преобразованиях важную роль играют формулы, выражающие одну из логических связей через другую. Наиболее широко используются следующие две формулы:

$$\overline{A \& B} = \bar{A} \vee \bar{B}; \tag{1.18}$$

$$\overline{A \vee B} = \bar{A} \& \bar{B}. \tag{1.19}$$

Докажем формулу (1.18) с помощью таблицы истинности:

A	B	$\bar{A}$	$\bar{B}$	$A \& B$	$\overline{A \& B}$	$\bar{A} \vee \bar{B}$	$\overline{A \& B} = \bar{A} \vee \bar{B}$
1	1	0	0	1	0	0	1
1	0	0	1	0	1	1	1
0	1	1	0	0	1	1	1
0	0	1	1	0	1	1	1

Из таблицы видно, что значения истинности в шестом и седьмом столбцах совпадают, следовательно, формула в восьмом столбце всегда истинна, что и требовалось доказать. Аналогично доказывается и формула (1.19).

С помощью формул (1.18) и (1.19) конъюнкция может быть выражена через дизъюнкцию и наоборот. Формулы (1.18) и (1.19) называют *законом инверсии* или *формулами де Моргана*.

Докажем формулу

$$A \to B = \overline{A \& \bar{B}}, \tag{1.20}$$

с помощью которой импликация выражается через конъюнкцию.

Построим таблицу истинности:

A	B	$\bar{B}$	$A \to B$	$A \& \bar{B}$	$\overline{A \& \bar{B}}$	$A \to B = \overline{A \& \bar{B}}$
1	1	0	1	0	1	1
1	0	1	0	1	0	1
0	1	0	1	0	1	1
0	0	1	1	0	1	1

Сравнивая четвертый и шестой столбцы, замечаем совпадение их значений. Это и доказывает справедливость формулы (1.20).

Следующую группу формул получим путем преобразования уже известных. Заменив в формуле (1.18) B и $\overline{B}$, получим

$$\overline{A\&\overline{B}} = \overline{A} \vee \overline{\overline{B}}.$$

Учитывая, что $\overline{\overline{B}} = B$, и сравнивая с формулой (1.20), получим соотношение, связывающее импликацию с дизъюнкцией:

$$A \to B = \overline{A} \vee B. \tag{1.21}$$

Исследуем, как изменится импликация, если в ней поменять местами посылку и следствие. В формуле (1.20) подставим вместо A высказывание $\overline{B}$, а вместо B — высказывание $\overline{A}$, получим

$$\overline{B} \to \overline{A} = \overline{\overline{B}\&\overline{\overline{A}}} = \overline{\overline{B}\&A} = \overline{\overline{B}} \vee \overline{A} = \overline{A} \vee B$$

(здесь использованы формулы (1.8), (1.18) и переместительный закон дизъюнкции). Таким образом,

$$\overline{B} \to \overline{A} = \overline{A} \vee B. \tag{1.22}$$

Замечая, что правые части формул (1.21) и (1.22) одинаковы, получим еще одну формулу

$$A \to B = \overline{B} \to \overline{A}, \tag{1.23}$$

рассматривая которую, замечаем, что импликация переместительным законом не обладает.

Выразим соотношение эквивалентности через другие логические связи. Сравним таблицы истинности для импликаций, в которых посылка и следствие поменялись местами:

A	B	$A \to B$
1	1	1
1	0	0
0	1	1
0	0	1

A	B	$B \to A$
1	1	1
1	0	1
0	1	0
0	0	1

Рассматривая таблицы, замечаем, что если элементарные высказывания принимают одинаковые значения истинности, то обе импликации истинны; если в точности одно из элементарных высказываний ложно, то и соответствующая импликацияция ложна. Такое соотношение истинности реализует логическая формула $(A \to B) \& (B \to A)$, которая принимает значение «истинно», если оба элементарных высказывания принимают одинаковые значения истинности.

Такому же условию удовлетворяет определение эквивалентности $A \sim B$, следовательно,

$$A \sim B = (A \to B) \& (B \to A). \qquad (1.24)$$

Формула (1.24) дает возможность выразить эквивалентность через импликацию и конъюнкцию.

Из определения эквивалентности следуют две формулы:

$$A \sim B = B \sim A;$$

$$A \sim B = \overline{A} \sim \overline{B}.$$

Ниже будет указана определенная целенаправленность в преобразованиях логических выражений и сформулированы соответствующие правила, здесь же рассмотрим только пример.

Пример 2. Упростить логическую формулу

$$[(A \to B) \& (B \to C)] \to (A \to C).$$

Решение. Все импликации заменим по формуле (1.21). Очевидно, для знака импликации, стоящего между квадратной и круглой скобками, роль посылки будет играть вся квадратная скобка, а роль следствия — импликация $(A \to C)$. В преобразовании будем стремиться к тому, чтобы знак отрицания относился только к элементарным высказываниям, и повторяющиеся высказывания по возможности будем опускать.

$$[(A \to B) \& (B \to C)] \to (A \to C) = \overline{(\overline{A} \vee B) \& (\overline{B} \vee C)} \vee (\overline{A} \vee C) =$$
$$= \overline{\overline{A} \vee B} \vee \overline{\overline{B} \vee C} \vee (\overline{A} \vee C) = \overline{\overline{A}} \& \overline{B} \vee \overline{\overline{B}} \& \overline{C} \vee (\overline{A} \vee C) =$$
$$= A \& \overline{B} \vee B \& \overline{C} \vee (\overline{A} \vee C) = [A \& \overline{B} \vee \overline{A}] \vee (B \& \overline{C} \vee C) =$$
$$= [(A \vee \overline{A}) \& (\overline{B} \vee \overline{A})] \vee [(B \vee C) \& (\overline{C} \vee C)] =$$
$$= [\text{И} \& (\overline{B} \vee \overline{A})] \vee [(B \vee C) \& \text{И}] = (\overline{B} \vee \overline{A}) \vee (B \vee C) =$$
$$= (\overline{B} \vee B) \vee (\overline{A} \vee C) = \text{И} \vee \overline{A} \vee C = \text{И}.$$

Преобразования исходной формулы были сделаны в соответствии с формулами (1.21), (1.18), (1.19), (1.5), (1.8), (1.10), (1.13) (использованы сочетательные и распределительные законы).

Таким образом, оказалось, что сложная формула $[(A \to B) \& (B \to C)] \to (A \to C)$ всегда истинна.

В практике, выполняя сложные преобразования и обнаружив при этом члены, которые составляют известную тождественно истинную (ложную) формулу, объединяют их и заменяют соответствующим символом.

В большинстве случаев в результате преобразований формулы упрощаются, но остаются функциональными. Всякие преобразования имеют свою цель. В обычной алгебре такой целью может быть представление выражения в виде многочлена или одночлена. В алгебре высказываний наряду с уменьшением числа элементов, входящих в формулу, целью преобразования является представление ее с помощью ограниченного числа логических операций. В связи с этим вводится понятие системы базисных операций. Некоторую часть логических операций, через которые могут быть выражены все остальные, называют *системой базисных операций* (в технических приложениях — *системой базисных элементов*).

Формулы (1.18) — (1.24) указывают на возможность замены одних логических операций другими и выбора из них базисной системы. Оказывается, существуют логические операции, обладающие свойством универсальности, т. е. с помощью одной такой операций можно выразить все остальные. Таким свойством обладают операции *отрицание конъюнкции* и *отрицание дизъюнкции*, названные по имени ученых *штрихом Шеффера* и *стрелкой Пирса*.

Они задаются таблицами:

A	B	$A \& B$	$\overline{A \& B}$
0	0	0	1
0	1	0	1
1	0	0	1
1	1	1	0

A	B	$A \vee B$	$\overline{A \vee B}$
0	0	0	1
0	1	1	0
1	0	1	0
1	1	1	0

§ 1. 5. КОНЪЮНКТИВНАЯ И ДИЗЪЮНКТИВНАЯ НОРМАЛЬНЫЕ ФОРМЫ

Рассмотрим практические приемы преобразования логических формул и формы их представления.

Преобразования называют *равносильными*, если исходная формула $U(A_1, A_2, \ldots, A_n)$ и формула, полученная в результате преобразований $W(A_1, A_2, \ldots, A_n)$, принимают одинаковые значения истинности на каждом наборе значений аргументов.

Формулы (1.5) — (1.24) указывают равносильность их правых и левых частей.

В практике преобразований использование равносильностей основано на правиле подстановок и выполняется следующим образом: в исходной формуле $U(A_1, A_2, \ldots, A_n)$ отыскивается выражение, полностью совпадающее с левой (правой) частью одной из формул равносильности. Найденное выражение всюду в формуле $U(A_1, A_2, \ldots, A_n)$ заменяется правой (левой) частью. В результате получается формула $W(A_1, A_2, \ldots, A_n)$, по внешнему виду иная, но равносильная исходной.

Кроме указанных, на практике эффективно используются многие другие равносильности, в частности следующие:

$$A \vee A \& B = A, \tag{1.25}$$

$$A \& (A \vee B) = A, \tag{1.26}$$

$$A \& (\overline{A} \vee B) = A \& B, \tag{1.27}$$

$$A \vee \overline{A} \& B = A \vee B, \tag{1.28}$$

$$A \vee A \& B = A \& (A \vee B). \tag{1.29}$$

Равносильные преобразования используются для приведения логических формул к так называемой *нормальной форме*. Такая форма строится с помощью элементарных сумм и произведений.

Элементарным произведением (элементарной суммой) называют произведение (сумму) переменных высказываний и их отрицаний.

Например, для трех высказываний одним из элементарных произведений будет $A_1\overline{A}_2A_3$, а суммой $A_1 \vee \overline{A}_2 \vee A_3$. Всего из n высказываний может быть получено 2^n элементарных произведений (сумм).

Конъюнктивной нормальной формой (КНФ) данной формулы называется равносильная ей формула, состоящая из произведения элементарных сумм.

Дизъюнктивной нормальной формой (ДНФ) данной формулы называется равносильная ей формула, представляющая собой сумму элементарных произведений.

Из определений следует, что в нормальных формах используются только три логические операции — отрицание, конъюнкция и дизъюнкция. Поэтому приведение формулы $U(A_1, A_2, \ldots, A_n)$ к нормальной форме должно выполняться по следующему плану.

В формуле $U(A_1, A_2, \ldots, A_n)$ отношения эквивалентности заменяются по формулам (1.24), (1.16) и (1.17). Отношение импликации заменяется по формулам (1.20) — (1.22), (1.14) и (1.15). Если над некоторыми элементами формулы будет стоять общий знак отрицания, то его нужно заменить по формулам (1.18) и (1.19), чтобы действие отрицания было отнесено только к одному элементарному высказыванию. Двойные отрицания снимаются согласно формуле 1.5.

После этого в формуле останутся элементарные высказывания и их отрицания, объединенные операциями конъюнкции и дизъюнкции. По формулам (1.5) — (1.13), распределительному и сочетательному законам отдельные части формулы приводятся к виду элементарных сумм (произведений) и вместе с этим вся формула — к нормальной форме. Оказывается, результат представления логической формулы в нормальной форме не единственный, хотя бы потому, что в соответствии с равносильностями (1.6) и (1.7) в них могут входить по несколько одинаковых элементарных сумм (произведений). Нормальные формы допускают дальнейшие упрощения, в частности минимизацию. Речь о ней будет идти дальше.

Рассмотрим примеры.

Пример 1. Привести к нормальной конъюнктивной форме выражение

$$(F_1 \to F_2) \sim (\overline{B} \to \overline{A}).$$

Решение. Заменяя знак $\to$ по формуле (1.21), получим

$$\overline{F}_1 \vee F_2 \sim \overline{\overline{B}} \vee \overline{A}.$$

Знак $\sim$ заменим по формуле (1.24), где под A будем понимать $(\overline{F}_1 \vee F_2)$, под B — $(B \vee \overline{A})$;

$$[(\overline{F} \vee F_2) \to (B \vee \overline{A})] \,\&\, [(B \vee \overline{A}) \to (\overline{F}_1 \vee F_2)] =$$
$$=]\overline{\overline{F}_1 \vee F_2} \vee (B \vee \overline{A})] \,\&\, [\overline{B \vee \overline{A}}\; (\overline{F}_1 \vee F_2)] =$$
$$= [\overline{\overline{F}}_1 \,\&\, \overline{F}_2 \vee (B \vee \overline{A})] \,\&\, [\overline{B} \,\&\, \overline{\overline{A}} \vee (\overline{F}_1 \vee F_2)] =$$
$$= [F_1 \,\&\, \overline{F}_2 \vee (B \vee \overline{A})] \,\&\, [\overline{B} \,\&\, A \vee (\overline{F}_1 \vee F_2)].$$

Применим второй распределительный закон. Для облегчения преобразований обозначим $B \vee \overline{A} = Z_1$ и $\overline{F}_1 \vee F_2 = Z_2$

$$[(F_1 \,\&\, \overline{F}_2) \vee Z_1] \,\&\, [(\overline{B} \,\&\, A) \vee Z_2] =$$
$$= [(F_1 \vee Z_1) \,\&\, (\overline{F}_2 \vee Z_1)] \,\&\, [(\overline{B} \vee Z_2) \,\&\, (A \vee Z_2)] =$$
$$= (F_1 \vee B \vee \overline{A}) \,\&\, (F_2 \vee B \vee \overline{A}) \,\&\, (\overline{B} \vee \overline{F}_1 \vee F_2) \,\&\, (A \vee \overline{F}_1 \vee F_2).$$

Это и будет нормальная конъюнктивная форма выражения

$$(F_1 \to F_2) \sim (\overline{B} \to \overline{A}).$$

При приведении к дизъюнктивной нормальной форме порядок действий остается тем же самым, только в конце преобразования следует использовать первый распределительный закон.

Пример 2. Привести к дизъюнктивной нормальной форме сложное высказывание

$$A \,\&\, \overline{\overline{B} \,\&\, C} \,\&\, (M \vee \overline{N}).$$

Решение. По формуле (1.18) устраним знак отрицания над частью высказывания и применим распределительный закон:

$$[A \,\&\, (\overline{\overline{B}} \vee \overline{C})] \,\&\, (M \vee N) = (A \,\&\, B \vee A \,\&\, \overline{C}) \,\&\, (M \vee \overline{N}) =$$
$$= (A \,\&\, B \,\&\, M) \vee (A \,\&\, B \,\&\, \overline{N}) \vee (A \,\&\, \overline{C} \,\&\, M) \vee (A \,\&\, \overline{C} \,\&\, \overline{N}).$$

Это и будет дизъюнктивная нормальная форма.

Сравнивая конъюнктивные и дизъюнктивные нормальные формы, замечаем, что действия конъюнкции и дизъюнкции в них меняются ролями. В одной форме знак $\vee$ входит внутрь отдельных частей формы, а знак & соединяет такие части, в другой форме — наоборот. Оказывается, это перемещение знаков подчинено закономерности и может быть использовано для доказательства новых формул. Знаки $\vee$ и & называют *двойственными*.

Если в логической формуле $F(A_1, A_2, \ldots, A_n)$, приведенной к нормальной форме, все знаки конъюнкции и дизъюнкции заменить на взаимодвойственные, то получим новую формулу, зависящую от тех же аргументов-высказываний.

Формулу, полученную в результате такой замены, называют *двойственной* относительно данной и обозначают $F^{д}(A_1, A_2, A_3, \ldots, A_n)$. Например, двойственными формулами будут

$$(A \vee \overline{A}) \,\&\, B \equiv (A \,\&\, \overline{A}) \vee B;$$
$$\overline{A \vee B} \,\&\, (A \vee \overline{B \,\&\, C}) \equiv \overline{A \,\&\, B} \vee (A \,\&\, \overline{B \vee C}).$$

Принцип двойственности. *Если две формулы равносильны*

$$F_1(A_1, A_2, \ldots, A_n) = F_2(A_1, A_2, \ldots, A_n),$$

то и двойственные им формулы также равносильны

$$F_1^{д}(A_1, A_2, \ldots, A_n) = F_2^{д}(A_1, A_2, \ldots, A_n).$$

Прежде чем перейти к доказательству принципа двойственности, рассмотрим три вспомогательных вопроса.

1. Покажем, что формулы де Моргана (1.18) и (1.19) справедливы для любого конечного числа элементарных высказываний.

Обозначим число элементарных высказываний через n.

Тогда для $n=2$ формулы $\overline{A \& B} = \overline{A} \vee \overline{B}$ и $\overline{A \vee B} = \overline{A} \& \overline{B}$ верны.

Пусть такие формулы будут верны для n элементарных высказываний

$$\overline{A_1 \vee A_2 \vee \ldots \vee A_n} = \overline{A}_1 \& \overline{A}_2 \& \ldots \& \overline{A}_n.$$

Покажем, что они верны и при $n+1$ высказываниях.

Обозначим $A_1 \vee A_2 \vee \ldots \vee A_n = B$, тогда дизъюнкция $B \vee A_{n+1}$ будет состоять из $n+1$ элементарных высказываний.

Согласно формуле (1.19)

$$\overline{B \vee A_{n+1}} = \overline{B} \& \overline{A}_{n+1}.$$

Подставив вместо B его выражение, получим

$$\overline{A_1 \vee A_2 \vee \ldots \vee A_n \vee A_{n+1}} = \overline{A}_1 \& \overline{A}_2 \& \ldots \& \overline{A}_n \& \overline{A}_{n+1},$$

в котором замечаем ту же самую закономерность. Следовательно, формула верна и для $n+1$ элементарных высказываний.

Пользуясь математическими обозначениями для произведения $\prod\limits_{i=1}^{n}$ и суммы $\sum\limits_{i=1}^{n}$ и помня, что дизъюнкцию мы определили как логическое сложение, а конъюнкцию — как логическое умножение, формулы (1.18) и (1.19) запишем в виде

$$\overline{\prod_{i=1}^{n} A_i} = \sum_{i=1}^{n} \overline{A}_i; \tag{1.18'}$$

$$\overline{\sum_{i=1}^{n} A_i} = \prod_{i=1}^{n} \overline{A}_i. \tag{1.19'}$$

Полученный результат обобщает формулы де Моргана на любое конечное число элементарных высказываний.

2. Если две формулы равносильны

$$F_1(A_1, A_2, \ldots, A_n) = F_2(A_1, A_2, \ldots, A_n),$$

то будут равносильны и отрицания этих формул

$$\overline{F}_1(A_1, A_2, \ldots, A_n) = \overline{F}_2(A_1, A_2, \ldots, A_n).$$

Докажем это.

Возьмем произвольный набор значений аргументов A_1^*, A_2^*, ..., A_n^*. Логическая функция $F_1(A_1^*, A_2^*, \ldots, A_n^*)$ на этом наборе примет определённое значение истинности. Так как по условию функции $F_1(A_1, A_2, \ldots, A_n)$ и $F_2(A_1, A_2, \ldots, A_n)$ равносильны, то и $F_2(A_1^*, A_2^*, \ldots, A_n^*)$ на том же наборе примет такое же значение истинности.

Отрицания одинаковых фиксированных значений истинности будут также одинаковы.

Так как при доказательстве рассматривался любой из возможных наборов $A_1^*, A_2^*, \ldots, A_n^*$, то равносильность формулы

$$\overline{F_1}(A_1^*, A_2^*, \ldots, A_n^*) = \overline{F_2}(A_1^*, A_2^*, \ldots, A_n^*)$$

доказана.

3. Если две формулы равносильны:

$$F_1(A_1, A_2, \ldots, A_n) = F_2(A_1, A_2, \ldots, A_n),$$

то будут равносильны формулы, полученные из данных путем замены аргументов на их отрицания, т. е.

$$F_1(\overline{A}_1, \overline{A}_2, \ldots, \overline{A}_n) = F_2(\overline{A}_1, \overline{A}_2, \ldots, \overline{A}_n).$$

Возьмём произвольный набор $A_1^*, A_2^*, \ldots, A_n^*$. Так как формула равносильна, то имеет место

$$F_1(A_1^*, A_2^*, \ldots, A_n^*) = F_2(A_1^*, A_2^*, \ldots, A_n^*).$$

Построим теперь другой набор $\overline{A}_1^*, \overline{A}_2^*, \ldots, \overline{A}_n^*$, полученный из первого путем замены A_i^* на $\overline{A}_i^*$. В силу определения равносильности на этом наборе формулы так же равносильны:

$$F_1(\overline{A}_1^*, \overline{A}_2^*, \ldots, \overline{A}_n^*) = F_2(\overline{A}_1^*, \overline{A}_2^*, \ldots, \overline{A}_n^*).$$

Доказательство принципа двойственности. Пусть формулы $F_1(A_1, A_2, \ldots, A_n)$; $F_2(A_1, A_2, \ldots, A_n)$ — равносильны.

Не уменьшая общности, можно считать, что каждая формула приведена к конъюнктивной нормальной форме

$$F_1(A_1, \overline{A}_2, \ldots, A_n) = \prod_{k=1}^{m} \Big(\sum_{i=1}^{n} A_i^*\Big)_k,$$

где n — число элементарных высказываний в каждом логическом сомножителе, а m — число таких сомножителей.

Заменяя в последнем равенстве все A^* на их отрицания, получим

$$F_1(\overline{A}_1, \overline{\overline{A}}, \ldots, \overline{A}_n) = \prod_{k=1}^{m} \Big(\sum_{i=1}^{n} \overline{A}_i^*\Big)_k.$$

Возьмем отрицание

$$\overline{F}_1(\overline{A}_1, \overline{\overline{A}}_2, \ldots, \overline{A}_n) = \overline{\prod_{k=1}^{m}\Big(\sum_{i=1}^{n}\overline{A}_i^*\Big)_k} = \sum_{k=1}^{m}\Big(\prod_{i=k}^{n}\overline{\overline{A}}_i^*\Big) =$$

$$= \sum_{k=1}^{m}\Big(\prod_{i=1}^{n}A_i^*\Big)_k.$$

Полученный результат $\sum_{k=1}^{m}\Big(\prod_{i=1}^{u}A_i^*\Big)$ есть не что иное, как формула, двойственная относительно $F_1(A_1, A_2, \ldots, A_n)$, так как в ней знаки логических действий поменялись на двойственные.

Таким образом, мы доказали, что

$$\overline{F}_1(\overline{A}_1, \overline{A}_2, \ldots, \overline{A}_n) = F_1^{д}(A_1, A_2, \ldots, A_n).$$

Таким же путем можно доказать, что

$$\overline{F}_2(\overline{A}_1, \overline{A}_2, \ldots, \overline{A}_n) = F_2^{д}(A_1, A_2, \ldots, A_n).$$

Так как по условию

$$F_1(A_1, A_2, \ldots, A_n) = F_2(A_1, A_2, \ldots, A_n),$$

то согласно доказанному ранее будет справедливо

$$\overline{F}_1(\overline{A}_1, \overline{A}_2, \ldots, \overline{A}_n) = \overline{F}_2(\overline{A}_1, \overline{A}_2, \ldots, \overline{A}_n).$$

Мы получили, что двойственные формулы

$$F_1^{д}(A_1, A_2, \ldots, A_n) \text{ и } F_2^{(д)}(A_1, A_2, \ldots, A_n)$$

логически равны равным между собой формулам. Отсюда следует

$$F_1^{д}(A_1, A_2, \ldots, A_n) = F_2^{(д)}(A_1, A_2, \ldots, A_n),$$

что и требовалось доказать.

Принцип двойственности с успехом может быть использован для получения новых логических формул. Например, второй распределительный закон

$$A \vee (B \& C) = (A \vee B) \& (A \vee C)$$

был нами доказан с помощью таблицы истинности. Заменяя в нем все знаки на двойственные, получим формулу

$$A \& (B \vee C) = (A \& B) \vee (A \& C),$$

которая, согласно закону двойственности, также верна. Эта формула есть выражение первого распределительного закона.

§ 1. 6. ТОЖДЕСТВЕННО ИСТИННЫЕ И ТОЖДЕСТВЕННО ЛОЖНЫЕ НОРМАЛЬНЫЕ ФОРМЫ

В соответствии с определением двоичной функции $F(A_1, A_2, \ldots, A_n)$ ее значение истинности зависит от конкретного набора фиксированных значений аргументов. Всего наборов может быть 2^n. В общем случае функция $F(A_1, A_2, \ldots, A_n)$ на некоторой части наборов истинна, на другой части—ложна.

Пусть первых наборов было N_1, вторых—N_2; $N_1+N_2=2^n$, n — число аргументов функции $F(A_1, A_2, \ldots, A_n)$.

В практике особый интерес представляют два крайних случая, когда $N_1=2^n$, $N_2=0$ и $N_1=0$, $N_2=2^n$, т. е. двоичная функция $F(A_1, A_2, \ldots, A_n)$ на всех наборах либо истинна, либо ложна.

В первом случае функцию называют *тождественно истинной*, во втором — *тождественно ложной*. Тождественно истинные и тождественно ложные функции называют *тождественно постоянными*.

Разумеется, выполняя равносильные преобразования над какой-либо логической функцией, важно определить, является ли она тождественно постоянной. В последнем случае вся формула заменится одним символом И или Л. Решить этот вопрос можно путем проведения равносильных преобразований (см. пример 2 § 1.4). Можно использовать таблицы истинности. Но оба эти способа громоздки. Если функция $F(A_1, A_2, \ldots, A_n)$ приведена к конъюнктивной или дизъюнктивной нормальной форме, то определить принадлежность ее к типу тождественно постоянных функций можно по внешнему виду.

Докажем два вспомогательных предложения.

1. *Для того чтобы элементарная сумма* $A_1 \vee A_2 \vee \ldots \vee A_n$ *была тождественно истинной, необходимо и достаточно, чтобы среди ее слагаемых нашлось хотя бы одно элементарное высказывание и его отрицание, например* A_i *и* $\overline{A_i}$.

Достаточность. Пусть в элементарной сумме есть элементарное высказывание A_i и его отрицание $\overline{A_i}$. Тогда на любом наборе A_i будет истинным или ложным. Если A_i истинно, то в соответствии с определением дизъюнкции, будет истинна вся элементарная сумма. Если A_i ложно, то $\overline{A_i}$ истинно, и в силу этого элементарная сумма также истинна.

Необходимость. Пусть в элементарной сумме все высказывания различны. Среди них часть, например m_1, может быть с отрицанием, часть m_2 — без отрицания; $m_1+m_2=n$. Среди всех 2^n наборов существует такой, на котором все m_2 высказываний ложны, а m_1 истинны. Тогда в элементарной сумме каждое слагаемое окажется ложным и вместе с этим элементарная сумма будет ложна. На других наборах может оказаться, что хотя бы одно слагаемое истинно.

Таким образом, отсутствие в элементарной сумме высказывания и его отрицания создает обстановку, в которой элементарная сумма может быть истинной и ложной, что и требовалось доказать.

2. *Для того чтобы элементарное произведение было тождественно ложным, необходимо и достаточно, чтобы среди его сомножителей нашлось хотя бы одно элементарное высказывание и его отрицание, например A_i и $\overline{A_i}$.*

Доказательство строится на определении конъюнкции и протекает аналогично предыдущему.

Два последних предложения позволяют указать практические приемы определения тождественного постоянства логической формулы.

Делается это следующим образом.

Логическая формула приводится к нормальной форме (конъюнктивной или дизъюнктивной).

а) В случае приведения к конъюнктивной нормальной форме рассматривается по отдельности каждый сомножитель $\ldots \& (A_1 \vee \overline{A_2} \vee \ldots \vee A_i \vee \overline{A_i} \vee \ldots \vee A_n) \& \ldots$ Если в каждом из них одно из высказываний A_i входит в элементарную сумму вместе со своим отрицанием $\overline{A_i}$, то все сомножители будут принимать значение «истинно» и вследствие этого вся логическая формула будет тождественно истинной.

б) В случае приведения к дизъюнктивной нормальной форме рассматривается по отдельности каждое слагаемое $\ldots \ldots \vee A_1 \& \overline{A_2} \& \ldots \& A_i \& \overline{A_i} \& \ldots \& A_n \vee \ldots$ Если в каждом из них одно из высказываний A_i входит в элементарное произведение вместе со своим отрицанием $\overline{A_i}$, то слагаемые будут принимать значения «ложно» и вследствие этого вся логическая формула будет тождественно ложной.

Пример 1. Установить характер истинности формулы

$$[P \& (P \to Q)] \to Q.$$

Решение. Приведем формулу к конъюнктивной нормальной форме

$$[P \& (P \to Q)] \to Q = \overline{P \& (\overline{P} \vee Q)} \vee Q = (\overline{P} \vee \overline{\overline{P} \vee Q}) \vee Q =$$
$$= (\overline{P} \vee \overline{\overline{P}} \& \overline{Q}) \vee Q = [(\overline{P} \vee P) \& (\overline{P} \vee \overline{Q})] \vee Q = (\overline{P} \vee P \vee Q) \& (\overline{P} \vee \overline{Q} \vee Q)$$

(в преобразовании использованы формулы (1.5), (1.18), (1.19), (1.21) и второй распределительный закон дизъюнкции).

Рассматривая полученную конъюнктивную форму, замечаем, что первый логический сомножитель имеет в своем составе элементарное высказывание P и его отрицание $\overline{P}$, во втором сомножителе есть Q и $\overline{Q}$. Следовательно, рассматриваемая формула тождественно истинная.

Пример 2. Определить характер истинности формулы

$$\overline{B}\&\overline{A\&\overline{B}}\&\overline{\overline{A}\&\overline{B}}.$$

Решение. Приведем к конъюнктивной нормальной форме:

$$\overline{B}\&\overline{A\&\overline{B}}\&\overline{\overline{A}\&\overline{B}}=\overline{B}\&(\overline{A}\vee\overline{\overline{B}})\&(\overline{\overline{A}}\vee\overline{\overline{B}})=\overline{B}\&(\overline{A}\vee B)\&(A\vee B).$$

Логические сомножители не удовлетворяют предложению 1. В силу этого судить об их истинности нет оснований.

Приведем ту же формулу к дизъюнктивной нормальной форме:

$$\overline{B}\&(\overline{A}\vee B)\&(A\vee B)=(\overline{B}\&\overline{A}\vee\overline{B}\&B)\&(A\vee B)=$$
$$=\overline{B}\&\overline{A}\&A\vee\overline{B}\&B\&A\vee\overline{B}\&\overline{A}\&B\vee\overline{B}\&B\&B$$

(в преобразовании использован первый распределительный закон).

Каждое логическое слагаемое полученной дизъюнктивной формы удовлетворяет условию предложения 2. Вследствие этого рассматриваемая формула тождественно ложна.

Пример 3. Определить характер истинности формулы

$$P_1\&P_2\&P_3\vee[(P_1\vee P_2\vee\overline{P_3})\to\overline{\overline{P_1}\vee P_2\vee P_3}].$$

Решение. Для замены импликации и общего знака отрицания над следствием импликации применим формулы (1.21) и (1.18) соответственно:

$$P_1\&P_2\&P_3\vee[(P_1\vee P_2\vee\overline{P_3})\to\overline{\overline{P_1}\vee P_2\vee P_3}]=$$
$$=P_1\&P_2\&P_3\vee(\overline{P_1\vee P_2\vee\overline{P_3}}\vee\overline{\overline{P_1}\vee P_2\vee P_3})=$$
$$=P_1\&P_2\&P_3\vee\overline{P_1}\&\overline{P_2}\&\overline{\overline{P_3}}\vee\overline{\overline{P_1}}\&\overline{P_2}\&\overline{P_3}=$$
$$=P_1\&P_2\&P_3\vee\overline{P_1}\&\overline{P_2}\&P_3\vee P_1\&\overline{P_2}\&P_3.$$

В полученной дизъюнктивной нормальной форме в каждый сомножитель входят различные элементарные высказывания. Потому не выполняются условия предложений 1 и 2. Этим установлено, что логическая формула не является тождественно постоянной.

И в самом деле. На наборе $P_1=0$, $P_2=0$ и $P_3=1$ формула $U(P_1, P_2, P_3)$ истинна, так как среднее слагаемое на этом наборе истинно. На наборе $P_1=0$, $P_2=1$, $P_3=0$ все три элементарных слагаемых ложны и вместе с этим ложна вся формула.

§ 1. 7. СОВЕРШЕННЫЕ НОРМАЛЬНЫЕ ФОРМЫ. РАЗЛОЖЕНИЕ ЛОГИЧЕСКИХ ФОРМУЛ ПО КОНСТИТУЕНТАМ

Нормальные формы (конъюнктивная и дизъюнктивная) удобны для решения задачи определения тождественной истинности и ложности формул. При решении других задач они оказываются недостаточно удобными.

Недостаток их состоит в том, что представление логической формулы в нормальной форме не единственно. Это объясняется наличием ряда равносильностей в исчислении высказываний. Например, по формулам (1.6) и (1.7) конъюнкция (дизъюнкция) конечного числа одного и того же высказывания логически равна самому высказыванию. В соответствии с этим формула $A\&B \vee A\&B$ может быть заменена одним слагаемым $A\&B$. В соответствии с формулой (1.10) формула $A\&(B \vee \overline{B})$ логически равна одному высказыванию A, так как $(B \vee \overline{B})$ тождественно истинно. В соответствии с формулой (1.12) $C \vee D\&\overline{D} = C$, так как $D\&\overline{D}$ тождественно ложна и т. д. В соответствии с формулами (1.10 — 1.17) к каждой из нормальных конъюнктивных форм можно приписывать любое конечное число тождественно истинных сомножителей, а к дизъюнктивной — тождественно ложных слагаемых. Этим и объясняется неединственность представления данной формулы в нормальной форме. Отсюда возникает потребность найти такой вид равносильных преобразований, чтобы данной формуле соответствовала единственная нормальная форма. Такой вид преобразований существует.

Пусть мы имеем n элементарных высказываний A_1, $A_2, \ldots, A_n$. Будем выписывать из них строки так, чтобы удовлетворялись следующие условия:

1) в каждую строку входят все n элементарных высказываний;

2) каждое элементарное высказывание входит в строку только один раз;

3) в строку входит или высказывание A_i или его отрицание $\overline{A_i}$ (т. е. A_i и $\overline{A_i}$ входить в одну и ту же строку не могут).

Построим такие строки для одного, двух и трех элементарных высказываний. Для одного высказывания A всем условиям удовлетворяют две строки: A и $\overline{A}$.

Для двух элементарных высказываний A_1 и A_2 таких строк будет четыре:

A_1, A_2 — без отрицаний,

$\left.\begin{matrix}\overline{A_1}, A_2 \\ A_1, \overline{A_2}\end{matrix}\right\}$ — с одним отрицаемым высказыванием,

$\overline{A_1}, \overline{A_2}$ — с двумя отрицаемыми высказываниями;

для трех высказываний — восемь:

A_1, A_2, A_3 — без отрицаний,

$$\left.\begin{matrix}\overline{A}_1, & A_2, & A_3\\ A_1, & \overline{A}_2, & A_3\\ A_1, & A_2, & \overline{A}_3\end{matrix}\right\}\text{— с одним отрицаемым высказыванием,}$$

$$\left.\begin{matrix}\overline{A}_1, & \overline{A}_2, & A_3\\ A_1, & \overline{A}_2, & \overline{A}_3\\ \overline{A}_1, & A_2, & \overline{A}_3\end{matrix}\right\}\text{— с двумя отрицаемыми высказываниями,}$$

$\overline{A}_1, \overline{A}_2, \overline{A}_3$ — с тремя отрицаемыми высказываниями.

Легко заметить закономерность получения числа строк каждой группы и всех строк, получаемых из n элементарных высказываний. Без отрицаний и со всеми отрицаемыми элементарными высказываниями возникает всегда по одной строке. Другие группы строк отличаются друг от друга числом отрицаемых высказываний, а внутри группы — самим отрицаемым элементом (порядок элементов в строке роли не играет). Такое соединение элементарных высказываний удовлетворяет определению сочетания из n элементов по m, где n — число всех элементов, m — число элементов с отрицанием. Число всех строк равно сумме числа строк в каждой группе. Так для четырех элементарных высказываний A_1, A_2, A_3, A_4 найдем

$$C_4^0 + C_4^1 + C_4^2 + C_4^3 + C_4^4 = 16,$$

где C_4^0 — строка, состоящая из заданных элементарных высказываний без отрицаний; C_4^1 — четыре строки, в каждую из которых одно из высказываний входит с отрицанием; C_4^2—шесть строк, в каждую из которых два высказывания входят с отрицанием; C_4^3 — четыре строки с тремя отрицаемыми высказываниями в каждой; C_4^4 — строка, в которой все четыре высказывания с отрицаниями.

В общем случае для n элементарных высказываний A_1, A_2, ..., A_n количество строк определится суммой

$$C_n^0 + C_n^1 + C_n^2 + \ldots + C_n^{n-1} + C_n^n = 2^n,$$

представляющей собой сумму биноминальных коэффициентов.

Если элементарные высказывания A_1, A_2, ..., A_n, построенные в строки в соответствии с требованиями 1) — 3), объединить конъюнктивной или дизъюнктивной связью

$$A_1 \& \overline{A}_2 \& \ldots \& A_n,$$

$$A_1 \vee \overline{A}_2 \vee \ldots \vee \overline{A}_n,$$

то получим элементарные произведения и элементарные суммы соответственно.

Описанные логические конструкции имеют важное значение и получили наименование конституент.

Конституентой называют логическую конструкцию, построенную из элементарных высказываний $A_1, A_2, \ldots, A_n$ и их отрицаний путем выписывания строк, удовлетворяющих требованиям 1) — 3), и соединения элементов строки с помощью конъюнктивной (дизъюнктивной) связи.

Отличие конституенты от элементарной суммы и произведения, получаемых в процессе приведения логической формулы к конъюнктивной или дизъюнктивной нормальной форме, состоит в том, что элементарные суммы и произведения не обязаны удовлетворять требованиям 1) — 3), предъявляемым к конституенте.

Отличие конституенты от набора фиксированных значений аргументов $A_1^*, A_2^*, \ldots, A_n^*$ состоит в том, что набор — это фиксированное значение, а конституента — двоичная функция.

Число их совпадает и равно 2^n, но это внешнее совпадение, так как набор и конституента имеют различную логическую природу.

Конституенты используют для получения совершенных нормальных форм логических формул.

Совершенной конъюнктивной нормальной формой (СКНФ) логической функции $U(A_1, A_2, \ldots, A_n)$ называют представление ее в виде конъюнкции дизъюнктивных конституент, построенных из аргументов $A_1, A_2, \ldots, A_n$ рассматриваемой функции:

$$U(A_1, A_2, \ldots, A_n) = A_1 \vee A_2 \vee \ldots \vee A_n \,\&\, \overline{A}_1 \vee A_2 \vee \ldots \vee A_n \,\&\ldots$$
$$\ldots\&\,\overline{A}_1 \vee \overline{A}_2 \vee \ldots \vee \overline{A}_n.$$

Совершенной дизъюнктивной нормальной формой (СДНФ) логической функции $U(A_1, A_2, \ldots, A_n)$ называют представление ее в виде дизъюнкции конъюнктивных конституент, построенных из аргументов $A_1, A_2, \ldots, A_n$ рассматриваемой функции

$$U(A_1, A_2, \ldots, A_n) = A_1 \& A_2 \& \ldots \& A_n \vee \overline{A}_1 \& A_2 \& \ldots$$
$$\ldots \& A_n \vee \ldots \vee \overline{A}_1 \& \overline{A}_2 \& \ldots \& \overline{A}_n.$$

Каждая логическая формула может быть представлена единственным образом в соответствующей ей совершенной нормальной форме. В ней могут участвовать или все конституенты, которые можно построить из n высказываний-аргументов, или часть их. Но дважды одна и та же конституента входить в совершенную нормальную форму не может.

Укажем последовательность преобразований, приводящую логическую формулу к совершенной дизъюнктивной нормальной форме:

1. Логическая формула $F(A_1, A_2, \ldots, A_n)$ приводится к дизъюнктивной нормальной форме

$$F(A_1, A_2, \ldots, A_n) = \varphi_1(A_1, A_2, \ldots, A_n) \vee \varphi_2(A_1, A_2, \ldots, \ldots, A_m) \vee \ldots \vee \varphi_n(A_1, A_2, \ldots, A_k).$$

Каждое слагаемое представляет конъюнкцию некоторого числа элементарных высказываний $A_1, A_2, \ldots, A_n$ и их отрицаний (но не обязательно всех).

2. Пусть слагаемое $\varphi_i(A_1, A_2, \ldots, A_n)$ не содержит ни A_i, ни $\overline{A_i}$. В таком случае к нему можно конъюнктивно присоединить тождественно истинную формулу $A_i \vee \overline{A_i}$:

$$\varphi_i(A_1, A_2, \ldots, A_n) \& (A_i \vee \overline{A_i}),$$

отчего значение истинности согласно формуле (1.10) не изменится. Полученное выражение развернем по первому распределительному закону:

$$[\varphi_i(A_1, A_2, \ldots, A_n) \& A_i] \vee [\varphi_i(A_1, A_2, \ldots, A_n) \& \overline{A_i}].$$

В слагаемых последнего выражения появились отсутствовавшие ранее сомножители. (Если отсутствовало несколько логических сомножителей A_i, A_p, A_q, то указанную операцию следует провести для всех отсутствующих сомножителей.) В результате каждое логическое слагаемое будет удовлетворять условию 1).

3. Если в каком-либо слагаемом есть несколько одинаковых сомножителей, то все их можно заменить согласно формуле (1.5) одним. В результате каждое слагаемое будет удовлетворять условию 2). Например,

$$\varphi_j(A_1, A_2, A_3) = A_1 \& \overline{A_2} \& A_3 \& \overline{A_2} = A_1 \& \overline{A_2} \& A_3.$$

4. Если какие-либо слагаемые содержат A_i и $\overline{A_i}$, то по определению конъюнкции произведение $A_i \& \overline{A_i}$ будет тождественно ложным и в силу формулы (1.11) эти слагаемые можно отбросить. В результате форма будет удовлетворять условию 3). Например,

$$\varphi(A_1, A_2, A_3) = A_1 \& \overline{A_2} \& A_3 \vee A_1 \& A_2 \& \overline{A_1} =$$
$$= A_1 \& \overline{A_2} \& A_3 \vee Л = A_1 \& \overline{A_2} \& A_3.$$

5. Если среди всех логических слагаемых окажется два или несколько одинаковых, то их можно заменить одним (согласно формуле (1.7)).

В результате всех операций каждое логическое слагаемое будет являться конституентой от n элементарных высказываний, и так как все полученные конституенты объединяются с помощью дизъюнктивной связи, то полученная форма будет совершенной дизъюнктивной нормальной формой.

При приведении логической формулы к конъюнктивной совершенной форме последовательность преобразований остается той же, но все действия заменяются на двойственные. При этом конституенты примут вид логических сумм, объединенных конъюнктивной связью. Иногда действие приведения логической формулы к совершенной нормальной форме называют *разложением формулы по конституентам.*

Такое преобразование рассмотрим на примере.

Пример 1. Привести к совершенной дизъюнктивной нормальной форме функцию $\overline{A} \to [B\&(A \vee \overline{B})]$.

Решение. Прежде приведем формулу к дизъюнктивной нормальной форме

$$\overline{A} \to [B \& (A \vee \overline{B})] = \overline{\overline{A}} \vee [B \& (A \vee \overline{B})] = A \vee B \& A \vee B \& \overline{B}.$$

Полученная форма хотя и дизъюнктивна, но не совершенна, так как первое слагаемое не содержит B, а последнее содержит B и его отрицание.

Присоединим конъюнктивно к высказыванию A тождественно истинную формулу $B \vee \overline{B}$ и преобразуем:

$$A \& (B \vee \overline{B}) = A \& B \vee A \& \overline{B}.$$

Получим

$$A \& B \vee A \& \overline{B} \vee B \& A \vee B \& \overline{B}.$$

Первое и третье слагаемые одинаковы (с точностью до порядка расположения сомножителей), поэтому их можно заменить одним. Четвертое слагаемое тождественно ложно, поэтому его следует отбросить. В результате получим $A \& \overline{B} \vee B \& A$.

Это выражение отвечает требованиям определения совершенной нормальной дизъюнктивной формы логической формулы. Следовательно,

$$\overline{A} \to [B \& (A \vee \overline{B})] = A \& \overline{B} \vee B \& A.$$

Пример 2. Привести к совершенной конъюнктивной нормальной форме функцию $(A \sim B) \& (B \sim C)$.

Решение. Приведем к конъюнктивной нормальной форме:

$$(A \sim B) \& (B \sim C) = [(A \to B) \& (B \to A)] \& [(B \to C) \& (C \to B)] =$$

$$= (\overline{A} \vee B) \& (\overline{B} \vee A) \& (\overline{B} \vee C) \& (\overline{C} \vee B).$$

В каждом сомножителе отсутствует одно из высказываний-аргументов. Добавим с помощью дизъюнктивной связи недостающие высказывания в виде тождественно ложных выражений и развернем формулу по второму распределительному закону:

$$[\overline{A}\vee B\vee C\&\overline{C}]\&[\overline{B}\vee A\vee C\&\overline{C}]\&[\overline{B}\vee C\vee A\&\overline{A}]\&$$
$$\&[\overline{C}\vee B\vee(A\&\overline{A})]=(\overline{A}\vee B\vee C)\&(\overline{A}\vee B\vee\overline{C})\&(\overline{B}\vee A\vee C)\&$$
$$\&(\overline{B}\vee A\vee\overline{C})\&(\overline{B}\vee C\vee A)\&(\overline{B}\vee C\vee\overline{A})\&(\overline{C}\vee B\vee A)\&$$
$$\&(\overline{C}\vee B\vee\overline{A}).$$

В полученном выражении каждый сомножитель является конституентой. Второй и последний, третий и пятый сомножители попарно одинаковы, поэтому каждую пару заменим одним сомножителем. В результате получим совершенную конъюнктивную нормальную форму:

$$(A\sim B)\&(B\sim C)=(\overline{A}\vee B\vee C)\&(\overline{A}\vee B\vee\overline{C})\&(\overline{B}\vee A\vee C)\&$$
$$\&(\overline{B}\vee A\vee\overline{C})\&(\overline{B}\vee C\vee\overline{A})\&(\overline{C}\vee B\vee A).$$

Для полного набора конституент в рассмотренной формуле не хватает двух: $A\vee B\vee C$ и $\overline{A}\vee\overline{B}\vee\overline{C}$.

Имеет место следующее весьма важное для практики предложение.

Каждая двоичная функция $U(A_1, A_2, \ldots, A_n)$ может быть представлена в совершенной дизъюнктивной (или конъюнктивной) нормальной форме.

Докажем это.

Пусть имеем двоичную логическую функцию $U(A_1, A_2, \ldots, A_n)$. В общем случае можно полагать, что в формульное выражение функции $U(A_1, A_2, \ldots, A_n)$ входит аргумент A_i и его отрицание $\overline{A_i}$. Если такого явления нет, например отсутствует отрицание высказывания $\overline{A_j}$, то к рассматриваемой формуле добавляем с помощью конъюнктивной связи тождественно истинное высказывание $A_j\vee\overline{A_j}$:

$$U(A_1, A_2, \ldots, A_n)\&(A_j\vee\overline{A_j}).$$

Полученная функция, согласно формуле (1.10), равносильна данной.

Используя распределительный закон, получим равносильную функцию, содержащую все аргументы A_i и их отрицания $\overline{A_i}, i=1, \ldots, n$.

Приведем функцию $U(A_1, A_2, \ldots, A_n)$ к дизъюнктивной нормальной форме. Сгруппируем элементарные произведения так, чтобы в одну группу входили члены, содержащие A_1, в другую — $\overline{A}_1$. В каждой группе вынесем за скобку сомножители A_1 и $\overline{A}_1$. Получим:

$$U(A_1, A_2, \ldots, A_n) = A_1 \& M_1(A_2, A_3, \ldots, A_n) \vee \\ \vee \overline{A}_1 \& N_1(A_2, A_3, \ldots, A_n). \qquad (1.30)$$

Выражения $M_1(A_2, \ldots, A_n)$ и $N_1(A_2, \ldots, A_n)$ представляют собой дизъюнктивную нормальную форму, но не содержат A_1 и $\overline{A}_1$.

Рассмотрим значение функции $U(A_1, A_2, \ldots, A_n)$ при фиксированном значении $A_1 = 1$. Тогда $\overline{A}_1 = 0$. Подставив эти значения в формулу (1.30), получим

$$U(1, A_2, \ldots, A_n) = 1 \& M_1(A_2, \ldots, A_n) \vee 0 \& N_1(A_2, \ldots, A_n).$$

Отсюда

$$M_1(A_2, \ldots, A_n) = U(1, A_2, \ldots, A_n).$$

Полагая $A_1 = 0$, $\overline{A}_1 = 1$, получим

$$N_1(A_2, \ldots, A_n) = U(0, A_2, \ldots, A_n).$$

Подставляем в формулу (1.30) значения M_1 и N_1:

$$U(A_1, A_2, \ldots, A_n) = A_1 \& U(1, A_2, \ldots, A_n) \vee \\ \vee \overline{A}_1 \& U(0, A_2, \ldots, A_n).$$

Проделаем такие же операции разложения по A_2 и $\overline{A}_2$ над выражениями M_1 и N_1. Полагая поочередно $A_2 = 1$, $\overline{A}_2 = 0$ и $A_2 = 0$, $\overline{A}_2 = 1$, получим:

$$M_1 = A_2 \& U(1, 1, A_3, \ldots, A_n) \vee \overline{A}_2 \& U(1, 0, A_3, \ldots, A_n);$$

$$N_1 = A_2 \& U(0, 1, A_3, \ldots, A_n) \vee \overline{A}_2 \& U(0, 0, A_3, \ldots, A_n).$$

Выражения M_1 и N_1 подставим в (1.30):

$$U(A_1, A_2, \ldots, A_n) = A_1 \& [A_2 \& U(1, 1, A_3, \ldots, A_n) \vee \\ \vee \overline{A}_2 \& U(1, 0, A_3, \ldots, A_n)] \vee \overline{A}_1 \& [A_2 \& U(0, 1, A_3, \ldots, A_n) \vee \\ \vee \overline{A}_2 \& U(0, 0, A_3, \ldots, A_n)] = A_1 \& A_2 \& U(1, 1, A_3, \ldots, A_n) \vee \\ \vee A_1 \& \overline{A}_2 \& U(1, 0, A_3, \ldots, A_n) \vee \\ \vee \overline{A}_1 \& A_2 \& U(0, 1, A_3, \ldots, A_n) \vee \overline{A}_1 \overline{A}_2 U(0, 0, A_3, \ldots, A_n).$$

Такие же преобразования проделаем с A_3 и $\overline{A}_3$, с A_4 и $\overline{A}_4$ и т. д. Подметим закономерность возникновения очередных выражений и по индукции напишем результат.

1. При разложении по одному аргументу A_1 получили два слагаемых.

При разложении по двум аргументам получили $2^2 = 4$ слагаемых.

При разложении по трем аргументам найдем $2^3 = 8$ слагаемых, так как каждый из предыдущих слагаемых представится суммой двух новых слагаемых.

При разложении $U(A_1, A_2, \ldots, A_n)$ по всем n аргументам получим в разложении 2^n элементарных слагаемых.

2. Каждое слагаемое разложения представляет собой конъюнкцию некоторого числа аргументов функции $U(A_1, A_2, \ldots, A_n)$. При этом, если A_i вынесен в качестве сомножителя, то в выражении функции он заменен фиксированным значением 1 для A_i и 0 для $\overline{A}_i$.

Руководствуясь указанными закономерностями, можно записать полное разложение логической функции по конституентам:

$$\begin{aligned} U(A_1, A_2, \ldots, A_n) = A_1 A_2 \ldots\ A_n U(1, 1, 1, \ldots, 1) \vee \\ \vee \overline{A}_1 A_2 \ \ldots A_n U(0, 1, 1, \ldots, 1) \vee \ldots \vee A_1 A_2 \ldots \overline{A}_i \ldots \\ \ldots \overline{A}_k \ldots A_n U(1, 1, \ldots, 0_i, \ldots, 0_k, \ldots, 1) \vee \\ \vee \ldots \vee \overline{A}_1 \overline{A}_2 \ldots \overline{A}_n U(0, 0, \ldots, 0). \end{aligned} \tag{1.31}$$

Здесь знак & опущен. Анализируя выражение (1.31), замечаем, что в нем сомножители $U(A_1^*, A_2^*, \ldots, A_n^*)$ представляют собой числа 1 или 0, равные значению функции $U(A_1, A_2, \ldots, A_n)$ на определенном наборе фиксированных значений аргументов $A_1^*, A_2^*, \ldots, A_n^*$. Заменяя в формуле (1.31) сомножители $U(A_1^*, A_2^*, \ldots, A_n^*)$ соответствующим числовым значением, будем иметь дизъюнкцию элементарных произведений. Построение их таково, что в каждое произведение входят все n аргументов. Ни один из них дважды не повторяется и входит либо сам аргумент A_i, либо его отрицание, т. е. элементарные произведения представляют собой конституенты.

Мы доказали, что любая логическая функция $U(A_1, A_2, \ldots, A_n)$ может быть представлена в виде разложения по конституентам. Число конституент не больше 2^n. Отсутствовать в разложении будут те конституенты, у которых функция $U(A_1^*, A_2^*, \ldots, A_n^*)$ на наборе, соответствующем конституенте, обращается в нуль. Таким образом,

$$U(A_1, A_2, \ldots, A_n) = \sum_{j=1}^{2^n} \left(\prod_{i=1}^{n} \alpha_j \tilde{A}_i^* \right),$$

где $\tilde{A}_i^*$ обозначает одно из двух значений высказывания A_i или $\overline{A}_i$, коэффициент α_j равен 1 или 0 и символизирует наличие или отсутствие данного члена в разложении.

Аналогично производится разложение функции $U(A_1, A_2, \ldots, A_n)$ по дизъюнктивным конституентам.

В основе разложения лежит нормальная конъюнктивная форма, позволяющая представить разлагаемую формулу в виде конъюнкции

$$U(A_1, A_2, \ldots, A_n) =$$
$$= [\overline{A}_1 \vee M(A_2, A_3, \ldots, A_n)][A_1 \vee N(A_2, A_3, \ldots, A_n)].$$

Полагая $A_1 = 1$, $\overline{A}_1 = 0$, затем $A_1 = 0$, $\overline{A}_1 = 1$, находим $M(A_2, \ldots, A_n)$ и $N(A_2, \ldots, A_n)$. Продолжая разложение аналогично предыдущему, получим

$$U(A_1, A_2, \ldots, A_n) = [U(0, 0, \ldots, 0) \vee$$
$$\vee A_1 \vee A_2 \vee \ldots \vee A_n] \& [U(1, 0, \ldots, 0) \vee \overline{A}_1 \vee A_2 \vee \ldots \vee A_n] \&$$
$$\& \ldots \& [U(1, 1, \ldots, 1) \vee \overline{A}_1 \vee \overline{A}_2 \vee \ldots \vee \overline{A}_n].$$

Заменяя $U(A_1^*, A_2^*, \ldots, A_n^*)$ значением 0 или 1 в соответствии с набором $A_1^*, A_2^*, \ldots, A_n^*$, получим разложение функции по дизъюнктивным конституентам:

$$U(A_1, A_2, \ldots, A_n) = \prod_{j=1}^{2^n} \Big(\sum_{i=1}^{n} \alpha_j A_i\Big). \qquad (1.32)$$

Пример 3. Разложить по конституентам тождественно истинную формулу. Эту же задачу можно сформулировать иначе: разложить по конституентам логическую единицу.

Решение. Согласно формуле (1.1) и определению конъюнкции $A_1 \vee \overline{A}_1 = 1$ и $A_2 \vee \overline{A}_2 = 1$. Тогда

$$1 = 1 \& 1 = (A_1 \vee \overline{A}_1) \& (A_2 \vee \overline{A}_2).$$

Развернув полученное выражение по первому распределительному закону, получим

$$1 = A_1 \& A_2 \vee A_1 \& \overline{A}_2 \vee \overline{A}_1 \& A_2 \vee \overline{A}_1 \& \overline{A}_2.$$

Каждое слагаемое представляет конституенту, причем имеется полный набор конституент, которые могут быть получены из двух элементарных высказываний. Это не случайно, так как отсутствие хотя бы одной конституенты не позволит сделать вывод о тождественной истинности всей формулы.

В общем случае тождественно истинная функция, зависящая от n высказываний-аргументов, будет представляться совершенной дизъюнктивной нормальной формой, составленной из всех 2^n конституент, которые могут быть составлены из n элементарных высказываний-аргументов:

$$1 = И = A_1 \& A_2 \& \dots \& A_n \vee \overline{A}_1 \& A_2 \& \dots \& A_n \vee \dots$$
$$\dots \vee \overline{A}_1 \& \overline{A}_2 \& \dots \& \overline{A}_n.$$

Пример 4. Разложить по конституентам тождественно ложную формулу (разложить по конституентам логический нуль).

Решение. Задачу можно решить путем аналогичных рассуждений, используя двойственные действия. Можно решить ее еще проще: взять отрицание от обеих частей разложения по конституентам И·

$$\overline{И} = \overline{A_1 \& A_2 \& \dots \& A_n \vee \overline{A}_1 \& A_2 \& \dots \& A_n \vee \dots \vee \overline{A}_1 \& \overline{A}_2 \& \dots \& \overline{A}_n} =$$
$$= \overline{A_1 \& A_2 \& \dots \& A_n} \& \overline{\overline{A}_1 \& A_2 \& \dots \& A_n} \& \dots \& \overline{\overline{A}_1 \& \overline{A}_2 \& \dots \& \overline{A}_n} =$$
$$= (\overline{A}_1 \vee \overline{A}_2 \vee \dots \vee \overline{A}_n) \& (\overline{\overline{A}}_1 \vee \overline{A}_2 \dots \vee \overline{A}_n) \& \dots \& (\overline{\overline{A}}_1 \vee \overline{\overline{A}}_2 \vee \dots \vee \overline{\overline{A}}_n).$$

Таким образом,

$$Л = (\overline{A}_1 \vee \overline{A}_2 \vee \dots \vee \overline{A}_n) \& (A_1 \vee \overline{A}_2 \vee \dots \vee \overline{A}_n) \& \dots$$
$$\dots \& (A_1 \vee A_2 \vee \dots \vee A_n).$$

Совершенная нормальная форма построена так, что каждой логической формуле соответствует единственная совершенная нормальная форма и наоборот. Однако эти формы не являются наиболее простыми.

Можно указать логические преобразования, которые приведут данную совершенную нормальную форму к более простому виду, не нарушая равносильности. При этом утрачивается принцип единственности представления формулы, но во многих прикладных логических исследованиях этот прицип играет второстепенную роль.

§ 1. 8. ВЫВОД ВСЕХ СЛЕДСТВИЙ ИЗ ДАННЫХ ПОСЫЛОК

В математической логике и в ее технических приложениях важное значение имеет получение логических следствий из некоторых исходных двоичных функций.

Рассмотрим вначале следующие вопросы:

1. Что такое логическое следствие в строгой математической формулировке?

2. Как найти функцию, являющуюся логическим следствием нескольких данных?

3. Сколько может быть получено логических следствий из посылок?

В соответствии с правилом заключений, сформулируем следующее определение логического следствия.

Двоичная функция $U(A_1, A_2, \ldots, A_n)$ является *логическим следствием* посылок

$$F_1(A_1, A_2, \ldots, A_n);\ F_2(A_1, A_2, \ldots, A_n);\ \ldots;\ F_m(A_1, A_2, \ldots, A_n),$$

если импликация

$$F_1(A_1, A_2, \ldots, A_n) \& F_2(A_1, A_2, \ldots, A_n) \& \ldots$$
$$\ldots \& F_m(A_1, A_2, \ldots, A_n) \to U(A_1, A_2, \ldots, A_n)$$

тождественно истинная функция.

Функции $F_i(A_1, A_2, \ldots, A_n)$ в общем случае могут быть представлены с помощью любых логических связей и зависеть от различного числа аргументов.

Приемы практического нахождения логических следствий из данных посылок строятся на основе следующего предложения.

Для того чтобы двоичная функция $U(A_1, A_2, \ldots, A_n)$, представленная в совершенной конъюнктивной нормальной форме, являлась логическим следствием посылок $F_i(A_1, A_2, \ldots, A_n)$, $i = 1, \ldots, k$, необходимо и достаточно, чтобы конституенты, составляющие $U(A_1, A_2, \ldots, A_n)$, входили составными частями в конъюнкцию посылок

$$F_1(A_1, A_2, \ldots, A_n) \& F_2(A_1, A_2, \ldots, A_n) \& \ldots$$
$$\ldots \& F_k(A_1, A_2, \ldots, A_n).$$

Докажем это.

Не уменьшая общности, можно полагать, что в каждую из посылок $F_i(A_1, A_2, \ldots, A_m)$ входят все n аргументов. Если в какой-либо функции отсутствует тот или иной аргумент A_j, то его можно добавить с помощью конъюнкции тождественной единицы $(A_j \vee \bar{A}_j.)$

Необходимость. Конъюнкцию посылок представим в виде одной формулы $\Phi(A_1, A_2, \ldots, A_n)$:

$$\Phi(A_1, A_2, \ldots, A_n) = F_1(A_1, A_2, \ldots, A_n) \& F_2(A_1,$$
$$A_2, \ldots, A_n) \& \ldots \& F_m(A_1, A_2, \ldots, A_n).$$

Согласно ранее доказанному, любая логическая формула может быть представлена в совершенной конъюнктивной нормальной форме.

В соответствии с определением $U(A_1, A_2, \ldots, A_n)$ будет логическим следствием посылки $\Phi(A_1, A_2, \ldots, A_n)$, если будет доказана истинность импликации

$$\Phi(A_1, A_2, \ldots, A_n) \to U(A_1, A_2, \ldots, A_n).$$

Допустим, что в $U(A_1, A_2, \ldots, A_n)$ есть конституента, например $A_1 \vee \overline{A}_2 \vee \overline{A}_3 \vee A_4 \vee \ldots \vee A_n$ (с отрицанием только A_2 и A_3), которая в совершенной конъюнктивной нормальной форме формулы $\Phi(A_1, A_2, \ldots, A_n)$ отсутствует.

Подберем набор M, на котором выбранная нами конституента ложна. Таким набором будет $M = (A_1 — Л, A_2 — И, A_3 — И, A_i)$, где A_i, $i = 4, \ldots, n$, — ложны.

Среди всех 2^n конституент будет одна такая, в которой все элементарные высказывания A_i без отрицаний, и одна, в которой все A_i с отрицаниями. Обе эти конституенты на наборе M будут истинны.

C_n^1 конституент будут иметь по одному высказыванию с отрицанием. Все конституенты этой группы на наборе M будут истинны, так как без отрицания будут либо A_2, либо A_3, либо оба вместе.

C_n^2 конституент будут иметь по два элементарных высказывания с отрицанием. Среди них окажется и выбранная нами конституента $A_1 \vee \overline{A}_2 \vee \overline{A}_3 \vee A_4 \vee \ldots \vee A_n$, которая на наборе M принимает значение «ложно». Но все остальные конституенты этой группы на наборе N будут истинны, так как хотя бы один знак отрицания разместится над элементарным высказыванием, отличным от A_2 и A_3. А этого достаточно, чтобы вся конституента приняла значение «истинно».

Остальные группы с конституентами, имеющими три, четыре и т. д. отрицания на наборе M, будут истинны, так как среди отрицаемых высказываний найдется хотя бы одно отличное от A_2 и A_3, что даст конституенте истинное значение.

Таким образом, допущение, что конституента $A_1 \vee \overline{A}_2 \vee \vee \overline{A}_3 \vee A_4 \vee \ldots \vee A_n$ не входит в совершенную конъюнктивную нормальную форму формулы $\Phi(A_1, A_2, \ldots, A_n)$, создает условие, в котором посылка $\Phi(A_1, \ldots, A_n)$ истинна, а следствие $U(A_1, A_2, \ldots, A_n) = A_1 \vee \overline{A}_2 \vee \overline{A}_3 \vee \ldots \vee A_n$ ложно. Следовательно, вся импликация $\Phi(A_1, A_2, \ldots, A_n) \to U(A_1, A_2, \ldots, A_n)$ не является тождественно истинной.

Достаточность. Пусть все конституенты, составляющие следствие $U(A_1, A_2, \ldots, A_n)$, входят составными частями в совершенную конъюнктивную нормальную форму формулы $\Phi(A_1, A_2, \ldots, A_n)$.

Если все конституенты, входящие в $U(A_1, A_2, \ldots, A_n)$, истинны, то истинна и $U(A_1, A_2, \ldots, A_n)$, вместе с этим истинна и импликация

$$\Phi(A_1, A_2, \ldots, A_n) \to U(A_1, A_2, \ldots, A_n)$$

независимо от значения посылки $\Phi(A_1, A_2, \ldots, A_n)$.

Если в $U(A_1, A_2, \ldots, A_n)$ найдется хотя бы одна конституента $A_1 \vee A_2 \vee \ldots \vee A_n$, которая на некотором наборе принимает значение «ложно», то такая же конституента есть и в совершенной конъюнктивной нормальной форме посылки $\Phi(A_1, A_2, \ldots, A_n)$. В силу этого в импликации $\Phi(A_1, A_2, \ldots, A_n) \to U(A_1, A_2, \ldots, A_n)$ посылка и следствие примут значение «ложно», а импликация — «истинно» на любом наборе. Это обеспечивает тождественную истинность импликации, что и требовалось доказать.

На основе доказанных предположений сформулируем следующие практические приемы для получения двоичных функций, вытекающих как логические следствия из данных посылок.

1. Из данных посылок $F_1(A_1, A_2, \ldots, A_n)$; ...; $F_m(A_1, A_2, \ldots, A_n)$ с помощью объединения их действием конъюнкции строится функция-посылка $\Phi(A_1, A_2, \ldots, A_n)$.

2. Посылка $\Phi(A_1, A_2, \ldots, A_n)$ приводится к совершенной конъюнктивной нормальной форме.

3. Из полученной совершенной конъюнктивной нормальной формы берется любое количество различных входящих в нее конституент, которые и являются двоичными функциями, вытекающими как логические следствия из посылок $F_1(A_1, A_2, \ldots, A_n)$; ...; $F_m(A_1, A_2, \ldots, A_n)$.

Полученная формула может быть упрощена с помощью преобразований по формулам (1.5) — (1.29).

Пример 1. Найти все логические следствия из посылок

$$F_1(A) = A_1;\ F_2(A, B) = A \to B.$$

Решение. В соответствии с рекомендацией построим функцию-посылку

$$\Phi(A, B) = A \& (A \to B)$$

и приведем ее к совершенной конъюнктивной нормальной форме

$$A \& (A \to B) = A \& (\bar{A} \vee B) = (A \vee B \& \bar{B}) \& (\bar{A} \vee B) =$$
$$= (A \vee B) \& (A \vee \bar{B}) \& (\bar{A} \vee B).$$

Логическими следствиями посылки $\Phi(A, B)$ будут:

1) $A \vee B$;
2) $A \vee \bar{B}$;
3) $\bar{A} \vee B$:
4) $(A \vee B) \& (A \vee \bar{B})$:
5) $(A \vee B) \& (\bar{A} \vee B)$;
6) $(A \vee \bar{B}) \& (\bar{A} \vee B)$:
7) $(A \vee B) \& (A \vee \bar{B}) \& (\bar{A} \vee B)$.

Полученные следствия можно упростить, например

$$(A \vee B) \& (\bar{A} \vee B) = B \& (A \vee \bar{A}) = B.$$

Рассмотрим вопрос о числе двоичных функций, выводимых как логические следствия из m посылок.

Пусть наибольшее число элементарных высказываний A_i, входящих в функции-посылки $F_i(A_1, A_2, \ldots, A_n)$, будет n. Тогда посылка $\Phi(A_1, A_2, \ldots, A_n)$ в совершенной конъюнктивной нормальной форме может иметь не больше 2^n конституент. Если в качестве двоичной функции брать каждую конституенту по отдельности, то таких следствий будет $C_{2^n}^1 = 2^n$. Объединяя по две различные конституенты, получим $C_{2^n}^2$ функций-следствий. По три — $C_{2^n}^3$ и т. д.

Общее число функций-следствий:

$$N = C_{2^n}^0 + C_{2^n}^1 + \ldots + C_{2^n}^{2^n} = 2^{2^n}.$$

В этой сумме слагаемые $C_{2^n}^0$ и $C_{2^n}^{2^n}$ соответствуют случаям, когда в качестве функции-следствия берется сама функция-посылка $\Phi(A_1, A_2, \ldots, A_n)$ и ее отрицание.

Таким образом, если имеем два элементарных высказывания A и B, то из них могут быть построены $2^{2^2} = 16$ двоичных функций.

Если в качестве посылок рассмотреть три высказывания A, B, C, то получим $N = 2^{2^3}$; $N = 2^8 = 256$ двоичных функций.

В частном случае, когда посылки будут представлены в виде каких-либо логических связей, в силу которых в функцию-посылку $\Phi(A_1, A_2, \ldots, A_n)$ войдут не все из возможных конституент, то число функций, полученных как логические следствия, определится формулой $N^* = 2^m$, где m — число конституент, участвующих в посылке $\Phi(A_1, A_2, \ldots, A_n)$.

Для технических приложений особо важное значение приобретает рассмотрение всех двоичных функций, вытекающих как логические следствия из двух элементарных высказываний A и B. Этих функций будет $2^{2^2} = 16$. Найдем все их и представим в наиболее простом, компактном виде. Из двух высказываний A и B строятся $2^2 = 4$ конституенты $A \vee B$; $\overline{A} \vee B$; $A \vee \overline{B}$; $\overline{A} \vee \overline{B}$. Рассматривая их как посылку

$$\Phi(A, B) = (A \vee B) \& (\overline{A} \vee B) \& (A \vee \overline{B}) \& (\overline{A} \vee \overline{B}),$$

в качестве следствий получим следующие двоичные функции; по одной конституенте:

$$A \vee B; \ \overline{A} \vee B; \ A \vee \overline{B}; \ \overline{A} \vee \overline{B};$$

комбинации по две конституенты:

$$(A \vee B)\&(\bar{A} \vee B); \qquad (A \vee B)\&(\bar{A} \vee \bar{B});$$

$$(A \vee B)\&(A \vee \bar{B}); \qquad (\bar{A} \vee B)\&(\bar{A} \vee \bar{B});$$

$$(\bar{A} \vee B)\&(A \vee \bar{B}); \qquad (A \vee \bar{B})\&(\bar{A} \vee \bar{B});$$

комбинации по три конституенты:

$$(A \vee B)\&(\bar{A} \vee B)\&(A \vee \bar{B}); \qquad (A \vee B)\&(\bar{A} \vee B)\&(\bar{A} \vee \bar{B});$$

$$(A \vee B)\&(A \vee \bar{B})\&(\bar{A} \vee \bar{B}); \qquad (\bar{A} \vee B)\&(A \vee \bar{B})\&(\bar{A} \vee \bar{B});$$

комбинация всех четырех конституент и ее отрицание:

$$(A \vee B)\&(\bar{A} \vee B)\&(A \vee \bar{B})\&(\bar{A} \vee \bar{B});$$

$$\overline{(A \vee B)\&(\bar{A} \vee B)\&(A \vee \bar{B})\&(\bar{A} \vee \bar{B})}.$$

Рассмотрим полученные функции:

1) $A \vee B$ — дизъюнкция элементарных высказываний;

2) $\bar{A} \vee B = A \rightarrow B$ — импликация;

3) $A \vee \bar{B} = \bar{B} \vee A = B \rightarrow A$ — обратная импликация;

4) $\bar{A} \vee \bar{B} = \overline{A \& B}$ — отрицание конъюнкции (штрих Шеффера);

5) $(A \vee B)\&(\bar{A} \vee B) = B \vee A\&\bar{A} = B \vee Л = B$ — сложное высказывание, не зависящее от A;

6) $(A \vee B)\&(A \vee \bar{B}) = A \vee B\&\bar{B} = A \vee Л = A$ — сложное высказывание, не зависящее от B;

7) $(\bar{A} \vee B)\&(A \vee \bar{B}) = (A \rightarrow B)\&(B \rightarrow A) = A \sim B$ — высказывания A и B эквивалентны;

8) $(A \vee B)\&(\bar{A} \vee \bar{B}) = A\&\bar{A} \vee A\&\bar{B} \vee B\&\bar{A} \vee B\&\bar{B} = Л \vee A\&\bar{B} \vee B\&\bar{A} \vee Л$ или $A\&\bar{B} \vee \bar{A}\&B = \overline{\overline{A\&\bar{B}}\&\overline{\bar{A}\&B}} = \overline{(\bar{A} \vee B)\&(A \vee \bar{B})} = \overline{(A \rightarrow B)\&(B \rightarrow A)} = \overline{A \sim B}$ — противоположность высказываний A и B (отрицание их эквивалентности);

9) $(\bar{A} \vee B)\&(\bar{A} \vee \bar{B}) = \bar{A} \vee B\&\bar{B} = \bar{A} \vee Л = \bar{A}$ — отрицание высказывания A (инверсия A);

10) $(A \vee \bar{B})\&(\bar{A} \vee \bar{B}) = \bar{B} \vee A\&\bar{A} = \bar{B} \vee Л = \bar{B}$ — отрицание высказывания B (инверсия B);

(В преобразованиях формул, состоящих из трех конституент, используем полученные результаты.)

11) $(A \vee B) \& (\overline{A} \vee B) \& (A \vee \overline{B}) = B \& (A \vee \overline{B}) = B \& A \vee B \& \overline{B} = B \& A \vee Л = B \& A$ — конъюнкция элементарных высказываний;

12) $(A \vee B) \& (A \vee \overline{B}) \& (\overline{A} \vee \overline{B}) = A \& (\overline{A} \vee \overline{B}) = A \& \overline{A} \vee A \& \overline{B} = Л \vee A \& \overline{B} = A \& \overline{B}$ или (если использовать отрицание) $A \& \overline{B} = \overline{\overline{A} \vee B} = \overline{A \rightarrow B}$ — отрицание импликации (функция запрета A);

13) $(A \vee B) \& (\overline{A} \vee B) \& (\overline{A} \vee \overline{B}) = \overline{B \rightarrow A}$ — отрицание обратной импликации (преобразования те же самые, что и в предыдущей формуле) (функция запрета B);

14) $(\overline{A} \vee B) \& (A \vee \overline{B}) \& (\overline{A} \vee \overline{B}) = (\overline{A} \vee B) \& (\overline{A} \vee \overline{B}) \& (A \vee \overline{B}) = [\overline{A} \vee B \& \overline{B}] \& (A \vee \overline{B}) = (\overline{A} \vee Л) \& (A \vee \overline{B}) = \overline{A} \& (A \vee \overline{B}) = \overline{A} \& A \vee \overline{B} = \overline{A} \& A \vee \overline{A} \& \overline{B} = Л \vee \overline{A} \& \overline{B} = \overline{A \vee B}$ — отрицание дизъюнкции (стрелка Пирса);

15) $(A \vee B) \& (\overline{A} \vee B) \& (A \vee \overline{B}) \& (\overline{A} \vee \overline{B}) = [B \vee A \& \overline{A}] \& (\overline{B} \vee A \& \overline{A}) = (B \vee Л) \& (\overline{B} \vee Л) = B \& \overline{B} = Л$ — тождественно ложное высказывание;

16) $(A \vee B) \& (\overline{A} \vee B) \& (A \vee \overline{B}) \& (\overline{A} \vee \overline{B}) = И$ — отрицание тождественно ложного высказывания тождественно истинно.

В скобках указаны наименования двоичных функций, употребляемые в технических приложениях.

Итак, получены все шестнадцать двоичных функций, которые могут давать два элементарных высказывания в различных логических связях. Заметим, что из всех шестнадцати формул восемь (формулы 4), 8) — 10), 12) — 14), 16)) представляют отрицания других восьми формул (соответственно формул 11), 7), 6), 5), 2), 3), 1) 15)). Можно доказать, что половина 2^n функций, получающихся из n элементарных высказываний-аргументов, будет отрицанием другой половины этих функций.

Двоичные функции двух переменных в прикладных вопросах имеют особо важное значение. Некоторые из них получили специальные наименования.

Сведем эти функции в таблицу (табл. 1.2) с указанием их обозначения, названия и представления в виде совершенных нормальных форм (знак конъюнкции опущен).

Таблица 1.2

	Название функции	Обозначение функции	Совершенная дизъюнктивная нормальная форма	Совершенная конъюнктивная нормальная форма	Сокращенная дизъюнктивная нормальная форма
$f_0(x, y)$	Тождественно ложное высказывание	0	0	$(x \vee y)(x \vee \bar{y})(\bar{x} \vee y)(\bar{x} \vee \bar{y})$	0
$f_1(x, y)$	Конъюнкция элементарных высказываний	xy	xy	$(x \vee y)(x \vee \bar{y})(\bar{x} \vee y)$	xy
$f_2(x, y)$	Отрицание импликации	—	$x\bar{y}$	$(x \vee y)(x \vee \bar{y})(\bar{x} \vee \bar{y})$	$x\bar{y}$
$f_3(x, y)$	Сложное высказывание, не зависящее от y	x	$x\bar{y} \vee xy$	$(x \vee y)(x \vee \bar{y})$	x
$f_4(x, y)$	Отрицание обратной импликации	—	$\bar{x}y$	$(x \vee y)(\bar{x} \vee y)(\bar{x} \vee \bar{y})$	$\bar{x}y$
$f_5(x, y)$	Сложное высказывание, не зависящее от x	y	$\bar{x}y \vee xy$	$(x \vee y)(\bar{x} \vee y)$	y
$f_6(x, y)$	Отрицание эквивалентности	$x + y$	$\bar{x}y \vee x\bar{y}$	$(x \vee y)(\bar{x} \vee \bar{y})$	$\bar{x}y \vee x\bar{y}$
$f_7(x, y)$	Дизъюнкция элементарных высказываний	$x \vee y$	$\bar{x}y \vee x\bar{y} \vee xy$	$(x \vee y)$	$x \vee y$
$f_8(x, y)$	Отрицание дизъюнкции	$x \downarrow y$	$\bar{x}\bar{y}$	$(x \vee \bar{y})(\bar{x} \vee y)(\bar{x} \vee \bar{y})$	$\bar{x}\bar{y}$
$f_9(x, y)$	Эквивалентность высказываний	$x \sim y$	$\bar{x}\bar{y} \vee xy$	$(x \vee \bar{y})(\bar{x} \vee y)$	$\bar{x}\bar{y} \vee xy$
$f_{10}(x, y)$	Отрицание высказывания y	$\bar{y}$	$\bar{x}\bar{y} \vee x\bar{y}$	$(x \vee \bar{y})(\bar{x} \vee \bar{y})$	$\bar{y}$
$f_{11}(x, y)$	Обратная импликация	$y \rightarrow x$	$\bar{x}\bar{y} \vee x\bar{y} \vee xy$	$x \vee \bar{y}$	$x \vee \bar{y}$
$f_{12}(x, y)$	Отрицание высказывания x	$\bar{x}$	$\bar{x}\bar{y} \vee \bar{x}y$	$(\bar{x} \vee y)(\bar{x} \vee \bar{y})$	$\bar{x}$
$f_{13}(x, y)$	Импликация	$x \rightarrow y$	$\bar{x}\bar{y} \vee \bar{x}y \vee xy$	$\bar{x} \vee y$	$\bar{x} \vee y$
$f_{14}(x, y)$	Отрицание конъюнкции	x / y	$\bar{x}\bar{y} \vee \bar{x}y \vee x\bar{y}$	$\bar{x} \vee \bar{y}$	$\bar{x} \vee \bar{y}$
$f_{15}(x, y)$	Тождественно истинное высказывание	1	$\bar{x}\bar{y} \vee x\bar{y} \vee \bar{x}y \vee xy$	1	1

Глава 2. ЭЛЕМЕНТЫ ИСЧИСЛЕНИЯ ПРЕДИКАТОВ

§ 2. 1. ПОНЯТИЕ ПРЕДИКАТА

Исчисление предикатов является очередным шагом в развитии математической логики. Известное из предыдущего изложения исчисление высказываний входит в логику предикатов составной частью.

При построении исчисления высказываний мы уславливались рассматривать элементарное высказывание как единое нерасчленяемое целое, относительно которого можно однозначно ответить на вопрос значения его истинности.

Но, оказывается, этого недостаточно не только для формализованного описания какой-либо дедуктивной теории в целом, но даже для описания отдельных ее частей.

Дело в том, что для конкретного выражения мысли необходимо характеризовать предмет с точки зрения его качества или отношения к другим предметам.

Пусть имеется некоторое множество предметов, например множество действительных чисел. Обозначим его через $M = \{a, b, \ldots, m\}$, где $a, b, \ldots, m$ — элементы множества. Возможны различные высказывания, характеризующие какие-то качества каждого элемента. Например, «a — целое», «b — иррациональное». Высказывание о нескольких предметах выражает отношения между ними. Например, «a кратно b», «$a + b + c = 0$».

Свойства и отношения между предметами называют *предикатами*.

Раздел математической логики, в котором наряду с элементарными высказываниями рассматривают высказывания, содержащие предикаты, называют *исчислением предикатов*. Познакомимся с его основными понятиями.

Совокупность предметов, взятых для изучения какого-то определенного вопроса, называют *множеством* или *полем*. Предметы, принадлежащие данной совокупности, обычно обозначают малыми буквами латинского алфавита $a, b, c, \ldots, x, y, z$ с индексами или без них.

Всю совокупность предметов в целом обозначают буквой M или символом $\{a_i\}$, $i = 1, \ldots, n$. Говорят: «Рассмотрим явление на множестве M», «Возьмем предмет a_i из множества M», «Предмет a_i принадлежит множеству M». Иногда в таких оборотах речи слово «множество» заменяют словом «совокупность предметов».

Предикаты обозначают большими буквами латинского алфавита и скобками с таким количеством «пустых мест», к скольким предметам относится данный предикат. Например, $P(\ )$, $Q(\ )$, $F(\ ,\ )$, $R(,\ ,\ \ldots,\ )$. Подставляя на пустое место символ данного предмета, получим о нем определенное высказывание.

Так $P(a)$ обозначает высказывание о предмете a; $F(a,\ b)$ — высказывание об отношении предметов a и b между собой; $R(a_1,\ a_2,\ \ldots,\ a_n)$ — высказывание об отношении между предметами $a_1,\ a_2,\ \ldots,\ a_n$.

Такие высказывания могут быть истинными и ложными. В дальнейшем предикаты будем рассматривать с точки зрения их истинности. Для обозначения используем известные из предыдущего символы И и Л.

В отличие от исчисления высказываний значение истинности предиката будем определять по отношению к субъекту или группе субъектов. Пусть M — заданная совокупность предметов, а x — любой предмет из этой совокупности. Тогда выражение $F(x)$ обозначает какое-то высказывание о неопределенном предмете x. В силу этого не представляется возможным определить значение истинности предиката $F(x)$ и выражение $F(x)$ оказывается неопределенным. Но если для x указан определенный предмет из совокупности M, например $x=a$, то выражение $F(a)$ становится вполне определенным, так как для конкретного субъекта a можно определить, истинно высказывание $F(a)$ или ложно. Например, пусть M — совокупность плоских четырехугольников. Имеем высказывание $F(x)$: «В x диагонали взаимно перпендикулярны». О значении истинности такого высказывания ничего определенного сказать нельзя. Но если вместо x взять конкретный четырехугольник a, например ромб, то высказывание «В ромбе диагонали взаимно перпендикулярны» — определенное, истинное высказывание.

Так как предикат представляет собой высказывание, принимающее одно из двух значений И или Л, то выражение $F(x)$ означает, что каждому предмету x из совокупности M поставлено в соответствие определенное значение истинности. Иначе говоря, $F(x)$ представляет собой функцию, заданную на совокупности объектов M. Таким же образом неопределенные высказывания $Q(x,\ y)$, $R(a_1,\ a_2,\ \ldots,\ a_n)$ представляют собой функции двух и вообще n переменных. При этом $x,\ y,\ a_1,\ \ldots,\ a_n$ могут принимать любую комбинацию значений из соответствующей совокупности значений M, а сами функции $Q(x,\ y)$, $R(a_1,\ a_2,\ \ldots,\ a_n)$ двузначны, т. е. принимают значения И и Л.

По числу субъектов предикаты называют: $F(x)$ — *одноместным*, $Q(x,\ y)$ — *двуместным* и, наконец, $R(x_1,\ x_2,\ \ldots,\ x_n)$ — *n-местным*.

Принято определенные элементы совокупности M, на которой задана логическая функция, обозначать начальными буквами латинского алфавита a; b; c; a_1, a_2, ..., a_n и называть *индивидуальными предметами* или *предметными постоянными*; неопределенные элементы — обозначать буквами x; y; z; x_1, x_2, ..., x_n и называть *предметными переменными*.

Для элементарных высказываний остаются те же обозначения A; B; A_1, ..., A_n, которые были в исчислении высказываний. Если будет рассматриваться элементарное высказывание или предикат с определенным значением истинности, то этот факт отметим значком * над символом: A^*, $F^*(x, y)$.

Отдельные высказывания и предикаты называют *элементарными формулами*. Из элементарных формул с помощью логических связей конъюнкции, дизъюнкции, импликации, отрицания высказывания строятся сложные формулы.

Логика предикатов в сочетании с исчислением высказываний позволяет представлять в формализованном виде весьма широкий круг понятий, образов и различных высказываний, характеризующих свойства предметов и отношений между ними.

Рассмотрим несколько примеров.

Пример 1. Пусть требуется формально изобразить свойство делимости целых чисел на 3.

Решение. В качестве совокупности M возьмем натуральные числа. Введем предикат «делится» и обозначим его символом $P(\cdot, \ )$. Тогда элементарная формула $P(x, y)$ символизирует высказывание «Число x делится на y», а $P(x, 3)$ — «Число x делится на 3».

При подстановке вместо x различных натуральных чисел формула $P(x, 3)$ может быть истинной и ложной. Одновременно с выделением истинных значений $P(x, 3)$ из M выделяются числа, кратные 3.

Пример 2. Записать в виде формулы теорему «Если прямые a и b параллельны и a перпендикулярна плоскости θ, то и b перпендикулярна θ».

Решение. Введем предикаты: $H(\ ,\ )$ — быть параллельными; $R(\ ,\ )$ — быть перпендикулярными.

Тогда формулировка теоремы в символах будет

$$[H(a, b) \& R(a, \theta)] \to R(b, \theta).$$

Пример 3. Выразить формулой содержание высказывания «Если студент — отличник учебы и активен в общественной жизни, то он получает повышенную стипендию».

Решение. Введем предикаты: $A(\ )$ — быть отличником учебы; $B(\ )$ — быть активистом; $C(\ )$ — получать повышенную стипендию.

Требуемая формула будет

$$A(x) \& B(x) \to C(x).$$

При подстановке вместо предметной переменной x фамилии определенного студента вся формула принимает одно из значений И или Л.

Формула $[\bar{A}(x) \vee \bar{B}(x)] \to \bar{C}(x)$, рассматриваемая на той же совокупности объектов $\{x_i\}$ и с тем же смыслом предикатов, несет в себе смысл следующего высказывания: «Если студент x не отличник учебы или не является активистом, то он не получает повышенной стипендии».

Пример 4. Записать на языке символов выражение «Если x и y положительны и $x > y$, то $\frac{xy}{y-x}$ отрицательна».

Решение. В математике для выражения таких предикатов, как «быть параллельным», «быть перпендикулярным», «быть равным», «быть больше» и некоторых других, употребляются известные символы $\parallel$, $\perp$, $=$, $>$. В таких символах данное выражение запишется формулой

$$[(x > 0) \& (y > 0) \& (x > y)] \to \left(\frac{xy}{y-x} < 0\right).$$

§ 2. 2. КВАНТОРЫ

В логике предикатов, кроме операций, известных из исчисления высказываний, используются еще две, вводимые с помощью символов $(\forall x)$ и $(\exists x)$ и называемые *квантором общности* и *квантором существования* соответственно.

Квантор общности $(\forall x)$ читается так: для всякого x. Смысл этого квантора следующий.

Пусть на множестве M заданы предметная переменная x и одноместный предикат $P(x)$.

Тогда под выражением

$$(\forall x) P(x)$$

мы понимаем высказывание истинное, если $P(x)$ истинно для всех x из множества M.

Если $P(x)$ истинно не для всех x из множества M, то выражение $(\forall x)\ P(x)$ ложно. Таким образом, $(\forall x)\ P(x)$ не функция, а определенное высказывание, имеющее конкретное значение истинности в зависимости от того, для всех или не для всех значений x из множества M $P(x)$ истинно. Следовательно, $(\forall x) P(x)$ от x не зависит.

Пусть, например, M — множество действительных чисел. На M заданы две логические функции

$$P_1(x) = x + 2 = 5;$$

$$P_2(x) = x^2 > 0.$$

Выражение $(\forall x) P_1(x)$ ложно, так как $x + 2 = 5$ истинно при единственном $x = 3$, а не для всех действительных чисел.

Выражение $(\forall x) P_2(x)$ — истинно, так как $x^2 > 0$ истинно для всех действительных чисел.

Квантор существования по смыслу противоположен квантору общности. Пусть x и $P(x)$ заданы на множестве M. Тогда под выражением

$$(\exists x)\,P(x)$$

мы понимаем высказывание, истинное, если на M существует хотя бы одно значение x, при котором $P(x)$ — истинно, и ложное, если такого значения нет.

Выражение $(\exists x)P(x)$ есть высказывание и от x не зависит (так же, как и $(\forall x)P(x)$).

Кванторы $(\forall x)$ и $(\exists x)$ называют *взаимно двойственными*.

Если перед логической формулой $P(x)$ стоит квантор $(\forall x)$ или $(\exists x)$, то переменная x, расположенная в $P(x)$, называется *связанной*.

Если в логической формуле, кроме переменной x, есть еще и другие (например, $P(x, y, z)$), то переменные, не входящие в квантор, называются *свободными* переменными.

Каждый квантор имеет свою область действия, которая указывается с помощью скобок. Область действия может относиться ко всей формуле или только к ее части.

Рассмотрим формулу

$$(\forall x)\,[(x \neq 0)\,\&\,(y \neq 0) \to (\exists z)\,(x = yz)],$$

словесная формулировка которой будет: «Для всякого числа x, если $x \neq 0$ и $y \neq 0$, существует число z, такое, что $x = yz$». Здесь x — связанная переменная; область действия квантора $(\forall x)$ — вся формула; z — связанная переменная; область действия квантора $(\exists z)$ — часть формулы: $x = yz$; переменная y свободна.

Число кванторов, стоящих перед логической функцией, не ограничивается.

Наличие в формуле не связанной кванторами свободной переменной превращает ее из высказывания в логическую функцию.

Преобразование формул в исчислении предикатов опирается на понятие равносильности формул.

Если две формулы логики предикатов U и W, заданные на совокупности M, принимают одинаковые значения истинности И или Л при любой замене в обеих формулах переменных предикатов, высказываний и предметных переменных постоянными значениями этих же величин, то формулы U и W равносильны. Обозначают это символом $U \equiv W$.

Все формулы равносильных преобразований, имевших место в исчислении высказываний, переносятся в исчисление предикатов с заменой в них элементарных высказываний на предикаты. Это утверждение распространяется и на тот случай, когда вся формула или отдельные ее части связаны кванторами. Например,

$$(\exists x)[U(x)\to(\forall y)W(y)]=(\exists x)[\overline{U(x)}\vee(\forall y)W(y)];$$

$$(\forall x)F(x)\to[P(z)\to(\forall x)R(x)]=\overline{(\forall x)F(x)}\vee(\overline{P}(z)\vee(\forall x)R(x)).$$

Специфическими для исчисления предикатов являются формулы, определяющие действия с кванторами и отрицанием. Для одноименных кванторов имеют место равносильности

$$(\forall x)(\forall y)P(x,\ y)\equiv(\forall y)(\forall x)P(x,\ y);$$

$$(\exists x)(\exists y)P(x,\ y)\equiv(\exists y)(\exists x)P(x,\ y).$$

Перестановка местами разноименных кванторов недопустима.

Рассмотрим действие отрицания кванторов. Имеют место следующие две формулы:

$$\overline{(\forall x)U(x)}\equiv(\exists x)\overline{U}(x);$$

$$\overline{(\exists x)U(x)}\equiv(\forall x)\overline{U}(x),$$

т. е. знак отрицания, стоящий над квантором, переносится на предикат, а квантор меняется на двойственный. Не приводя формальных доказательств, проиллюстрируем справедливость формул логическими рассуждениями.

1. Смысл формулы $\overline{(\forall x)P(x)}$ выражается словами «Неверно, что для всех x истинно $P(x)$». Следовательно, есть такие x, для которых $P(x)$ ложно. Эту мысль и запишем формулой $(\exists x)\overline{P}(x)$, в которой ложное значение $P(x)$ заменено истинным $\overline{P}(x)$.

Отсюда $\overline{(\forall x)P(x)}\equiv(\exists x)\overline{P}(x)$.

2. Смысл формулы $\overline{(\exists x)P(x)}$ выражается словами «Неверно, что существует x, для которого $P(x)$ истинна», иначе говоря «Для всех $x\,P(x)$ ложна».

Последнее высказывание запишется формулой $(\forall x)\overline{P}(x)$, где ложное значение $P(x)$ заменено истинным $\overline{P}(x)$.

Следовательно,

$$\overline{(\exists x)P(x)}\equiv(\forall x)\overline{P}(x).$$

Эти формулы в сочетании с формулами равносильностей в исчислении высказываний позволяют представлять формулы исчисления предикатов в приведенной форме.

Формулу в исчислении предикатов, в которой логические связи выражены операциями &, $\vee$, —, а знак отрицания отнесен только к элементарным предикатам и высказываниям, называют *приведенной формой* данной формулы.

Можно доказать, что каждую формулу исчисления предикатов можно представить в приведенной форме.

Пример 1. Найти приведенную форму формулы

$$\overline{(\exists x)\,[P(x) \to (\forall y)\,R(y)]}.$$

Решение. Используя формулу (1.21), заменим знак $\to$ под общим знаком отрицания:

$$\overline{(\exists x)\,[\overline{P}(x) \vee (\forall y)\,R(y)]}.$$

Применив формулу отрицания квантора существования, получим

$$(\forall x)\,\overline{\overline{P}(x) \vee (\forall y)\,R(y)}.$$

Преобразуем выражение, содержащее знак отрицания над предикатом:

$$(\forall x)\,[\overline{\overline{P(x)}}\,\&\,\overline{(\forall y)\,R(y)}] = (\forall x)\,[P(x)\,\&\,\overline{(\forall y)\,R(y)}].$$

Применив формулу отрицания квантора общности, найдем

$$(\forall x)\,[P(x)\,\&\,(\exists y)\,\overline{R(y)}].$$

Полученная формула выражена с помощью конъюнкции, а отрицание отнесено к элементарному предикату. Следовательно, имеем формулу, равносильную исходной, но представленную в приведенной форме:

$$\overline{(\exists x)\,[P(x) \to (\forall y)\,R(y)]} \equiv (\forall x)\,[P(x)\,\&\,(\exists y)\,\overline{R}(y)].$$

Пример 2. Найти отрицание формулы

$$(\forall x)\,[A(x) \to B(x)]\,\&\,(\exists x)\,[P(x)\,\&\,\overline{H}(x)]$$

и представить его в приведенной форме.

Решение.

$$\overline{(\forall x)\,[A(x) \to B(x)]\,\&\,(\exists x)\,[P(x)\,\&\,\overline{H}(x)]} \equiv$$

$$\equiv \overline{(\forall x)\,[A(x) \to B(x)]} \vee \overline{(\exists x)\,[P(x)\,\&\,\overline{H}(x)]} \equiv$$

$$\equiv (\exists x)\,\overline{\overline{A(x)} \vee B(x)} \vee (\forall x)\,[\overline{P(x)} \vee \overline{\overline{H(x)}}] \equiv$$

$$\equiv (\exists x)\,[A(x)\,\&\,\overline{B}(x)] \vee (\forall x)\,[\overline{P(x)} \vee H(x)].$$

Приведенную форму формулы исчисления предикатов называют *нормальной*, если она или не содержит кванторов или построена так, что все кванторы вынесены из формул предикатов и размещены впереди логической формулы, не содержащей кванторов. Например, приведенная формула

$$(\forall x)\,(\forall y)\,(\exists z)\,(\exists t)\,R(x, y, z, t)$$

нормальна, если $R(x, y, z, t)$ не содержит кванторов.

Глава 3. ПРИЛОЖЕНИЕ ЭЛЕМЕНТОВ МАТЕМАТИЧЕСКОЙ ЛОГИКИ В ТЕХНИКЕ

В предыдущем изложении мы старались дать представление о математической логике как о строгой математической дисциплине, построенной аксиоматическим методом.

Аксиоматическое построение науки предполагает наличие интерпретации. Это означает, что существуют совокупности объектов, которые, замещая соответствующие символы в аксиомах, превращают их в истинные высказывания в реальном, содержательном смысле.

Мощным источником интерпретаций являлись математические науки и прежде всего теория множеств.

Вследствие этого математическая логика долгое время была наукой, обслуживающей внутренние потребности математики, утверждая истинность ее оснований и стимулируя развитие.

В § 2.1 говорилось, что в 30-х годах В. И. Шестаков указал новую интерпретацию исчисления высказываний на электрических цепях.

Смысл интерпретации следующий: элементарным высказываниям «Цепь замкнута», «Цепь разомкнута» ставят в соответствие физические явления наличия тока в цепи. Закон исключенного третьего выражается в том, что рассматриваются только установившиеся состояния. Переходные процессы момента замыкания и размыкания цепи не рассматриваются. Потому высказывание «Цепь замкнута» может иметь только два значения: быть истинным, если по цепи течет ток, и ложным в противном случае. Третьего быть не может. Логическая связь конъюнкция истолковывается как последовательное соединение цепей, дизъюнкция — как параллельное соединение.

Различные сложные цепи, состоящие из участков последовательных и параллельных соединений, могут быть изображены в виде формул математической логики, представляющих собой в таком виде двоичные функции. К последним применим аппарат алгебры логики, позволяющий анализировать и синтезировать сложнейшие электро- и радиотехнические устройства формальным путем. Идеи электротехнической интерпретации оказались настолько плодотворными, что они быстро нашли применение в теории конечных автоматов, при конструировании электронных вычислительных машин, в исследовании цепей телефонной и телеграфной связи и в ряде других технических наук.

§ 3. 1. МОДЕЛИРОВАНИЕ ЛОГИЧЕСКИХ ФУНКЦИЙ ПЕРЕКЛЮЧАТЕЛЬНЫМИ УСТРОЙСТВАМИ

Условимся рассматривать различные электротехнические и электронные конструкции как электрические цепи, состоящие из переключательных устройств и соединяющих их проводников. Такими устройствами могут быть различные контактные переключатели, реле, кнопки и бесконтактные сочленения в виде электронных ламп и полупроводников.

Роль таких устройств состоит в том, что они способны пропускать ток или прекращать его прохождение на каком-то участке цепи, создавая на этом участке определенный эффект. Для отображения различных переключателей примем символ x_i. Эффект их действия представим как двоичную функцию, у которой фиксированные значения обозначим $x_i^* = 1$ и $x_i^* = 0$.

Например,

$$x_i = \begin{cases} 1 \text{ — высокое напряжение;} \\ 0 \text{ — низкое напряжение;} \end{cases}$$

$$[x_i = \begin{cases} 1 \text{ — наличие импульса;} \\ 0 \text{ — отсутствие импульса;} \end{cases}$$

$$[x_i = \begin{cases} 1 \text{ — наличие положительного импульса;]} \\ 0 \text{ — наличие отрицательного импульса.} \end{cases}$$

Если контакт x_i находится в одном из своих фиксированных состояний, но неизвестно, в каком именно, то обозначают это символом x^*.

Обычно в состав любого электротехнического устройства входит некоторое число переключателей $x_1, x_2, \ldots, x_n$. Функционирование устройства будет зависеть от состояний замыканий и размыканий контактов, поэтому его работу можно рассматривать как двоичную функцию $F(x_1, x_2, \ldots, x_n)$ двоичных переменных $x_1, x_2, \ldots, x_n$. Если состояние замыкания и размыкания переключателей известно, то каждое x_i примет одно из двух значений 1 или 0, и вместе с тем все аргументы обратятся в набор нулей и единиц. В этом случае говорят, что функция $F(x_1, x_2, \ldots, x_n)$ рассматривается на наборе, например, $(10010101)_j$.

Рассмотрим физический смысл логических операций отрицания, конъюнкции и дизъюнкции на электрических цепях. Следуя сложившейся в технических науках системе обозначений, при написании логических формул дизъюнкцию будем обозначать символом $+$ вместо $\vee$. Знак конъюнкции будем опускать вообще, как это делается в обычной алгебре при обозначении произведения.

Отрицание. Логический смысл операции отрицания **не** состоит в том, что она присваивает высказыванию противоположное значение истинности. В приложении к цепям получим: если x_i — замкнутый контакт, то $\bar{x}_i$ — разомкнутый, и наоборот, если x_i — разомкнутый контакт, то $\bar{x}_i$ — замкнутый.

Существует несколько различных типов устройств, реализующих операцию **не.** Эффект таких устройств должен состоять в преобразовании данного действия в противоположное. Так, если на вход устройства поступает сигнал, то на выходе он должен отсутствовать, и наоборот, если на входе нет сигнала, то на выходе он должен появляться.

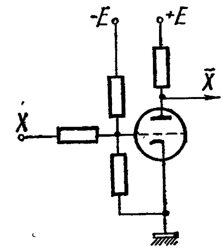

Рис. 3.1

Одним из примеров такого устройства может быть триод (рис. 3.1). Если на входе x_i сигнала нет ($x_i = 0$), то напряжения на сетке нет. Триод окажется запертым и все напряжение E окажется на выходе x_i ($\bar{x}_i = 1$). Если на вход подать напряжение $x_i = 1$, то триод отпирается и на выходе окажется сигнал $\bar{x}_i = 0$.

Можно указать несколько устройств, реализующих в электрических схемах действие отрицания. На функциональных

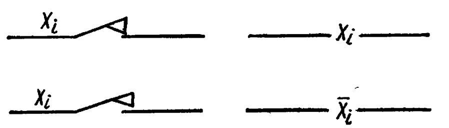

Рис. 3.2

Рис. 3.3

схемах принято их обозначать символом (рис. 3.2), который указывает только на действие элемента в схеме, а не на его устройство. Иногда в схемах логический элемент **не** изображают проще (рис. 3.3).

Конъюнкция. Согласно определению конъюнкция выражает смысл союза **и** и принимает истинное значение, если истинны все элементарные высказывания $x_1, x_2, \ldots, x_n$, входящие в нее. Рассмотрим цепь последовательного соединения нескольких проводников. Пусть в точках соединения их находятся переключатели $x_1, x_2, \ldots, x_n$. По цепи будет идти ток, если будет замкнут переключатель x_1 **и** замкнут x_2 **и... и** замкнут переключатель x_n. Достаточно размыкания хотя бы одного переключателя, как ток прекратится.

Сравнивая определение конъюнкции и физическую картину явлений замыкания и размыкания цепи последовательного соединения, замечаем совпадение смысла значений «истинно» и «Цепь замкнута» (ток течет); «ложно» и «Цепь разомкнута» (ток отсутствует). Это совпадение дает право рассматривать конъюнкцию как формальное изображение явления протекания электрического тока по цепи последовательного соединения

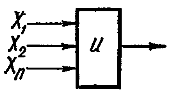

Рис. 3.4

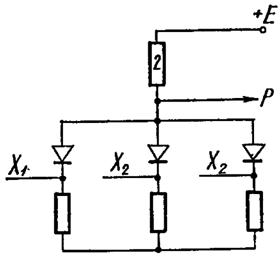

Рис. 3.5

проводников с n контактами и, наоборот, последовательную цепь рассматривать как интерпретацию операции конъюнкции. При таком подходе приобретает конкретный смысл формула

$$a\&a\&a\& \ \dots \ \&a = a. \tag{3.1}$$

Пусть a — какой-то переключатель. Левая часть формулы представляет собой последовательное соединение n одинаковых переключателей. Если a замкнут, то замкнуты все n переключателей, составляющих последовательную цепь, а вместе с этим замкнута вся цепь. Если a разомкнут, то разомкнута и вся цепь. Но ведь так же работает и один переключатель. Мы показали, что формула (3.1) изображает последовательное соединение проводников, и наоборот, последовательное соединение проводников может быть представлено формулой (3.1). Устройство, реализующее действия, соответствующие смыслу конъюнкции, должно обладать следующим свойством: в нем может быть несколько входов x_1, x_2, ..., x_n и один выход, причем на выходе будет сигнал 1 в том случае, если будут сигналы 1 на входе x_1 и на входе x_2 и... и на входе x_n. На функциональных схемах такие устройства изображают символом, показанным на рис. 3.4.

Простейшим примером устройства, реализующего конъюнкцию, может быть схема, составленная из нескольких параллельно включенных диодов с постоянным источником питания (рис. 3.5). Если на все устройства x_1, x_2 и x_3 подано напряжение, то диоды заперты и на выходе P будет все напряжение E. Если хотя бы на одном входе напряжение снять ($x_i = 0$), то соответствующий диод станет проводником и на выходе P будем наблюдать резкое падение напряжения ($P = 0$).

В технической терминологии устройства, реализующие конъюнкцию, называют *схемами совпадения* или *вентилями*.

Дизъюнкция. Рассмотрим участок цепи с параллельным разветвлением, на каждой ветви которого есть переключатель.

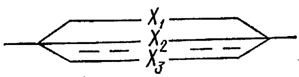

Рис. 3.6

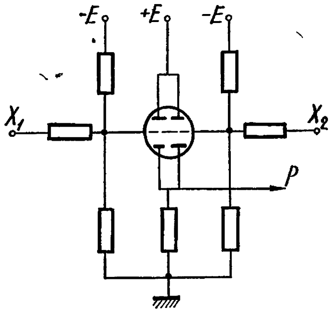

Рис. 3.7

Весь участок цепи будет проводить ток в том случае, когда окажется замкнутой хотя бы одна ветвь (рис. 3.6). Сигнал на выходе цепи будет в том случае, когда замкнута цепь **или** x_1, **или** x_2, **или**..., **или** x_n. Здесь также замечаем совпадение смысла операции дизъюнкции с законами электрического тока в цепи параллельного разветвления. В силу этого формально участки параллельного разветвления цепей можно изображать дизъюнкцией соответствующих контактов.

Существует несколько схем, реализующих логическую связь **или**. Их называют *собирательными схемами*. Рассмотрим, для примера, собирательную схему на двойном триоде (рис. 3.7).

Если на входах x_1 и x_2 сигнал 0, то лампа заперта и на выходе сигнал будет 0. Если подать напряжение на вход x_1 **или** на вход x_2, то возникает катодный ток и на выходе появится сигнал 1.

Можно построить несколько подобных устройств на различных лампах или полупроводниках. На схемах собирательные устройства изображают символом, показанным на рис. 3.8.

Кроме рассмотренных *базовых операций*, в технике используются универсальные операции: отрицание конъюнкции и отрицание дизъюнкции (на функциональных схемах их обозначают символами Г и Ш соответственно).

В соответствии с определениями смысл операции отрицание конъюнкции состоит в следующем: рассматривается логическая формула, состоящая из элементарных высказываний, объединенных операциями конъюнкции (элементарное произведение). Если все высказывания принимают одновременно значение «истинно», то формула в целом принимает значение «ложно» и, наоборот, если хотя бы одно из элементарных высказываний ложно, то формула в целом принимает значение «истинно».

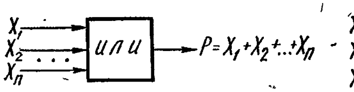

Рис. 3.8

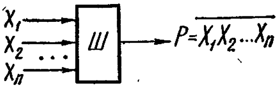

Рис. 3.9

Моделировать такую операцию может устройство, обладающее следующей способностью: если на все его входы подать сигнал, то на выходе сигнал будет отсутствовать; если хотя бы на одном входе сигнал исчезнет, то на выходе сигнал появится.

Операцию отрицания конъюнкции (штрих Шеффера) на функциональных схемах изображают символом, показанным на рис. 3.9. На рис. 3.10 изображена схема электронного устройства,

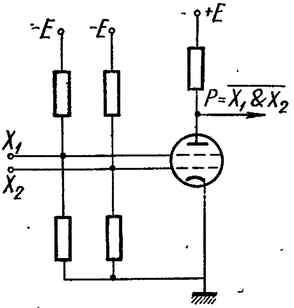

Рис. 3.10

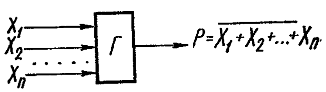

Рис. 3.11

реализующего эту операцию. Если на входах x_1 и x_2 такого устройства напряжение отсутствует ($x_1 = x_2 = 0$), то лампа заперта и на выходе будет высокий потенциал ($P = 1$). Сопротивления и сетки подобраны так, что появление напряжения на одной из них еще не отпирает лампу. Но если напряжение будет подано на оба входа ($x_1 = x_2 = 1$), то лампа отпирается и на выходе потенциал исчезнет ($P = 0$).

Аналогично строится вторая универсальная операция — отрицание дизъюнкции (стрелка Пирса). Рассмотрим ее логический смысл. Имеется элементарная сумма. Если все входящие в нее высказывания одновременно принимают значение ложно, то вся формула в целом истинна. И если хотя бы одно слагаемое истинно, то формула в целом ложна. На функциональных схемах эту операцию изображают символом, показанным на рис. 3.11. Схема одного из электронных устройств, способных реализовать операцию отрицания дизъюнкции, показана на рис. 3.12.

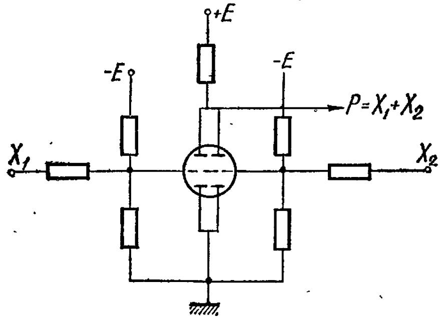

Рис. 3.12

Если на обоих входах устройства, работающего по такой схеме, сигналы отсутствуют ($x_1 = x_2 = 0$), то лампа заперта и на выходе будет высокий потенциал ($P = 1$). Если подать напряжение хотя бы на один вход, то лампа отпирается и потенциал на выходе падает ($P = 0$).

Перейдем к построению изображений некоторых формул алгебры логики на схемах цепей последовательного и параллельного соединения.

Каждая из формул соединяет знаком логического равенства два различных по виду выражения, которые принимают одинаковые значения истинности на одних и тех же наборах значений истинности элементарных высказываний.

В рассматриваемой интерпретации понятие логического равенства переходит в понятие равносильности двух схем, имеющих различную структуру, но выполняющих одинаковые функции. Формула (1.8), выражающая всегда истинное высказывание, моделируется в виде участка цепи параллельного соединения (рис. 3.13).

Такой участок цепи будет всегда проводящим, так как если контакт x замкнут, то ток потечет по верхней ветви, если разомкнут, то контакт $\bar{x}$ окажется замкнутым, и ток потечет по нижней ветви.

Так как понятие «проводимость участка цепи» эквивалентно понятию «истинно» для логической формулы, то проводимость

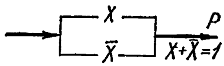

Рис. 3.13

Рис. 3.14

участка цепи во всех случаях замыкания и размыкания эквивалентна всегда истинной формуле.

Аналогично тождественно ложная формула $x \cdot \bar{x} = 0$ может быть представлена участком последовательного соединения

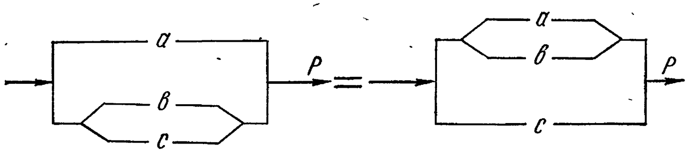

Рис. 3.15

двух контактов, работающих в таком режиме, когда при замыкании одного размыкается другой и наоборот (рис. 3.14).

Сочетательный закон дизъюнкции $a + (b + c) = (a + b) + c$ моделируется равносильностью двух цепей, изображенных на рис. 3.15.

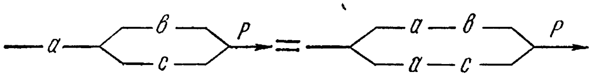

Рис. 3.16

Первый и второй распределительные законы $a(b+c)=ab+bc$; $a + bc = (a + b)(a + c)$ моделируются равносильностью схем, изображенных на рис. 3.16 и 3.17 соответственно.

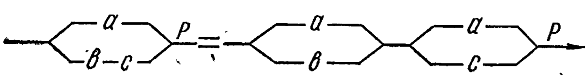

Рис. 3.17

В технических приложениях широко используются так называемые *инверсные схемы*, изображающие работу двух устройств, для которых действие одного устройства равносильно отрицанию действия другого. Такие схемы моделируют, в частности, формулы де Моргана.

Простейшие инверсные схемы изображены на рис. 3.18 и 3.19.

Рассмотренные модели дают возможность представить логические формулы в виде электрических схем. Для построения таких схем логическая формула приводится к нормальному виду.

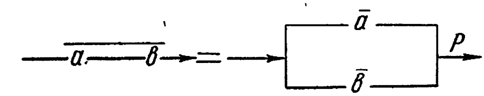

Рис. 3.18

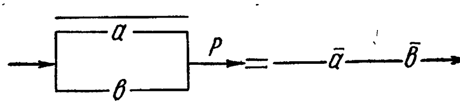

Рис. 3.19

Далее, изображая дизъюнктивную связь элементарных высказываний в виде параллельных разветвлений, а конъюнктивную— в виде последовательных соединений, получим схему, соответствующую данной формуле.

Например, для формулы

$$F(a, b, c, d, e) = [(a + b + c)\,\bar{e}] + [(\bar{a} + \bar{b} + \bar{c})\,d] + \\ + [(b + \bar{c} + \bar{e})\,\bar{a} + (\bar{c} + \bar{e})\,\bar{b} + \bar{c}e]\,d$$

такой будет схема, изображенная на рис. 3.20.

Значение истинности формулы и наличие соответствующего сигнала на выходе схемы (0 или 1) будут совпадать при взаимно однозначном соответствии значений истинности элементарных высказываний в формуле и замыкания и размыкания контактов в схеме.

Различные кибернетические устройства от небольших автоматов до крупных электронных вычислительных машин монтируются из отдельных деталей, которыми являются вентили, инверторы, собирательные схемы, универсальные элементы Г и Ш и некоторые другие.

В схемах различных технических устройств для изображения деталей, выполняющих определенные функции, используют символические обозначения, часть которых показана на рис. 3.2, 3.4, 3.8, 3.9, 3.11. В таких схемах не указывается, на каких именно элементах выполнена та или иная деталь, а показано только, какую функцию она выполняет в устройстве. Такие схемы называют *функциональными*.

Функциональные схемы могут быть представлены в виде формул алгебры логики и наоборот — формула алгебры логики может быть изображена в виде функциональной схемы определенного устройства. Если по этой схеме смонтировать реально действующее устройство, то оно в зависимости от наборов значений сигналов, поступающих на входы, будет давать на выходах строго определенные комбинации сигналов.

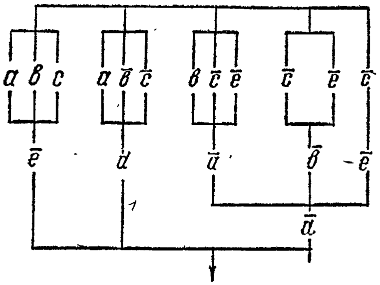

Рис. 3.20

Эти комбинации будут точно такими же, какими будут полученные значения истинности формулы алгебры логики, соответствующей данному устройству, если в нее подставить набор значений истинности элементарных высказываний, совпадающий с набором входных сигналов.

Рассмотрим простые примеры изображения формул алгебры логики в виде функциональных схем устройств и, наоборот, представления схемы данного устройства в виде формулы алгебры логики.

1) Задачу сформулируем следующим образом: требуется начертить функциональную схему устройства, реализующего формулу алгебры логики $P = x_1 + x_1x_2$.

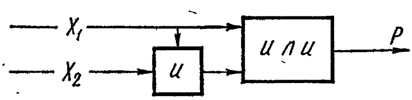

Рис. 3.21

В заданной формуле имеем дизъюнкцию двух слагаемых, одно из которых есть конъюнкция x_1x_2. В соответствии с этим в схеме будет вентиль **и** и собирательная схема **или**. На вход вентиля должны быть поданы сигналы x_1 и x_2, а в собирательную схему — сигнал x_1 и результат срабатывания вентиля x_1x_2. Такая схема изображена на рис. 3.21.

Сопоставим значение истинности формул $P = x_1 + x_1x_2$ при всевозможных наборах значений аргументов и значений сигналов на выходе схемы в зависимости от комбинаций сигналов на входах. Совпадение значений истинности формулы и значений сигналов на выходе схемы укажет на их эквивалентность.

По таблице истинности

x_1	x_2	x_1x_2	$x_1 + x_1x_2$	$x_1 = x_1 + x_1x_2$
0	0	0	0	1
0	1	0	0	1
1	0	0	1	1
1	1	1	1	1

замечаем, что значения истинности формулы и переменной x_1 совпадают. А это соответствует преобразованию $x_1 + x_1x_2 = x_1(1 + x_2) = x_1$, так как $1 + x_2 = И$.

Проследим появление сигнала 1 на выходе схемы (рис. 3.21), руководствуясь при этом физическими закономерностями работы элементов **и** и **или** при прохождении по цепи электрического тока.

Если в проводнике x_1 тока нет, то на выходе P сигнала не будет ($P = 0$), так как на обоих входах элемента **или** будут сигналы, равные 0.

Если $x_1 = 1$, то на выходе будет сигнал $P = 1$ независимо от x_2, так как для срабатывания элемента **или** достаточно сигнала хотя бы на одном входе. Таким образом, схема (рис. 3.21) работает так же, как x_1.

Этим мы показали эквивалентность формулы и схемы.

2) Изобразить функциональную схему устройства, заданного формулой

$$P(x_1, x_2, x_3, x_4) = x_1 + \overline{x_1}x_3 + \overline{x_2}x_4 + \overline{x_3x_4}.$$

В практике удобно изображать каждую переменную в виде питающей шины и замыкать ответвления от шин на соответствующие устройства.

Функциональная схема представлена на рис. 3.22. Проанализировать и показать эквивалентность схемы и формулы предлагаем читателю.

3) Рассмотрим обратную задачу. Дана функциональная схема какого-то устройства (рис. 3.23). Требуется написать соответствующую ей формулу алгебры логики.

Функциональной схеме (рис. 3.23) будет соответствовать формула алгебры логики

$$P = x_1 [x_1 x_2 x_3 + x_4 x_5] [\bar{x}_2 x_3 + \overline{x_4 x_5}].$$

Рассуждения при написании такой формулы выглядят примерно так.

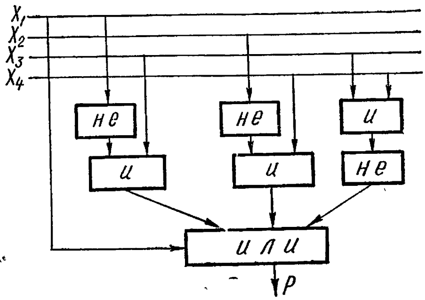

Рис. 3.22

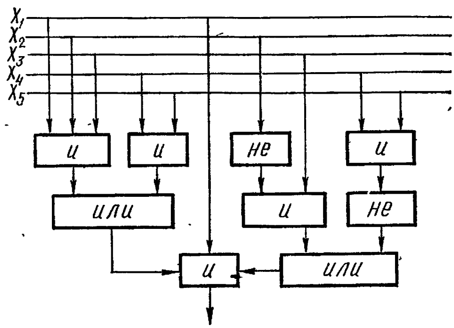

Рис. 3.23

Выход схемы исходит от элемента **и**, который имеет три входа. Следовательно, формула должна представлять конъюнкцию трех сомножителей. Один из сомножителей замыкается непосредственно с x_1 на выходной элемент **и**; он будет одним из трех сомножителей конъюнкции.

Два других сомножителя сложные. Левый в схеме вход объединяет с помощью элемента **или** два вентиля. Первый из них объединяет три входа $x_1 x_2 x_3$, второй — два $x_4 x_5$. Отсюда сомножитель: $[x_1 x_2 x_3 + x_4 x_5]$. Аналогично строится третий сомножитель.

В технических приложениях алгебра логики используется при решении различных вопросов, в том числе при решении проблем синтеза новых устройств и при анализе схем уже существующих устройств с целью их усовершенствования.

При решении обоих этих вопросов преимущество использования специального математического аппарата состоит в том, что, применяя его, можно выполнять различные эквивалентные преобразования формул, получать и оценивать различные варианты и следствия логического вывода, и на этом пути получать оптимальные функциональные схемы устройств. Наиболее интересными и практически важными являются вопросы синтеза схем кибернетических устройств с наперед заданными свойствами. В следующем параграфе описывается постановка задачи и принципы решения задач синтеза функциональных схем.

Подробное описание методов синтеза и оптимизации схем относится к специальным вопросам и излагается в соответствующих учебных курсах и литературе.

§ 3. 2. СИНТЕЗ ФУНКЦИОНАЛЬНЫХ СХЕМ ПЕРЕКЛЮЧАТЕЛЬНЫХ УСТРОЙСТВ

Как указывалось выше, переключательными устройствами мы назвали конструкции, которые автоматически или полуавтоматически выполняют различного рода операции за счет переключения контактных и бесконтактных сочленений, размещенных в соединяющих сетях. К контактным сочленениям отнесем различные реле, выключатели, контакторы и т. д.; к бесконтактным — электронные лампы, полупроводники и т. д.

Представим себе объем и виды работы, которую должен выполнить конструктор, создавая переключательное устройство, способное выполнять заданные операции. Не претендуя на полноту и точность, ограничим перечень работ следующими видами:

1) уяснение конструктором поставленной перед ним задачи, оценка состава и способов передачи входной и выходной информации;

2) разработка функциональной схемы устройства, способного выполнять заданные операции;

3) выполнение расчетов в соответствии с функциональной схемой;

4) подбор материалов и элементов конструкции с учетом требований габаритов, стоимости, технологии изготовления, эксплуатационной надежности и т. д.

В настоящем пособии мы рассмотрим только способы разработки функциональных схем устройств, т. е. те работы, где успешно используются методы математической логики.

В общей постановке задание на разработку какой-либо конструкции описывается словами обычной разговорной речи.

Для изображения содержания этого задания на формальном языке с последующими преобразованиями и схемными решениями успешно используется аппарат математической логики.

Он применяется для:

1) формального изображения в виде формулы исчисления высказываний всех процессов, которые должны функционировать в проектируемом устройстве;

2) эквивалентных преобразований полученной формулы с целью ее минимизации и получения оптимальной схемы устройства;

3) построения по логической формуле функциональных и принципиальных схем с учетом технических требований (габариты, стоимость, надежность и т. д.).

Рассмотрим по отдельности каждый из видов работ.

Начнем с рассмотрения практических приемов получения начальной логической формулы, которая отражала бы функционирование процессов в конструируемом устройстве.

В практике положительно зарекомендовал себя метод составления так называемых *таблиц срабатывания* или *таблиц задания* логических функций. Построение этих таблиц требует содержательного анализа физических закономерностей процессов, которые должны протекать в проектируемом устройстве. Систематизация анализа и символическая запись его результатов в таблице дает возможность получить соответствующую логическую формулу и этим облегчает труд разработки схемы конструкции.

Методика построения таблиц следующая. Из анализа условия задачи определяется входная ($x_1, x_2, \ldots, x_n$) и выходная ($P_1, P_2, \ldots, P_m$) информации. Далее, анализируются все комбинации, в которых может быть подана входная информация $x_1^*, x_2^*, \ldots, x_n^*$, и оценивается состояние физических явлений при каждом фиксированном значении $x_1^*, x_2^*, \ldots, x_n^*$. На основании такого анализа указывается, что должно быть на выходе устройства.

Таблица содержит $n+m$ столбцов: n столбцов для обозначения входной информации и m столбцов для обозначения выходов $P_1, P_2, \ldots, P_m$.

Рассматривая входы $x_1, x_2, \ldots, x_n$ как двоичные переменные, получим для них 2^n комбинаций фиксированных значений. В соответствии с этим в таблице должно быть 2^n строк.

При построении таблицы первые n столбцов по всем 2^n строкам заполняются наборами 0 и 1, символизирующими фиксированные значения входных сигналов x_i^*. Заполнение наборов производится формально в порядке возрастания двоичного числа от 0 до $2^n - 1$, т. е. точно так же, как это делалось при построении таблиц истинности.

Принципиальное отличие таблиц срабатывания от таблиц истинности состоит в различии областей, из которых черпаются соображения по установлению соответствий символов в n и m столбцах.

В случае таблиц срабатывания столбцы m заполняются на основе анализа физических явлений и обнаруженных закономерностей, соответствующих набору значений входных переменных, отмеченных в рассматриваемой строчке, а в таблицах истинности соответствия отображают законы логических операций математической логики.

Для примера приведена таблица срабатывания, в которой отображено функционирование некоторого устройства с четырьмя входами x_1, x_2, x_3, x_4 и одним выходом P:

x_1	x_2	x_3	x_4	P
0	0	0	0	1
0	0	0	1	0
0	0	1	0	0
0	0	1	1	1
0	1	0	0	0
0	1	0	1	1
0	1	1	0	1
0	1	1	1	0
1	0	0	0	0
1	0	0	1	1
1	0	1	0	1
1	0	1	1	0
1	1	0	0	1
1	1	0	1	0
1	1	1	0	0
1	1	1	1	1

В первых четырех столбцах с помощью символов 1 (включено) и 0 (отключено) обозначены все комбинации, которые могут возникнуть из состояний включения и выключения на входах устройства. Так как входов 4 и каждый из них — двоичная переменная, то всего комбинаций может быть $2^4 = 16$. В столбце P с помощью символов 1 и 0 изображено состояние на выходе устройства. Физический смысл символов 1 и 0 в столбце P может быть различный. Например, 1 означает включение двигателя станка, 0 — его отключение. Или 1 — поворот устройства управления автоматом на угол α; 0 — сохранение неподвижного состояния и т. д.

В соответствии с этим заполнен столбец P в таблице срабатывания. Читается таблица по строкам, например: в первой строке указано, что при отсутствии сигналов на входах, на выходе будет сигнал 1; в шестой указано, что на выходе сигнал будет равен 1, если поступят сигналы на второй и четвертый входы ($x_2 = 1$; $x_4 = 1$) при отсутствии сигналов на первом и третьем входах ($x_1 = x_3 = 0$).

Рассуждения, отображенные в таблице срабатывания, легко переводятся на язык исчисления высказываний. Входы x_i будем рассматривать как элементарные высказывания. Если высказывание x_i истинно, то $x_i = 1$; если ложно, то $x_i = 0$. Тогда набор нулей и единиц в строке таблицы можно рассматривать как набор значений истинности элементарных высказываний $x_1, x_2, \ldots, x_n$. Объединяя конъюнктивной связью элементарные высказывания в каждой строке, получим сложное высказывание, значение истинности которого на данном наборе отмечено символом 0 или 1 в столбце P. При таком истолковании каждой строки смысл таблицы в целом будет выражаться высказыванием «Устройство будет срабатывать (иметь на выходе сигнал $P = 1$), когда на вход поступят комбинации сигналов, отображенные в i-й строке **или** $(i+k)$-й строке, **или** ...» Запишем наличие сигналов на входе как x_i, отсутствие — как $\bar{x}_i$, всю строку — как конъюнкцию элементарных высказываний. Если строки с одинаковыми значениями истинности на выходе ($P = 1$), представленные в виде конъюнкций, объединить с помощью операции дизъюнкции, то получим совершенную нормальную дизъюнктивную форму такой формулы, которая определяет условие срабатывания устройства. Аналогично, объединяя строки таблицы при $P = 0$, получим формулу, определяющую отсутствие сигнала на выходе (не срабатывание).

В нашем примере логическая формула по условию срабатывания будет иметь вид:

$$P(x_1, x_2, x_3, x_4) = \bar{x}_1\bar{x}_2\bar{x}_3\bar{x}_4 \vee \bar{x}_1\bar{x}_2x_3x_4 \vee \bar{x}_1x_2\bar{x}_3x_4 \vee$$
$$\vee\ \bar{x}_1x_2x_3\bar{x}_4 \vee x_1\bar{x}_2\bar{x}_3x_4 \vee x_1\bar{x}_2x_3\bar{x}_4 \vee x_1x_2\bar{x}_3\bar{x}_4 \vee x_1x_2x_3x_4.$$

Во многих практических случаях построение таблиц срабатывания не представляет большого труда, так как известны физические условия состояния входной информации и связь ее с выходной. В нашем примере такой связью является следующее соответствие: сигнал на выходе будет $P = 1$, если на входах сигнал будет отсутствовать или подаваться на четное число входных каналов.

Таблицы срабатывания позволяют получить логические формулы, отображающие функционирование переключательного устройства, в совершенной нормальной дизъюнктивной форме. Как известно, эта форма допускает значительные упрощения.

Существует несколько способов преобразований минимизации логических формул.

Рассмотрим наиболее употребительные.

Метод Квейна. В общем случае могут быть использованы любые эквивалентные преобразования, выполняемые в алгебре логики. Примеры таких преобразований мы видели в § 2.7. На практике для минимизации логических формул оказался удобным *метод Квейна*. Исходной для преобразования является совершенная нормальная дизъюнктивная форма, над ней выполняются операции склеивания и поглощения.

Операцией поглощения называются преобразования, выполненные по формуле (1.25):

$$x \vee xy = x(1 \vee y) = x.$$

Операция склеивания выполняется по аналогичной формуле:

$$xy \vee x\bar{y} = x,$$

справедливость которой легко доказывается с помощью распределительного закона и формул (1.8) и (1.10):

$$xy \vee x\bar{y} = x(y \vee \bar{y}) = x\,\text{И} = x.$$

Преобразования минимизации по методу Квейна выполняются в следующей последовательности:

1) минимизируемая формула приводится к совершенной нормальной дизъюнктивной форме;

2) выполняются все операции склеивания и затем поглощения.

В результате будем иметь сокращенную дизъюнктивную нормальную форму данной формулы.

Пример 1. Минимизировать формулу

$$F(x_1,\ x_2,\ x_3,\ x_4) = \bar{x}_1 x_2 \bar{x}_3 \vee x_1 \bar{x}_3 \bar{x}_4 \vee x_1 \bar{x}_2 \bar{x}_3 x_4 \vee \bar{x}_1 \bar{x}_2 \bar{x}_3 \bar{x}_4.$$

Решение. Заметим, что формула не является совершенной дизъюнктивной нормальной формой, так как в первом слагаемом отсутствует x_4, а во втором x_2. Добавим их путем перемножения соответствующих слагаемых на тождественно истинную формулу:

$$(\bar{x}_1 x_2 \bar{x}_3) \& (x_4 \vee \bar{x}_4) = \bar{x}_1 x_2 \bar{x}_3 x_4 \vee \bar{x}_1 x_2 \bar{x}_3 \bar{x}_4,\ \ x_1 \bar{x}_3 \bar{x}_4 \& (x_2 \vee \bar{x}_2) =$$
$$= x_1 x_2 \bar{x}_3 \bar{x}_4 \vee x_1 \bar{x}_2 \bar{x}_3 \bar{x}_4.$$

Подставляя в исходную формулу, получим

$$F(x_1,\ x_2,\ x_3,\ x_4) = \bar{x}_1 x_2 \bar{x}_3 x_4 \vee \bar{x}_1 x_2 \bar{x}_3 \bar{x}_4 \vee x_1 x_2 \bar{x}_3 \bar{x}_4 \vee$$
$$\vee x_1 \bar{x}_2 \bar{x}_3 \bar{x}_4 \vee x_1 \bar{x}_2 \bar{x}_3 x_4 \vee \bar{x}_1 \bar{x}_2 \bar{x}_3 \bar{x}_4.$$

Выполним операции склеивания. Для этого будем выбирать такие слагаемые, которые отличаются отрицанием над одним сомножителем. Напомним, что в алгебре логики можно одно и то же слагаемое использовать несколько раз (см. формулу (1.7)). Произведем склеивание первого слагаемого со вторым по x_4; получим

$$\bar{x}_1 x_2 \bar{x}_3 x_4 \vee \bar{x}_1 x_2 \bar{x}_3 \bar{x}_4 = \bar{x}_1 x_2 \bar{x}_3 (x_4 \vee \bar{x}_4) = \bar{x}_1 x_2 \bar{x}_3.$$

Аналогично:

$$\begin{array}{ll} 2-3 = x_2 \bar{x}_3 \bar{x}_4 & \text{по } x_1; \\ 3-4 = x_1 \bar{x}_3 \bar{x}_4 & \text{по } x_2; \\ 4-5 = x_1 \bar{x}_2 \bar{x}_3 & \text{по } x_4; \\ 2-6 = \bar{x}_1 \bar{x}_3 \bar{x}_4 & \text{по } x_2; \\ 4-6 = \bar{x}_2 \bar{x}_3 \bar{x}_4 & \text{по } x_1. \end{array}$$

После этого можно записать

$$F(x_1,\ x_2,\ x_3,\ x_4) = \bar{x}_1 x_2 \bar{x}_3 \vee x_2 \bar{x}_3 \bar{x}_4 \vee x_1 \bar{x}_3 \bar{x}_4 \vee x_1 \bar{x}_2 \bar{x}_3 \vee$$
$$\vee \bar{x}_1 \bar{x}_3 \bar{x}_4 \vee \bar{x}_2 \bar{x}_3 \bar{x}_4.$$

Возможны еще склеивания:

$$\begin{array}{ll} 2-6 = \bar{x}_3 \bar{x}_4 & \text{по } x_2; \\ 3-5 = \bar{x}_3 \bar{x}_4 & \text{по } x_1. \end{array}$$

Окончательно получим

$$F(x_1,\ x_2,\ x_3,\ x_4) = \bar{x}_1 x_2 \bar{x}_3 \vee \bar{x}_3 \bar{x}_4 \vee x_1 \bar{x}_2 \bar{x}_3.$$

Метод Гаврилова. Метод минимизации, предложенный М. А. Гавриловым, удобен в том случае, когда логическая формула задана в виде таблицы срабатывания.

Метод имеет строгое теоретическое обоснование, опирающееся на законы равносильных преобразований. Для практического использования он дается в виде следующего правила.

1. Строится логическая функция:

1) в таблице срабатывания выделяются строки, в которых $P = 1$;

2) по каждой такой строке строится логическая формула в виде элементарного произведения. Для этого 1 заменяется x_i, 0 — $\bar{x}_j$;

3) элементарные произведения объединяются действием дизъюнкции.

В результате возникает совершенная дизъюнктивная нормальная форма логической функции, которая определяет условие срабатывания данного устройства.

2. Над полученной совершенной дизъюнктивной нормальной формой выполняются преобразования минимизации.

Делается это последовательными этапами:

1) берется очередное (начиная с первого) элементарное слагаемое $x_1x_2 \ldots x_i \ldots x_n$ и в нем очередной сомножитель x_i (также начиная с первого);

2) производится проверка, реализует ли данный сомножитель хотя бы одну строку с нерабочим состоянием.

Если не реализует, то сомножитель x_i включается в минимизированную форму, заменяя собой все слагаемые $x_1x_2 \ldots x_i \ldots x_n$. Если реализует, то сомножитель отбрасывается. Таким же путем анализируется следующий сомножитель. Если окажется, что ни один из сомножителей, взятых по одному, не входит в минимизированную форму, то таким же путем производится проверка на реализуемость нерабочих состояний таблицы всех комбинаций, составленных из элементов, входящих в данное слагаемое, взятых по два, затем для всех комбинаций по три элемента и т. д.

Увеличение числа сомножителей прекращается тогда, когда в минимизированную формулу войдет сокращенный «представитель» от каждого слагаемого исходной формулы. Если какое-то слагаемое не дает «представителя» при данном числе сомножителей, то из него строятся комбинации с числом элементов на 1 больше предыдущего и производится отбор «представителя» в минимизированную форму.

Отобранные таким путем минимизированные слагаемые объединяются операцией дизъюнкции и в результате возникает так называемая общая минимизированная форма.

Смысл слов «реализует нерабочее состояние» состоит в следующем.

Пусть таблицей задана логическая функция, зависящая от n переменных $x_1, x_2, \ldots, x_n$. Рассмотрим комбинацию, составленную из m элементов ($m < n$). В нее могут входить переменные без отрицаний и с отрицаниями $x_i \ldots \overline{x_k} \ldots x_l$. Будем говорить: комбинация $x_i \ldots \overline{x_k} \ldots x_l$ *реализует нерабочее состояние*, если среди строк с нерабочим состоянием найдется хотя бы одна такая, в которой в столбцах, соответствующих $x_i, \ldots, x_l$ (без отрицаний), будет стоять 1, а в столбцах x_k — 0 (столбцы, не участвующие в комбинации $x_i \ldots \overline{x_k} \ldots x_l$, во внимание не принимаются).

Рассмотрим процедуру минимизации на примере.

Функционирование некоторого устройства задано таблицей рабочих состояний:

№ строки	x_1	x_2	x_3	x_4	P
1	1	1	1	0	1
2	1	1	0	0	1
3	1	0	1	0	1
4	0	1	0	1	1
5	1	0	0	1	0
6	0	1	1	0	0
7	0	0	1	0	0
8	0	0	0	1	0

Требуется построить соответствующую логическую формулу и минимизировать ее.

Замечание. Таблица всех комбинаций значений четырех аргументов должна содержать 16 строк. В примере даны только восемь. Очевидно, другие восемь комбинаций не представляют интереса для конструктора. Иначе говоря, здесь мы имеем дело с техническим приложением, а не с логическим исследованием.

Исходную формулу построим по первым четырем строкам:

$$P(x_1,\ x_2,\ x_3,\ x_4) = x_1x_2x_3\bar{x}_4 + x_1x_2\bar{x}_3\bar{x}_4 + x_1\bar{x}_2x_3\bar{x}_4 + \bar{x}_1x_2\bar{x}_3x_4.$$

Возьмем первое слагаемое и построим из него все комбинации по одному элементу. Это будет x_1, x_2, x_3, $\bar{x}_4$. Все они реализуют строки с нерабочим состоянием: так $x_1 = 1$ в пятой строке, $x_2 = 1$ в шестой строке, $x_3 = 1$ в шестой и седьмой и $x_4 = 0$ в пятой и восьмой строках. Комбинаций по два элемента будет шесть:

$$x_1x_2,\ x_1x_3,\ x_1\bar{x}_4,\ x_2x_3,\ x_2\bar{x}_4,\ x_3\bar{x}_4.$$

Определим, входит ли в минимизированную форму первая комбинация.

Среди строк с нерабочими состояниями нет такой, в которой было бы $x_1 = 1$ и $x_2 = 1$. Следовательно, x_1x_2 не реализует нерабочее состояние и потому входит в минимизированную форму. Также не реализуют ни одну из строк таблицы и входят в минимизированную форму комбинации x_1x_3 и $x_1\bar{x}_4$; комбинации x_2x_3, x_2x_4, x_3x_4 реализуют шестую, шестую и седьмую строки соответственно и потому должны быть отброшены.

Таким образом, после отбора по элементам первого слагаемого в минимизированную форму включаются три комбинации

$$x_1x_2,\ x_1x_3,\ x_1\bar{x}_4.$$

Аналогично проведем отбор по комбинациям второго слагаемого исходной формулы. Рассмотрим шесть комбинаций:

x_1x_2 — не реализует — остается в минимизированной форме;
$x_1\bar{x}_3$ — реализует пятую строку — отбрасываем;
$x_2\bar{x}_3$ и $x_1\bar{x}_4$ — не реализуют — остаются;
$x_2\bar{x}_4$ — реализует шестую строку — отбрасываем;
$\bar{x}_3\bar{x}_4$ — не реализует — остается.

От второго слагаемого остаются четыре комбинации:

$$x_1x_2,\ x_2\bar{x}_3,\ x_1\bar{x}_4,\ \bar{x}_3\bar{x}_4.$$

Из третьего слагаемого проверим шесть комбинаций:

$\bar{x}_1\bar{x}_2$ — реализует пятую строку — отбрасываем;
x_1x_3 — не реализует — оставляем;
$x_1\bar{x}_4$ — не реализует — оставляем;
$\bar{x}_2x_3$ — реализует седьмую строку — отбрасываем;
$\bar{x}_2\bar{x}_4$ — реализует седьмую строку — отбрасываем;
$x_3\bar{x}_4$ — реализует шестую и седьмую строки — отбрасываем.

От третьего слагаемого остаются только две комбинации:

$$x_1x_3,\ x_1\bar{x}_4.$$

Четвертое слагаемое:

$\bar{x}_1x_2$ — реализует шестую строку — отбрасываем;
$\bar{x}_1\bar{x}_3$ — реализует восьмую строку — отбрасываем;
$\bar{x}_1x_4$ — реализует восьмую строку — отбрасываем;
$x_2\bar{x}_3$ — не реализует — остается;
x_2x_4 — не реализует — остается;
x_3x_4 — реализует восьмую строку — отбрасываем.

Остались две комбинации:

$$x_2\bar{x}_3 \text{ и } x_2x_4.$$

Рассматривать комбинации по три нет необходимости, так как каждое слагаемое исходной формулы уже представлено хотя бы одним минимизированным членом. Объединив с помощью операции дизьюнкции остатки всех слагаемых, получим

$$P(x_1,\ x_2,\ x_3,\ x_4) = x_1x_2 \vee x_1x_3 \vee x_1\overline{x_4} \vee x_1x_2 \vee x_1\overline{x_4} \vee$$
$$\vee\ x_2\overline{x_3} \vee \overline{x_3}\,\overline{x_4} \vee x_1x_3 \vee x_1\overline{x_4} \vee x_2\overline{x_3} \vee x_2x_4.$$

Объединив одинаковые слагаемые (формула 1.7), получим общую минимизированную форму

$$P(x_1,\ x_2,\ x_3,\ x_4) = x_1x_2 \vee x_1x_3 \vee x_1\overline{x_4} \vee x_2\overline{x_3} \vee \overline{x_3}\,\overline{x_4} \vee x_2x_4.$$

Оказывается, эта форма допускает дальнейшие упрощения за счет способности некоторых слагаемых реализовывать одни и те же строки условия срабатывания. Принцип отбора состоит в том, чтобы из нескольких слагаемых, реализующих одно и то же условие срабатывания, отобрать любое одно и его оставить в минимизированной форме, а остальные отбросить. Такая операция легко осуществляется с помощью следующей таблицы, в которой столбцы озаглавливаются символами каждого слагаемого общей минимальной формулы; строчки нумеруются теми же номерами, какие они получили в таблице задания функции по условию срабатывания; на пересечении строки и столбца ставится крестик в том случае, если данное слагаемое реализует данную строку:

№ строки	x_1x_2	x_1x_3	$x_1\overline{x_4}$	$x_2\overline{x_3}$	$\overline{x_3}\,\overline{x_4}$	x_2x_4
1	+	+	+			
2	+		+	+	+	
3		+	+			
4				+		+

Выберем слагаемое. Очевидно, выгодно в минимизированную форму включить такие слагаемые, которые реализуют максимальное число строк. В нашем случае можно заметить несколько комбинаций такого типа. Например, второе слагаемое x_1x_3 с четвертым $x_2\overline{x_3}$ в совокупности охватывают все четыре строки условия срабатывания. Объединяя их, получим минимизированную форму, отображающую функционирование устройства, заданного таблицей срабатывания.

$$P(x_1,\ x_2,\ x_3,\ x_4) = x_1x_3 \vee x_2\overline{x_3}.$$

В этой формуле первое слагаемое реализует первую и третью строки таблицы, а второе — вторую и четвертую, и в то же время ни одно из них не реализует условия несрабатывания.

Кроме рассмотренной комбинации, может быть выбрана такая:

$$P(x_1, x_2, x_3, x_4) = x_1\overline{x_4} \vee x_2x_4.$$

Остальные комбинации хуже, так как требуют не менее трех слагаемых.

Кроме описанных, существуют и другие методы минимизации логических формул, в частности так называемые диаграммы Вейча (или карты Карно). Описания их есть в технической литературе.

Недостатками всех методов минимизации является возрастающая громоздкость преобразования с увеличением числа переменных.

В последнее время в литературе появился новый метод минимизации, так называемый *канонический метод*, предложенный А. Ш. Блохом. Он имеет следующие преимущества:

1) способность решать задачи минимизации логических функций двоичного, троичного и вообще k-ичного измерения, а также логических переменных смешанного состава;

2) допускает выполнение преобразований минимизации для сравнительно большого числа переменных, избегая при этом громоздкости операций;

3) первоначальной формой задания функции является полная таблица срабатывания (названная характеристической). Преобразования характеристической таблицы производятся по формальным правилам и могут быть поручены ЭЦВМ;

4) он одинаково успешно может применяться для синтеза функциональных схем переключательных устройств, для синтеза блок-схем алгоритмов сложных вычислительных и логических задач, для синтеза схем конечных автоматов.

К сожалению, в рамках настоящего пособия мы не можем осветить канонический метод достаточно подробно и вынуждены интересующихся читателей отослать к специальной литературе [3].

Рассмотрим пример синтеза функциональной схемы переключательного устройства.

Требуется разработать функциональную схему устройства, выполняющего операцию сложения действительных чисел.

Такие устройства называют *сумматорами*.

Всю работу проведем по следующему плану:

1. Определим входную и выходную информацию, т. е. что и в каком виде нам будет дано и что мы должны получить.

2. Установим закономерности, которые должен реализовать сумматор.

3. Разработаем функциональную схему сумматора. Для этого:

1) построим таблицу срабатывания;
2) получим логическую формулу;
3) выполним преобразования минимизации;
4) вычертим функциональную схему сумматора.

Мы в нашем примере на этом закончим работу. А в практике после получения функциональной схемы она анализируется, оценивается с точки зрения конкретной реализуемости: в соответствии с требованиями габаритов, надежности, технологии изготовления и т. д. подбираются материал и детали, из которых может быть смонтирован сумматор. В ходе такой работы могуть вновь производиться различные преобразования минимизации над логической формулой.

Затем вычерчивается принципиальная схема устройства, рабочие чертежи, пользуясь которыми, устройство можно изготовить в металле.

Приступим к выполнению нашей части работы.

На вход поступают два числа a и b. Очевидно, необходимо оговорить разрядность чисел: на бумаге можно записать число с большим количеством разрядов, а в устройстве каждый разряд будет требовать увеличения габаритов. Потому ограничим разрядность числа, например, семью разрядами.

Далее, числа удобно представить в двоичной системе счисления, так как в таком виде они изображаются набором нулей и единиц, а наборы легко переводятся в последовательность импульсов напряжений тока.

На выходе устройства должны получить одно число c — результат сложения a и b, представленный в двоичной системе счисления.

Напомним закономерности сложения чисел в двоичной системе счисления (справа указаны соответствующие действия):

a	b	S	
0	0	0	$0+0=0$
0	1	1	$0+1=1$
1	0	1	$1+0=1$
1	1	10	$1+1=10$

где 10 — изображение числа 2 в двоичной системе.

Из таблицы видим, что в результате сложения в трех случаях получаем одноразрядные числа, а в четвертом — двухразрядное. Для него в соответствии с законом сложения в разряде, в котором производится сложение, записывается 0, а 1 переходит в следующий числовой разряд. Пользуясь таблицей, выполним для примера сложение двух четырехразрядных двоичных чисел.

Пусть в десятичной системе слагаемые суть $a = 5$ и $b = 15$, тогда в двоичном счислении $a = 0101$, $b = 1111$. Сложим их:

$$\begin{array}{r} 1\,1\,1 \\ +\begin{array}{r} 0\,1\,0\,1 \\ 1\,1\,1\,1 \end{array} \\ \hline 1\,0\,1\,0\,0 \end{array}$$

Здесь вверху символом 1 обозначена единица, переходящая из предыдущего разряда в случае, если в результате сложения получается двухразрядное число:

$$1 + 1 = 10 = [2]_{10}, \quad 1 + 1 + 1 = 11 = [3]_{10}.$$

Из закона сложения вытекает, что действие должно производиться последовательно, поразрядно, начиная с младшего разряда, так как при определенных сочетаниях значений цифр может быть переход единицы в очередной числовой разряд.

В соответствии с этим на вход каждого разряда сумматора должны поступать три сигнала:

1) сигнал соответствующего разряда от первого слагаемого (0 или 1);

2) сигнал того же разряда от второго слагаемого (0 или 1);

3) 0 или 1 от предыдущего разряда.

Таким образом, одноразрядный сумматор в целом должен иметь три входа и два выхода: S и P (S — обозначает результат сложения в том же разряде; P — сигнал о переходе 1 в очередной разряд).

Последовательные действия n одноразрядных сумматоров создадут сумматор, выполняющий сложение двух n-разрядных двоичных чисел.

В литературе для обозначения одноразрядного сумматора используют символ, изображенный на рис. 3.24, и n-разрядного — на рис. 3.25.

Перейдем к синтезу функциональной схемы одноразрадного сумматора. Составим таблицу срабатывания. В ней будет пять столбцов: три для входов a, b, p и два для выходов P и S — и восемь строк:

a	*b*	*p*	*S*	*P*
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

Столбцы S и P заполнены в соответствии с законом сложения трех одноразрядных двоичных чисел, численные значения которых представлены в каждой строке. Так, в третьей

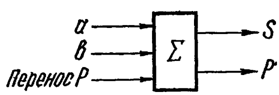

Рис. 3.24

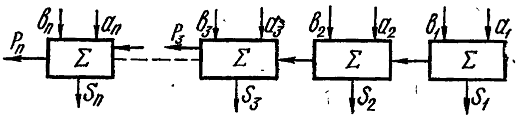

Рис. 3.25

строке имеем $0+1+0=1$ — одноразрядное число. Потому в столбце S записываем 1, в столбце P — 0. В четвертом столбце $0+1+1=10$; в столбце S записываем 0, в P — единицу, переходящую в очередной разряд. В восьмой строке $1+1+1=11$; единицы в обоих столбцах S и P.

Заполнив таблицу срабатывания, исходя из соображений закономерностей формализуемого процесса (в данном случае руководствуясь законами сложения), рассмотрим ту же таблицу как способ задания двоичной логической функции. Для этого в столбцах a, b и p символ 1 будем рассматривать как высказывания a, b и p соответственно, а 0 — как их отрицания $\bar{a}$, $\bar{b}$, $\bar{p}$.

Связь символов в строке изобразим в виде конъюнкции, а строк между собой — в виде дизъюнкции тех двоичных переменных a, b, p, которыми озаглавлены столбцы.

При таком истолковании в каждой строке первых трех столбцов имеем конституенту, составленную из логических переменных a, b и p в конъюнктивной форме.

Во всех строках имеем полный набор конституент. Объединяя их дизъюнктивно, получим совершенную дизъюнктивную форму логической формулы, с помощью которой формализован алгоритм сложения двух одноразрядных двоичных чисел.

При синтезе функциональной схемы сумматора необходимо построить две логические формулы: по одной для каждого выхода. Для этого достаточно из таблицы выбрать те конституенты, на которых данный выход срабатывает. Формально это будут те строки, в которых на данном выходе проставлена единица. Руководствуясь указанным соображением, выпишем две логические формулы:

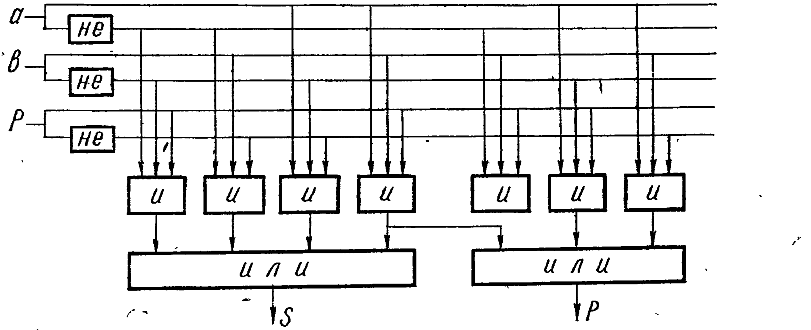

Рис. 3.26

$$S = \bar{a}\,\bar{b}p + \bar{a}b\bar{p} + a\bar{b}\,\bar{p} + abp,$$

$$P = \bar{a}bp + a\bar{b}p + ab\bar{p} + abp.$$

По формулам может быть построена функциональная схема одноразрядного сумматора (рис. 3.26).

Из алгебры логики известно, что совершенные нормальные формы допускают преобразования минимизации. Выполнив их, получим минимизированную функциональную схему такого же устройства.

В нашем примере возможны следующие эквивалентные преобразования: рассматривая формулу для P, замечаем, что первое и последнее слагаемые можно упростить, вынеся за скобку множитель bp: $\bar{a}bp + abp = bp(\bar{a} + a)$. Нетрудно заметить, что такое же упрощение дали бы два других слагаемых, если бы в формуле нашлись слагаемые вида abp. Алгебра логики позволяет дизъюнктивно добавить нужное число таких слагаемых. Пользуясь этим, выполним преобразования минимизации:

$$P = \bar{a}bp + a\bar{b}p + ab\bar{p} + abp + abp + abp =$$
$$= \bar{a}bp + abp + a\bar{b}p + abp + ab\bar{p} + abp =$$
$$= bp(\bar{a} + a) + ap(b + \bar{b}) + ab(p + \bar{p}) = ab + ap + bp.$$

Формула для S преобразуется сложнее.

Прежде упростим сумму трех слагаемых, в которой есть члены с отрицанием. Для этого с помощью действия отрицания представим их в конъюнктивной форме и произведем перемножение:

$$\bar{a}\bar{b}p + \bar{a}b\bar{p} + a\bar{b}\bar{p} = \overline{(a + b + \bar{p})(a + \bar{b} + p)(\bar{a} + b + p)} =$$
$$= \overline{[\bar{p} + (a + b)][\bar{b} + (a + p)][\bar{a} + (b + p)]} = \overline{\bar{a}\bar{b}\bar{p} + \bar{a}\bar{p}(a + p) +}$$
$$\overline{+ \bar{a}\bar{b}(a + b) + \bar{a}(a + b)(a + p) + \bar{p}\bar{b}(b + p) + \bar{p}(a + p)(b + p) +}$$
$$\overline{+ \bar{b}(a + b)(b + p) + (a + b)(a + p)(b + p)}.$$

Заметим, что некоторые слагаемые имеют одинаковую структуру. Так, второе, третье и пятое слагаемые представляют конъюнкцию трех членов. Причем одно из них есть дизъюнкция членов без отрицаний, а два других — эти же члены с отрицанием. Оказывается, такие слагаемые тождественно ложны.

Покажем это:

$$\bar{a}\bar{p}(a + p) = \bar{a}\bar{p}a + \bar{a}\bar{p}p = 0,$$

так как в конъюнкции есть элемент и его отрицание.

Четвертое, шестое и седьмое слагаемые также одинаковой структуры. Преобразуем одно из них:

$$\bar{a}(a + b)(a + p) = \bar{a}aa + \bar{a}ap + \bar{a}ba + \bar{a}bp = \bar{a}bp,$$

так как в трех первых слагаемых есть a и его отрицание $\bar{a}$, связанные конъюнктивно. Такое выражение тождественно ложно. Аналогично шестое и седьмое слагаемые дадут $\bar{p}ab$ и $\bar{b}ap$. Упростим последнее слагаемое:

$$(a + b)(a + p)(b + p) = aab + aap + apb + app + bab +$$
$$+ bap + bpb + bpp = ab + ap + abp + ap + ab + abp +$$
$$+ bp + pb = ab + ap + abp + pb = ab(1 + p) + ap + bp =$$
$$= ab + ap + bp.$$

Продолжим преобразования, учтя при этом выполненные упрощения:

$$\overline{(a+b+\bar p)(\bar a+b+p)(a+\bar b+p)}=$$
$$=\overline{\bar a\,\bar b\,\bar p+\bar a bp+a\bar b p+ab\bar p+ab+ap+bp}=$$
$$=\overline{\bar a\,\bar b\,\bar p+bp(\bar a+1)+ap(\bar b+1)+ab(\bar p+1)}=$$
$$=\overline{\bar a\,\bar b\,\bar p+ab+ap+bp}=\overline{\bar a\,\bar b\,\bar p+(ab+ap+pb)}=$$
$$=\overline{\bar a\,\bar b\,\bar p}\;\overline{ab+ap+bp}=(a+b+p)\bar p.$$

Таким образом, в результате эквивалентных преобразований формулы для S и P получат вид:

$$S=(a+b+p)\bar P+abp,$$
$$P=ab+ap+bp.$$

По преобразованным формулам построим минимизированную функциональную схему одноразрядного сумматора (рис. 3.27).

Сравнивая функциональные схемы на рис. 3.26 и 3.27, замечаем существенные упрощения схемы после эквивалентных преобразований формулы. Дальнейшая работа конструктора по синтезу сумматора состоит в преобразовании функциональной схемы в принципиальную. Для этого символы

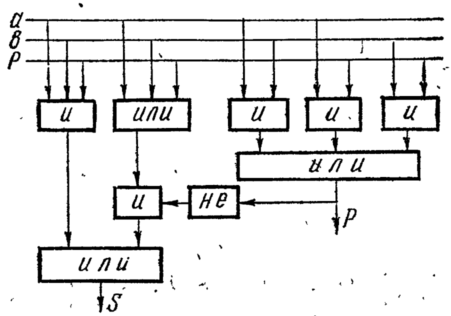

Рис. 3.27

и, **или** и **не** должны быть заменены схемами конкретных устройств. При этом, в соответствии с техническим заданием из различных видов устройств выбираются наиболее целесообразные, удовлетворяющие требованиям габаритов, стоимости, надежности и т. д.

На рассмотренном примере мы показали только идею синтеза функциональных схем переключательных устройств. В практике эта работа значительно сложнее.

Трудности вытекают из следующих обстоятельств. Число логических переменных $x_1, x_2, \ldots, x_n$, действие которых нужно учитывать при синтезе устройства, бывает сравнительно большим. Вследствие этого становятся громоздкими таблицы срабатывания и соответствующие им логические формулы. Становятся сложными и громоздкими эквивалентные преобразования. В практике разработан целый ряд упрощенных удобных приемов составления таблиц срабатывания и минимизации логических формул (булевых функций).

Систематизированное изложение этих методов относится к специальным курсам и имеется в соответствующей литературе.

§ 3. 3. ПРИМЕНЕНИЕ МЕТОДОВ МАТЕМАТИЧЕСКОЙ ЛОГИКИ В ВЫЧИСЛИТЕЛЬНОЙ ТЕХНИКЕ

Бурное развитие вычислительной техники и численных методов анализа, проникновение их в различные естественные и гуманитарные науки во многом обязано математической логике. Укажем на некоторые из направлений, в которых используются ее идеи и методы.

1. Конструирование электронных цифровых и аналоговых вычислительных машин.

2. Разработка алгоритмов решения инженерно-расчетных задач и математического моделирования.

3. Программирование, минимизация и автоматизация программ.

Общие черты первого направления мы осветили в предыдущем изложении.

Второй вопрос — разработка алгоритмов — является достаточно большим, сложным и в ряде случаев проблемным. Дело в следующем. Если предполагается решить на ЭЦВМ инженерно-расчетную задачу, заданную математическими формулами, то этим самым уже задан алгоритм ее решения. Программирование в этом случае принципиальных трудностей не вызывает.

Но в последнее время осуществляются успешные попытки решать с помощью ЭЦВМ задачи, которые не представляются математическими соотношениями. К таковым относятся: различные информационно-логические задачи, моделирование производственных процессов, управление боевыми действиями войск и вообще управление динамическими процессами; задачи биологии, медицины, лингвистики и т. д.

Оказывается, в этих нематематических областях использование ЭЦВМ наиболее эффективно.

Выше подчеркивалось, что для решения перечисленных типов задач еще нет достаточно хороших и общих математических методов, нет символики для обозначения различных понятий, объектов и отношений между ними. На практике все эти задачи решаются усилиями творческой работы человека, который руководствуется опытом, логическими соображениями, оценками складывающихся ситуаций, интуицией, принимает волевые решения. Если мы хотим решать такие задачи на ЭЦВМ, то ей должен быть указан алгоритм решения. Идеи построения таких алгоритмов, способных отображать любые (конечные) процессы, независимо от их природы и системы обозначений, были сфомулированы и получили законченное развитие в математической логике. Это так называемые нормальные алгоритмы А. А. Маркова (и эквивалентные им системы — «машина Тьюринга», рекурсивные функции).

Перенесение этих идей на практическую почву позволяет создавать алгоритмы и успешно решать на ЭЦВМ множество труднейших задач. В практике при разработке алгоритмов многих задач и в настоящее время встречаются принципиальные трудности. Можно сказать, что законченной теории разработки алгоритмов логических задач нет и до сих пор. И, очевидно, математической логике еще предстоит внести свой вклад в решение подобных проблем.

В решении третьего вопроса — в программировании и минимизации программ — уже сейчас достигнуты значительные результаты. Суть применения идей и методов математической логики в этой работе состоит в следующем. Допустим, для решения на ЭЦВМ готовится какая-то определенная задача. Решить ее на ЭЦВМ означает: разработать соответствующий алгоритм, составить по нему машинную программу, отладить ее на машине и решить по ней задачу.

Как и в любом деле, алгоритм может быть построен лучше или хуже, и естественно возникает вопрос об оптимизации алгоритмов и составленных по ним программ.

Методы математической логики находят приложение и в этих вопросах.

Познакомимся прежде с основными понятиями, которыми пользуются при разработке алгоритмов. Это операторы — арифметические, логические, служебные (машинные), — и блок-схема алгоритма (операторные формулы).

Для разъяснения понятий используем условие конкретной задачи, например планирование перевозок в аэропорту.

В самых общих чертах постановка задачи может быть сформулирована следующим образом. В порту есть k типов самолетов. Известно количество их по типам. Известен поток заявок на перевозки с указанием места перевозок, типа груза (срочный, малогабаритный, скоропортящийся и т. д.). Требуется составить расписание для оптимального удовлетворения запросов на перевозки.

Не вдаваясь в подробности построения алгоритма, можно указать математические методы, используемые для решения частных вопросов. Так, определение времени перевозки с учетом расстояния и типа самолета решается с помощью систем дифференциальных уравнений; задачи планирования ремонтных и профилактических работ решаются с использованием вероятностных методов; задачи выбора оптимального распределения маршрутов и типов самолетов в зависимости от поступивших на перевозки заявок могут решаться в рамках линейного программирования или игровыми методами и т. д.

Таким образом, мы приходим к необходимости использования различных математических теорий со своей спецификой и приемами решений.

Спрашивается: как, в каких символах изобразить взаимную связь и обусловленность отдельных частей в общем процессе. Как формулировать задачу, чтобы решение ее могло быть выполнено автоматически на ЭЦВМ. Для этой цели используются операторы и блок-схемы алгоритмов. *Арифметическими операторами* называют такие, в которых выполняются вычислительные или специфические машинные операции. Обозначаются они A_i.

Содержание операций можно выразить с помощью элементарного высказывания, например «Найти корни уравнения $y = f(x)$», «Вычислить численное значение выражения». Смысл «истинно», «ложно» понимается как выполнение указанных действий: если корни найдены, то высказывание «Найти корни уравнения» — истинно.

Логические операторы символизируют условия, в которых производятся вычисления или протекает моделируемый процесс. Выражаются они с помощью предикатов $a > b$, $a < b$, $a = b$, $a \neq b$. В них никаких вычислений не производится, а определяется путь дальнейших вычислений в зависимости от выполнения или не выполнения условия, которое символизирует данный оператор. Если условие выполнено, то по физическому смыслу решение должно проходить в определенном арифметическом операторе, например A_k. Если не выполнено, то в другом, например A_l.

В буквенных символах используются следующие обозначения: логическое условие, выраженное одним из предикатов $a > b$, $a < b$, $a = b$, $a \neq b$, обозначают значком ω. Числа a и b вычисляются предшествующими арифметическими операторами. И если, сравнивая числа, получим, что предикат ω оказался истинным, то полагаем, что условие, символизированное в нем, выполняется. В противном случае — не выполняется.

Логические операторы обозначают P_i. Иногда удобно указывать условие, которое реализует оператор. Это записывается так:

$$P_i(\omega_i).$$

Смысл логического оператора выражается формулой

$$P_i(\omega_i) \Longrightarrow \begin{cases} A_k, & \text{если условие } \omega_i \text{ выполняется;} \\ A_l, & \text{если условие } \omega_i \text{ не выполняется.} \end{cases}$$

Знак $\Longrightarrow$ не определяет математического равенства операторов а указывает, который из арифметических операторов должен реализоваться в алгоритме в случае выполнения и невыполнения логического условия ω_i.

Из формулы следует, что логический оператор есть двоичная функция.

Алгоритм любой задачи в формальном изображении может быть представлен блок-схемой, состоящей из арифметических и логических операторов. Первые указывают, какие вычислительные операции следует выполнять, вторые в зависимости от выполнения или невыполнения логического условия ω_i указывают однозначно определенный арифметический оператор, подлежащий реализации.

Так как логические операторы — двоичные функции, то существует 2^n (n — число логических операторов) различных комбинаций, составленных из фиксированных значений логических условий. Каждой такой комбинации соответствует единственный набор арифметических операторов, который в численных выражениях характеризует течение алгоритмизированного процесса.

Схематическое отображение всех таких комбинаций решения называют *блок-схемой алгоритма*.

По внешнему начертанию блок-схема алгоритма выглядит следующим образом.

Арифметические операторы обозначаются какой-либо геометрической фигурой, например прямоугольником. Внутри нее словами или формулой записывается содержание действий, выполняемых данным оператором.

Логические операторы, для отличия от арифметических, обозначаются другой фигурой (кружок, овал, ромб) и внутри нее записывается формулировка логического условия, которое проверяется.

Символы операторов располагают в строгой последовательности, соответствующей законам математических действий и логическому построению решения.

Последовательность обозначается с помощью стрелок. Количество входящих в блок стрелок не ограничивается. Число выходных стрелок строго ограничивается: из арифметического оператора выходит одна стрелка, из логического строго две: одна указывает последовательность решения в случае выполнения условия, записанного в данном логическом операторе, другая — при невыполнении его.

В целом блок-схема представляет замкнутый контур с одним входом и одним выходом на останов, когда решение закончено. Внутри контура определены несколько путей решения задачи в соответствии с различными колебаниями значений логических условий.

В блок-схемах возможны перестановки некоторых логических и арифметических операторов. Поэтому конструкция блок-схем носит отражение субъективных воззрений разработчика алгоритма и допускает оптимизирующие ее преобразования.

Однако по самой блок-схеме такие преобразования выполнять трудно. Значительно удобнее представить эту же блок-схему в виде так называемых *операторных формул*. Внешний вид такой формулы:

$$A_0A_1A_2P_3A_4A_5P_6A_7Я.$$

Читается она так: операторы располагаются в строку и нумеруются индексами. Выполняются все вычислительные операции в A_0 и управление передается соседнему оператору A_1, затем A_2 и т. д. до тех пор, пока встретится логический оператор P_i. Логический оператор имеет два выхода. Принято считать: при выполнении условия в P_i управление передается на соседний оператор, стоящий справа, при невыполнении — на тот арифметический оператор, к которому ведет нижняя стрелка. В нашем примере оператор P_3 передает управление при выполнении условия на A_4 и невыполнении — на A_5; оператор P_6 при выполнении условия — на A_7 и невыполнении возвращает счет к началу на оператор A_1, образуя цикл.

На практике операторные формулы бывают громоздкие, содержат десятки и даже сотни операторов, соединенных сложной системой передачи управления, указанной стрелками. Оказывается, операторные формулы при некоторых допущениях могут быть истолкованы как формулы исчисления высказываний. Тогда к ним применима алгебра высказываний.

Преобразования минимизации операторных формул позволяют исключить из блок-схемы алгоритмов повторяющиеся части, отыскать кратчайшие пути решений и тем самым сократить программу и время реализации ее на ЭЦВМ.

Аппарат таких преобразований разработан в трудах Ю. И. Янова и описан в литературе по программированию.

В заключение отметим, что в настоящем пособии мы познакомились с небольшим числом направлений, в которых успешно применяется математическая логика. Очевидно, в самом ближайшем будущем круг таких наук значительно пополнится. Произойдет это не только за счет внедрения методов математической логики в научные исследования, но и в большей степени за счет включения в программы высшей и средней школы некоторых идей и методов математической логики как средства становления и развития логического мышления.

Вопросы и упражнения к разделу II

§ 1.1

1. Что мы называем формальной или классической логикой?
2. В чем смысл закона исключенного третьего?
3. Какую роль в развитии математической логики сыграли Лейбниц, Буль, Джевансон, Порецкий и другие ученые XVIII и первой половины XIX вв?
4. Какую роль в развитии математики сыграли работы Н. И. Лобачевского?
5. Что мы называем кризисом оснований математики?
6. Приведите примеры парадоксов в основаниях математики.
7. В чем различие кризиса оснований математики XVIII и конца XIX вв?
8. Какую роль в развитии математической логики сыграли работы Гильберта?
9. Что называется метатеорией и, в частности, метаматематикой? Что она изучает?
10. Имена каких ученых конца XIX и XX вв., занимавшихся математической логикой, вы знаете?

11. Смогла ли математическая логика решить все противоречия оснований математики? В чем смысл теоремы Гёделя?

12. Какова роль математической логики в настоящее время?

§ 1.2

1. Что называется аксиомой? Системой аксиом?
2. Каким требованиям должна удовлетворять система аксиом? Все ли требования обязательны?
3. Что такое интерпретация системы аксиом?
4. Что понимается под термином «высказывание»?
5. Какие отношения между высказываниями рассматриваются в математической логике?
6. В чем особенность аксиоматики математической логики?
7. Что такое правило вывода? Сформулируйте правило подстановок и правило заключений.

§ 1.3

1. Каким требованиям должно удовлетворять высказывание? Все ли предложения речевого общения являются высказываниями в смысле математической логики?
2. Что такое сложное высказывание? Какие значения истинности оно может принимать? Какими символами они обозначаются?
3. Что такое двоичная функция?
4. Что называется областью существования и областью изменения двоичной функции?
5. Что называется набором значений аргументов двоичной функции?
6. Если значения истинности обозначить 0 и 1, то является ли набор числом в двоичной системе счисления?
7. В следующих предложениях укажите высказывания и определите значения истинности:

1) гипотенуза длиннее катета;
2) $(3+4)(7-5)=15$;
3) $\frac{4}{7}$ — четное число;
4) Волга впадает в Черное море;
5) да здравствует мир!
6) лошадь — домашнее животное;
7) квадратное уравнение имеет два действительных или комплексных корня;
8) какая сегодня погода?

Ответы: 1), 6), 7) — истинные высказывания; 2) — 4) — ложные высказывания; 5), 8) — не являются высказываниями.

8. В следующих примерах найдите переменные высказывания и их фиксированные значения. Подберите значения x, при которых переменное высказывание становится истинным и ложным:

1) 50 кратно 7; x кратно 7;
2) в x месяце холодно; в январе месяце холодно;
3) город x стоит на берегу моря.

§ 1.4

1. Дать определения логическим операциям. Написать для них таблицы задания.

2. Доказать с помощью таблиц истинности переместительные, сочетательные и распределительные законы алгебры высказываний.

3. Доказать с помощью равносильных преобразований справедливость следующих формул:

1) $x \,\&\, y = \overline{\overline{x} \vee \overline{y}}$;

2) $\overline{A \to B} = A \,\&\, \overline{B}$;

3) $(x \,\&\, y) \vee (x \,\&\, \overline{y}) = x$;

4) $x_1 \,\&\, x_2 \vee \overline{x_1} \,\&\, x_2 \vee \overline{x_1} \,\&\, \overline{x_2} = x_1 \to x_2$;

5) $A \,\&\, (\overline{A} \vee B) = A \,\&\, B$;

6) $\overline{x \,\&\, \overline{y}} \to (\overline{y} \to x) = \overline{x \to y} \vee x \vee y$.

4. Упростить следующие формулы:

1) $(\overline{\overline{x} \vee y} \to x \vee y) \,\&\, y$;

2) $x_1 \& x_3 \vee x_1 \& \overline{x_3} \vee x_2 \,\&\, x_3 \vee \overline{x_1} \,\&\, x_2 \& x_3$;

3) $\overline{(A \to B) \,\&\, (B \to \overline{A})}$;

4) $(x \to \overline{y}) \vee \overline{(x \vee y)}$.

Ответы. 1) y; 2) $x_1 \vee x_2 \,\&\, x_3$; 3) A; 4) $\overline{X} \vee \overline{y}$.

5) Следующие формулы преобразовать так, чтобы знак отрицания относился к элементарным высказываниям:

1) $\overline{\overline{x} \vee y}$;

2) $\overline{A \,\&\, B \vee \overline{C}}$;

3) $\overline{x_1 \,\&\, (\overline{x_2} \vee \overline{x_3}) \,\&\, x_3}$.

Ответы. 1) $x \,\&\, \overline{y}$; 2) $(\overline{A} \vee \overline{B}) \,\&\, C$; 3) $\overline{x_1} \vee x_2 \,\&\, x_3 \vee \overline{x_3}$.

6. Следующие формулы преобразовать так, чтобы они содержали только действия $\&$, $\vee$, $\overline{}$:

1) $(x \to y) \,\&\, (y \to x) \to (x \sim y)$;

2) $(x \to y) \,\&\, (\overline{y} \to \overline{z}) \to (z \to \overline{x})$;

3) $(A \sim B) \,\&\, (\overline{A} \sim \overline{B}) \to (A \vee B) \,\&\, (\overline{A} \vee \overline{B})$.

Ответы. 1) $x \,\&\, \overline{y} \vee y \,\&\, \overline{x} \vee \overline{x} \,\&\, \overline{y} \vee x \,\&\, y \equiv И$; 2) $x \,\&\, \overline{y} \vee \overline{y} \,\&\, z \vee \vee \overline{z} \vee \overline{x}$; 3) $(\overline{A} \,\&\, B) \vee (A \,\&\, \overline{B})$.

§ 1.5

1. Привести к дизъюнктивной нормальной форме:

1) $x \& (x \to y) \to y$;

2) $(x \sim y) \& (x \to z)$;

3) $(A \vee B \vee C) \& (\overline{A} \vee \overline{B})$;

4) $(x_1 \vee x_2) \& (\overline{x_1 \& x_2} \vee x_3) \vee x_3$.

Ответы. 1) $\bar{x} \vee x \& \bar{y} \vee y \equiv И$; 2) $\bar{x} \& \bar{y} \vee x \& y \& z$; 3) $A \& \overline{B} \vee \overline{A} \& C \vee \overline{A} \& B \vee \overline{B} \& C$; 4) $x_1 \& \bar{x}_2 \vee x_2 \& \bar{x}_1 \vee x_3 \vee x_1 \& x_3 \vee x_2 \& x_3$.

2. Привести к конъюнктивной нормальной форме:

1) $A \& (B \vee C) \& \overline{B \vee A \& C}$;

2) $(x_1 \vee x_2) \to (\bar{x}_2 \vee x_3)$;

3) $\overline{[(x_1 \vee x_2) \to x_3] \vee (\bar{x}_1 \& x_2 \& \bar{x}_3}$;

4) $[A \sim (B \vee C)] \& (\overline{A} \vee \overline{B})$;

5) $\overline{A \vee B \vee C} \to D \vee (\overline{\overline{A} \vee B})$.

Ответы. 1) $(A \vee B) \& (A \vee C) \& (\overline{B} \vee \overline{A}) \& (\overline{B} \vee \overline{C})$; 2) $(\bar{x}_1 \vee \bar{x}_2 \vee x_3) \& (\bar{x}_2 \vee x_3)$; 3) $(\bar{x}_3 \vee x_1) \& (\bar{x}_3 \vee x_2) \& (x_1 \vee \bar{x}_2 \vee x_3)$; 4) $(\overline{A} \vee B \vee C) \& (\overline{B} \vee A) \& (\overline{C} \vee A) \& (\overline{A} \vee \overline{B})$; 5) $(A \vee B \vee C \vee D \vee \overline{A}) \& (A \vee B \vee C \vee D \vee \overline{B}) \equiv И$.

§ 1.6

1. С помощью таблиц истинности установить тождественную истинность, ложность или выполнимость следующих формул:

1) $A \& B \to A \vee B$;

2) $(x \to y) \to (\bar{y} \to \bar{x})$;

3) $\overline{(x_1 \sim x_2) \vee x_1}$;

4) $x \& y \& z \vee x \& \bar{y} \& z \vee x \& y \& \bar{z} \vee \bar{x} \& \bar{y} \& \bar{z}$.

2. Доказать тождественную истинность формул с помощью приведения их к конъюнктивной нормальной форме:

1) $x \& y \to x$;

2) $(x_1 \to x_2) \& x_1 \to (x_1 \sim x_2)$;

3) $(A \to B) \& A \to B$;

4) $(x \to y) \& \bar{y} \to \bar{x}$;

5) $x \& y \to \overline{x \leftrightarrow y}$.

3. Доказать тождественную ложность формул путем приведения их к дизъюнктивной нормальной форме:

1) $(A \to B) \& A \& \overline{B}$;

2) $\overline{x \& y \to x}$;

3) $\overline{x_1 \sim x_2} \& (\bar{x}_1 \vee x_2) \& (x_1 \vee \bar{x}_2)$;

4) $(x \& y \vee \bar{x} \& \bar{y}) \sim (\bar{x} \vee \bar{y}) \& (x \vee y)$.

§ 1.7

1. Привести к совершенной дизъюнктивной нормальной форме:

1) $x \to y \& z$;

2) $x \to (y \to x)$;

3) $x \& \bar{y} \to \overline{x \to y}$;

4) $(x \to y) \& (y \to z) \to (x \to z)$.

Ответы. 1) $x \& y \& z \vee \bar{x} \& y \& z \vee \bar{x} \& y \& \bar{z} \vee \bar{x} \& \bar{y} \& z \vee \bar{x} \& \bar{y} \& \bar{z}$; 2) $x \& y \vee \bar{x} \& y \vee x \& \bar{y} \vee \bar{x} \& \bar{y}$; 3) $x \& y \vee \bar{x} \& y \vee x \& \bar{y} \vee \bar{x} \& \bar{y}$; 4) $x \& \& y \& z \vee \bar{x} \& y \vee z \vee x \& \bar{y} \& z \vee x \& y \& \bar{z} \vee \bar{x} \& \bar{y} \& z \vee \bar{x} \& y \& \bar{z} \vee x \& \& \bar{y} \& \bar{z} \vee \bar{x} \& \bar{y} \vee z$.

2. Привести к совершенной конъюнктивной нормальной форме:

1) $(x \to y) \& x \& y$;

2) $\overline{x \sim y} \& (\bar{x} \vee y) \& (x \vee \bar{y})$;

3) $(x \to y) \& (\bar{y} \vee z) \& \overline{(x \to y)}$.

Ответы. 1) $(x \vee y) \& (\bar{x} \vee y) \& (x \vee \bar{y})$; 2) $(x \vee y) \& (\bar{x} \vee y) \& (x \vee \vee \bar{y}) \& (\bar{x} \vee \bar{y})$; 3) $(x \vee y \vee z) \& (\bar{x} \vee y \vee z) \& (x \vee \bar{y} \vee z) \& (x \vee y \vee \bar{z}) \& \& (\bar{x} \vee \bar{y} \vee z) \& (\bar{x} \vee y \vee \bar{z}) \& (x \vee \bar{y} \vee \bar{z}) \& (\bar{x} \vee \bar{y} \vee \bar{z})$.

3. Разложить по конституентам следующие формулы:

1) $(x \sim y) \vee x$;

2) $(x \vee y) \to (x \to z)$.

Ответы. 1) $\bar{x} \& \bar{y} \vee x \& y \vee x \& \bar{y}$; $x \vee \bar{y}$; 2) $x \& y \& z \vee x \& \bar{y} \& z \vee \vee \bar{x} \& y \& z \vee \bar{x} \& y \& \bar{z} \vee \bar{x} \& \bar{y} \& z \vee \bar{x} \& \bar{y} \& \bar{z}$; $(\bar{x} \vee y \vee z (\& (\bar{x} \vee \bar{y} \vee z)$.

§ 1.8

1. Найти все логические следствия из посылок F_1; F_2; F_3:

1) $F_1 = x \sim y$; $F_2 = \bar{x}$;

2) $F_1 = A \to B$; $F_2 = B \to C$;

3) $F_1 = x \vee y$; $F_2 = x$; $F_3 = \bar{y}$.

Ответы. 1) $x \vee \bar{y}$; $\bar{x} \vee y$; $\bar{x} \vee \bar{y}$; $\bar{x}$; $\bar{y}$; $\bar{x} \& \bar{y}$; $x \sim y$; 2) $\bar{A} \vee B \vee C$; $\bar{A} \vee B \vee \bar{C}$; $A \vee \bar{B} \vee C$; $\bar{A} \vee \bar{B} \vee C$; $\bar{A} \vee B$; $C \vee (A \sim B)$; $\bar{A} \vee C$; $(A \sim B) \vee \vee (B \sim C) \vee \overline{(A \sim C)}$; $\bar{A} \vee (B \sim C)$; $\bar{B} \vee C$; $(\bar{A} \vee B) \& (A \vee \bar{B} \vee C)$; $\bar{A} \vee \vee B \& C$; $C \vee \bar{A} \& \bar{B}$; $(\bar{B} \vee C) \& (\bar{A} \vee B \vee \bar{C})$; $(\bar{A} \vee B) \& (\bar{B} \vee C)$.

§ 2.1

1. В следующих выражениях указать одноместные, двуместные предикаты; установить, истинны они или ложны:

1) $A(10, 5)$; $A(25, 12)$, где A означает «делиться»;

2) $B(\sqrt{2})$; $B\left(\frac{10}{3}\right)$; $B(0{,}5)$, где B означает «быть рациональным числом»;

3) C (железо), C (сера), где C означает «быть металлом»;

4) $R(1, 2) \& H(2)$; $R(1, 2) \& H(1)$, где R — «предшествует»; H — «быть четным».

Ответы. 1) Двуместный; истинный; ложный; 2) одноместный; ложный; истинный; истинный; 3) одноместный; истинный; ложный; 4) двуместный и одноместный; И, Л.

2. Прочитать следующие формулы:

1) $P(x, y) \& \theta(y, z) \to P(x, z)$, где P означает «делиться», θ — «совпадать»; x, y — любые числа;

2) $[(S(x, 2) \& S(x, 3) \to S(x, 6)] \sim V(2) \& V(3)$, где S означает «делиться», V — «быть простым числом»;

3) $[A(a) \& A(b) \& A(c) \& (a \perp c) \& (b \perp c)] \to (a \parallel b)$; где A означает «принадлежать плоскости α».

Ответы. 1) Если x делится на y и z совпадает с y, то x делится на z; 2) утверждение о том, что если x делится на 2 и x делится на 3, то x делится на 6, равносильно утверждению, что 2 и 3 простые числа; 3) если прямые a, b, c принадлежат плоскости α и a перпендикулярна c и b перпендикулярна c, то a параллельна b.

3. Составить формулы, соответствующие следующим предложениям:

1) если $x = 0$ или $y = 0$, то $xy = 0$;

2) если $a < b$ и $b < c$, то $a < c$;

3) если $a \perp M$ и $b \perp M$, то $a \parallel b$,

где a, b — прямые, M — плоскость.

Ответы. 1) $(x = 0) \vee (y = 0) \to (xy = 0)$; 2) $(a < b) \& (b < c) \to (a < c)$; 3) $(a \perp M) \& (b \perp M) \to (a \parallel b)$.

§ 2.2

1. Прочитать следующие выражения. Установить для них значения истинности (x, y, z — произвольные числа):

1) $(\forall x)(|x| = -x)$;

2) $(\forall x)(x + 1 = x)$;

3) $(\exists x)(\forall y)(x + y = y)$;

4) $(\forall x)(\forall y)(\exists z)(x + y = z)$.

Ответы. 1) Для всякого x абсолютная величина числа противоположна самому числу; ложно; 2) для всякого x значение числа от прибавления единицы не меняется; ложно; 3) существует число x такое, что от прибавления его к любому числу y значение y не меняется; истинно. (Такое число $x = 0$.); 4) для любых чисел x и y существует число, равное их сумме; истинно.

2. Прочитать следующие выражения. Установить для них значения истинности (a, b, c — произвольные прямые на плоскости; $\theta(a, b)$ означает, что a пересекает b):

1) $(\forall a)(\exists b)\,\theta(a, b)$;

2) $(\forall a)(\forall b)[\theta(a, b) \to \overline{a \parallel b}]$;

3) $(\forall a)(\forall b)[\theta(a, b) \vee a \parallel b]$;

4) $(\forall a)(\forall b)[\theta(a, b) \& (\exists c)(c \parallel b) \to \theta(c, a)]$.

Ответы. 1) Для любой прямой a существует прямая b, которая ее пересекает; истинно; 2) для любых прямых справедливо: если a и b пересекаются, то они не параллельны; истинно; 3) для любых прямых справедливо: a и b пересекаются или параллельны; истинно; 4) для любых прямых a и b справедливо: если a пересекает b и существует прямая c параллельная b, то c пересекает a.

3. Следующие высказывания записать с помощью символов. Установить их истинность:

1) существует число x такое, что для каждого числа a справедливо $ax = a$;

2) для всяких чисел a и b существует число x такое, что $a + b = x$;

3) не существует рационального числа x такого, что $x^2 = 2$.

Ответы. 1) $(\exists x)(\forall a)(ax = a)$; истинно; 2) $(\forall a)(\forall b)(\exists x)(a + b = x)$; истинно; 3) $\overline{(\exists x)}(x \in R \,\&\, x^2 = 2)$; истинно ($R$ —множество рациональных чисел).

4. Следующие предложения записать в виде формул, введя кванторы и предикаты. Указать связанные и свободные переменные:

1) существует число x, которое делится на y;

2) если $x < y$, то существует число z такое, что $x < z$.

Ответы. 1) $(\exists x)(x : y)$; y — свободная переменная; 2) $(x < y) \to$ $\to (\exists z)(x < z)$; x, y — свободные переменные.

5. Найти отрицания следующих формул:

1) $(\exists x)[A(x) \,\&\, B(x) \vee C(x)]$;

2) $(\forall x)[A(x) \to (\forall y) B(y)]$;

3) $(\forall x)[A(x) \vee (\exists y) B(y)]$;

4) $(\forall x)[A(x) \to B(x)] \,\&\, (\exists x)[S(x) \,\&\, \overline{R}(x)]$.

Ответы. 1) $(\forall x)\{[\overline{A(x)} \vee \overline{B}(x)] \,\&\, \overline{c}(x)\}$; 2) $(\exists x)[A(x) \,\&\, (\exists y)\overline{B}(y)]$; 3) $(\exists x)[\overline{A}(x) \,\&\, (\forall y)\overline{B}(y)]$; 4) $(\exists x)[A(x) \,\&\, \overline{B}(x)] \vee (\forall x)[S(x) \to R(x)]$.

6. Найти приведенную форму следующих формул:

1) $(\forall x)(\exists y)[A(x) \vee \overline{(\forall z)(B(x, z) \to C(x, z))}]$;

2) $(\forall x) A(x) \vee (\exists x)[B(x) \to P(x)]$;

3) $(\forall x)(\forall y)[A(x) \,\&\, B(y) \to (\exists z) B(z)]$.

Ответы. 1) $(\forall x)(\exists y)\{A(x) \vee (\exists z)[B(x, z) \,\&\, \overline{C}(y, z)]\}$; 2) $(\exists x)\overline{A}(x) \,\&$ $\&\, (\forall x)[B(x) \,\&\, \overline{P}(x)]$; 3) $(\forall x)(\forall y)[\overline{A}(x) \vee \overline{B}(y) \vee \exists(z) B(z)]$.

§ 3.1

1. Как изображаются конъюнкция, дизъюнкция и отрицание в электросетях?

2. Какими символами изображаются логические операции в функциональных схемах переключательных устройств?

3. Построить функциональные схемы переключательных устройств, предварительно упростив их:

1) $P = AB\overline{C} \vee \overline{A}B\overline{C} \vee A\overline{B}\overline{C}$;

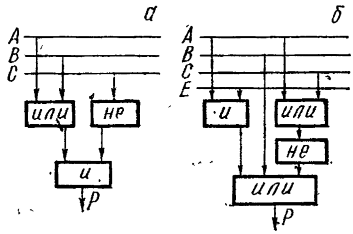

Рис. 3.28

2) $P = A\,[E \vee B(\overline{C}\overline{D}B \vee B \vee C \vee D)] \vee \overline{A}\,(B \vee \overline{C})$.

Ответы. 1) $P = (A \vee B)\,\&\,\overline{C}$ (рис. 3.28, *а*); 2) $P = B \vee AE \vee \overline{A \vee C}$ (рис. 3.28, *б*).

§ 3.2

1. Произвести синтез дешифратора с четырьмя входами A, B, C, D и тремя выходами P_1, P_2, P_3, функционирование которого задано таблицей срабатывания:

N	A	B	C	D	P_1	P	P_3
1	0	0	0	1	1	0	0
2	0	0	1	0	1	0	0
3	0	0	1	1	0	1	0
4	1	1	0	0	0	1	0
5	1	1	0	1	0	0	1
6	1	1	1	0	0	0	1

Ответ. $P_1 = \overline{A \vee CD}$; $P_2 = CD \vee \overline{C \vee D}$; $P_3 = A\,(D \vee C)$ (рис. 3.29).

2. По таблице задания функций составить исходные формулы, минимизировать их по методу Гаврилова и составить функциональную схему:

1		2		3		4	
ABCD	*P*	*ABCD*	*P*	*ABCD*	*P*	*ABCD*	*P*
0 1 0 1	1	1 0 0 0	1	0 0 1 0	1	1 1 1 0	1
0 1 1 1	1	1 1 0 1	0	1 0 1 0	0	1 1 0 0	1
1 0 1 0	1	1 0 0 1	1	0 1 1 0	1	0 0 1 1	1
0 0 1 1	0	0 0 1 0	0	1 1 1 0	1	0 1 0 1	1
0 1 0 0	0	1 0 1 0	1	0 0 0 1	0	0 0 0 0	0
1 0 0 0	0	0 1 1 1	0	1 0 0 1	1	1 0 1 0	0
				0 1 0 1	0	0 0 0 1	0
				1 1 0 1	1	1 1 0 1	0

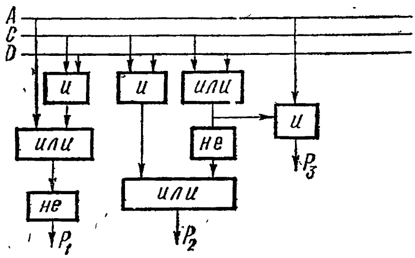

Рис. 3.29

Ответы. 1) $P = AC \vee BD$; 2) $P = \overline{A}B$; 3) $P = AD \vee BC \vee \overline{A}C$; 4) $P = \overline{A}(B \vee C) \vee B\overline{D}$.

Литература к разделу II

1. Бекетов В. С. и др. Основы вычислительной техники. Киев, 1968.
2. Бекетов В. С. и др. Сборник задач по курсу «Основы вычислительной техники». Мн., 1968.
3. Блох А. Ш. Синтез переключательных схем. Мн., 1966.
4. Вавилов Е. Н., Портной Г. П. Элементы теории электронных цифровых машин. Киев, 1960.
5. Гаврилов М. А. Теория релейно-контактных схем. М., 1950.
6. Гильберт Д., Аккерман В. Основы теоретической логики. М., 1947.
7. Глушков В. М. Синтез цифровых автоматов. М., 1962.
8. Драбкина М. Е. Логические упражнения по элементарной математике. Мн., 1965.
9. Дьяченко В. Ф., Лазарев В. Г., Саввин Г. Г. Управление на сетях связи. М., 1967.

10. Ефимов Н. В. Высшая геометрия. М., 1961.
11. Китов А. И., Криницкий Н. А. Электронные цифровые машины и программирование. М., 1961.
12. Клини С. К. Введение в метаматематику. М., 1957.
13. Колужин Л. А. Об алгоритмизации математических задач. В сб.: Проблемы кибернетики, вып. 2. М., 1959.
14. Ляпунов А. А. О логических схемах программ. В сб.: Проблемы кибернетики, вып. 1. М., 1958.
15. Ляпунов А. А., Яблонский С. В. Теоретические проблемы кибернетики. В сб.: Проблемы кибернетики, вып. 9. М., 1963.
16. Молодший В. Н. Очерки по вопросам обоснования математики. М., 1958.
17. Неверов Г. С. Элементы математической теории электронных цифровых вычислительных машин. Мн., 1962.
18. Новиков П. С. Элементы математической логики. М., 1959.
19. Парамонов В. С., Чеботарев Н. А. Основы вычислительной техники. Харьков, 1966.
20. Применение логики в науке и технике (сб. ст.). М.
21. Рвачев В. Л. Геометрические приложения алгебры логики. Киев, 1967.
22. Черч А. Введение в математическую логику. М., 1960.
23. Шиханович Ю. А. Введение в современную математику. М., 1965.
24. Яглом И. М. Необыкновенная алгебра. М., 1968.
25. Янов Ю. И. О логических схемах алгоритмов. В сб.: Проблемы кибернетики, вып. 1. М., 1958.

ОГЛАВЛЕНИЕ

Раздел II

ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ ЛОГИКИ

Колобов Александр Михайлович,
Неверов Георгий Степанович

Избранные главы высшей математики. Ч. 3. Методы математической физики (дифференциальные уравнения в частных производных второго порядка). Элементы математической логики

Редактор Т. Майборода
Худож. редактор В. Валентович
Техн. редактор Г. Романчук
Корректор Н. Бондаренко

АТ 07262. Сдано в набор 16/II 1971 г. Подписано к печати 28/VI 1971 г. Бумага $60 \times 90^1/_{16}$ типогр. № 3. Печ. л. 19,5. Уч.-изд. л. 19,75. Изд. № 69-60. Тип. зак. 80. Тираж 9000. Цена 68 коп.

Издательство «Вышэйшая школа» Государственного комитета Совета Министров БССР по печати. Редакция физико-математической литературы. Минск, ул. Кирова, 24.

Ордена Трудового Красного Знамени типография издательства ЦК КП Белоруссии, Минск, Ленинский проспект, 79.

К61 **Колобов А. М. и Неверов Г. С.**

Избранные главы высшей математики. Ч. 3. Методы математической физики (дифференциальные уравнения в частных производных второго порядка). Элементы математической логики. Мн., «Вышэйш. школа», 1971.

311 с. с илл.

Учебное пособие для радиотехнических, энергетических и машиностроительных специальностей втузов по высшей математике. Адресовано прежде всего лицам, самостоятельно изучающим курс высшей математики. Содержит большое число задач с подробным разъяснением методов их решения, вопросы и задачи для повторения. Как и в предыдущих двух частях книги под аналогичным названием, намечены пути по использованию изучаемых математических методов и понятий в радио- и электротехнике, автоматике.

$\frac{2\text{-}2\text{-}3}{6\text{-}71}$ 517

68 к.