

Пол БАМБЕРГ  
Шломо СТЕРНБЕРГ

КУРС МАТЕМАТИКИ  
для  
СТУДЕНТОВ – ФИЗИКОВ

Том II



Предлагаемый курс лекций известных профессоров Гарвардского университета П. Бамберга и Ш. Стернберга пользуется большой популярностью среди физиков и инженеров, считающих для себя необходимым освоить основные элементы современной математики.

Помимо традиционных вопросов этот курс содержит изложение основ топологии и геометрии, топологический смысл уравнений Максвелла, комплексный анализ, асимптотический метод Лапласа и начала термодинамики. Объединение всех этих разделов математики и физики в одной монографии, несомненно, полезно для студентов и аспирантов российских университетов и позволит им быстро включиться в проходящие во всём мире исследования в теоретической и математической физике.

*Я. Г. Синай*  
академик РАН



Я. Г. Синай  
на ступенях МГУ  
(фотс



и В. И. Арнольд  
в сентябре 1963 года  
(J. Moser)

История преподавания математического анализа, начиная с его изобретения Ньютоном, изобиловала крутыми поворотами. К середине XX века традиционные изложения стали отличать полное исключение всех геометрических понятий и представлений, с одной стороны, и полный отрыв от физики и всех приложений, с другой.

Предлагаемая книга знаменует возвращение преподавания математики к естественнонаучному подходу, для которого характерны явная геометричность и прямое представление математики как части физики, непосредственно пригодной для многочисленных приложений как в естествознании, так и в инженерном деле.

Это возвращение математики в русло естественных наук происходит сейчас и на уровне новых научных исследований, а его проявление в преподавании давно стало абсолютно необходимым.

*В. И. Арнольд*  
академик РАН

Пол БАМБЕРГ  
Шломо СТЕРНБЕРГ

# КУРС МАТЕМАТИКИ

для

# СТУДЕНТОВ-ФИЗИКОВ

Том II

Главы 12–22

ПЕРЕВОД С АНГЛИЙСКОГО  
К. Ф.-М. Н., доц. Л. П. Котовой (МФТИ)

ПОД РЕДАКЦИЕЙ  
К. Ф.-М. Н., доц. А. Н. Якивчика (МГУ)



МОСКВА  
2006

УДК 51



Издание осуществлено при поддержке  
Центрального аэрогидродинамического  
института имени Н. Е. Жуковского,  
директор В. Г. Дмитриев  
и  
Фонда «КНИГА–НАУКА–КУЛЬТУРА»,  
директор В. Б. Филиппов

Бамберг П., Стернберг Ш.  
**Курс математики для студентов-физиков. Том II, главы 12–22.**  
Перев. с англ. Л. П. Котовой под ред. А. Н. Якивчика.  
— М.: ФАЗИС, 2006. 664 с. (XIX–XXXVI, 575–1220)  
ISBN 5-7036-0110-X

Paul Bamberg & Shlomo Sternberg  
**A Course in Mathematics for Students of Physics**  
© Cambridge University Press 1990 (Vol. 2)

Издательство ФАЗИС  
123557 Москва, Пресненский вал, 42-44  
e-mail: publisher@phasis.ru <http://www.phasis.ru>

-----  
Типография «Наука» Академиздатцентра РАН  
121099 Москва Г-99, Шубинский пер., 6  
Заказ № 3639

ISBN 5-7036-0110-X

© ФАЗИС, 2006

---

## Оглавление тома II

---

<b>Предисловие</b> .....	<b>xxv</b>
<b>От издателя русского перевода</b> .....	<b>xxxv</b>
<b>Глава 12. Теория электрических цепей</b> .....	<b>575</b>
Введение .....	575
12.1. Линейные цепи с резисторами .....	580
12.2. Топология одномерных комплексов .....	590
12.3. Коцепи и $d$ -оператор .....	604
12.4. Базисы и дуальные базисы .....	607
12.5. Методы Максвелла .....	610
12.6. Матричные выражения для операторов .....	614
12.7. Теорема Кирхгофа .....	626
12.8. Цепи переменного тока в стационарном режиме и фильтры .....	630
Резюме .....	636
Задачи .....	637
<b>Глава 13. Метод ортогональных проекций</b> .....	<b>645</b>
13.1. Метод ортогональных проекций Вейля .....	645
13.2. Метод Кирхгофа .....	650
13.3. Теорема взаимности Грина .....	656
13.4. Цепи с конденсаторами .....	661
13.5. Краевые задачи .....	667
13.6. Решение краевых задач методом ортогональных проекций Вейля .....	671
13.7. Функции Грина .....	678
13.8. Ядро Пуассона и случайное блуждание .....	683
13.9. Теорема взаимности Грина в электростатике .....	686
Резюме .....	692
Задачи .....	693
<b>Глава 14. Комплексы высших размерностей</b> .....	<b>705</b>
14.1. Комплексы и гомологии .....	705
14.2. Дуальные пространства и когомологии .....	728

Резюме .....	737
Задачи .....	737
<b>Глава 15. Комплексы в пространстве <math>\mathbb{R}^n</math> .....</b>	<b>745</b>
Введение .....	745
15.1. Внешняя алгебра .....	749
15.2. $k$ -формы и оператор $d$ .....	757
15.3. Интегрирование $k$ -форм .....	760
15.4. Теорема Стокса .....	777
15.5. Дифференциальные формы и когомологии .....	792
Резюме .....	805
Задачи .....	806
<b>Глава 16. Электростатика в пространстве <math>\mathbb{R}^3</math> .....</b>	<b>817</b>
16.1. От дискретности к непрерывности .....	817
16.2. Граничный оператор .....	820
16.3. Телесный угол .....	822
16.4. Напряженность электрического поля и электрическое смещение .....	824
16.5. Диэлектрическая проницаемость .....	836
16.6. Звездный оператор в трехмерном евклидовом пространстве .....	837
16.7. Формулы Грина .....	842
16.8. Гармонические функции .....	844
16.9. Метод ортогональной проекции .....	848
16.10. Функции Грина .....	850
16.11. Интегральная формула Пуассона .....	853
Резюме .....	859
Задачи .....	859
<b>Глава 17. Токи, потоки и магнитостатика .....</b>	<b>863</b>
17.1. Токи .....	863
17.2. Потоки и векторные поля .....	865
17.3. Внутреннее произведение .....	871
17.4. Производные Ли .....	879
17.5. Магнетизм .....	883
Приложение: альтернативное доказательство фундаментальной формулы дифференциального исчисления .....	890
Резюме .....	893
Задачи .....	893
<b>Глава 18. Звездный оператор .....</b>	<b>897</b>
18.1. Скалярные произведения и внешняя алгебра .....	897
18.2. Звездный оператор .....	901
18.3. Интеграл Дирихле и лапласиан .....	909
18.4. Оператор $\square$ в четырехмерном пространстве-времени .....	916
18.5. Алгебра Клиффорда .....	920
18.6. Звездный оператор и геометрия .....	931
18.7. Звездный оператор и векторный анализ .....	933
Приложение: тензорные произведения .....	936

Резюме .....	953
Задачи .....	953
<b>Глава 19. Уравнения Максвелла .....</b>	<b>969</b>
19.1. Уравнения .....	969
19.2. Одномерное однородное волновое уравнение .....	974
19.3. Однородное волновое уравнение в пространстве $\mathbb{R}^3$ .....	977
19.4. Неоднородное волновое уравнение в пространстве $\mathbb{R}^3$ .....	981
19.5. Лагранжиан электромагнитного поля и тензор энергии-импульса .....	983
19.6. Волновые фронты и принцип Гюйгенса .....	988
Резюме .....	994
Задачи .....	994
<b>Глава 20. Комплексный анализ .....</b>	<b>997</b>
Введение .....	997
20.1. Комплекснозначные функции .....	999
20.2. Комплекснозначные дифференциальные формы .....	1001
20.3. Голоморфные функции .....	1004
20.4. Вычисление вычетов .....	1010
20.5. Приложения и следствия .....	1021
20.6. Локальное отображение .....	1031
20.7. Контурные интегралы .....	1040
20.8. Пределы и ряды .....	1047
Добавление: связь между голоморфными и гармоническими функциями .....	1052
Резюме .....	1055
Задачи .....	1056
<b>Глава 21. Асимптотическое вычисление интегралов .....</b>	<b>1065</b>
Введение .....	1065
21.1. Метод Лапласа .....	1066
21.2. Метод стационарной фазы .....	1072
21.3. Гауссовы интегралы .....	1076
21.4. Групповая скорость .....	1080
21.5. Формула обратного преобразования Фурье .....	1082
21.6. Асимптотическая оценка формулы Гельмгольца .....	1084
Резюме .....	1086
Задачи .....	1087
<b>Глава 22. Термодинамика .....</b>	<b>1089</b>
22.1. Теорема Каратеодори .....	1090
22.2. Классическая термодинамика по Борну и Каратеодори .....	1100
22.3. Энтропия и абсолютная температура .....	1108
22.4. Системы с одной конфигурационной переменной .....	1115
22.5. Условия равновесия .....	1131
22.6. Системы и состояния в статистической механике .....	1137
22.7. Произведения и образы .....	1144

22.8. Наблюдаемые, ожидаемые значения и внутренняя энергия .....	1149
22.9. Энтропия .....	1158
22.10. Равновесие в статистических системах .....	1161
22.11. Квантовые и классические газы .....	1171
22.12. Определители и следы .....	1175
22.13. Квантовые состояния и квантовая логика .....	1182
Резюме .....	1188
Задачи .....	1189
Приложение: доказательство теоремы о нормальной форме .....	1192
<b>Рекомендуемая литература .....</b>	<b>1201</b>
<b>Предметный указатель .....</b>	<b>1207</b>

---

## Предисловие

---

В этой книге излагается курс математики, который мы преподавали в Гарварде в течение восьми лет. Курс предназначен для студентов, интересующихся физикой и имеющих хорошую подготовку по математическому анализу функций одной переменной. Полезно некоторое знакомство с линейной алгеброй, но это не обязательно. Большинство наших студентов одновременно с этим курсом изучают сложные курсы по физике, так что они могут соединить излагаемые здесь сведения со своими физическими знаниями. Такое совмещение полезно, но не обязательно. Основное содержание нашего курса — теория и физические приложения линейной алгебры и математического анализа функций нескольких переменных, включая внешнее исчисление. Наш педагогический метод — изложение по «спирали», т. е. мы рассматриваем одну и ту же тему несколько раз, возвращаясь к ней на все более высоком уровне, одновременно расширяя области ее применения. Такое построение мы предпочитаем «прямолинейному» строгому логическому порядку. Надеемся, что при этом мы избежим логических ошибок типа «порочного круга», но часто будем начинать с простого случая того или иного понятия, а затем возвращаться к нему для обобщений, делая это только после ознакомления с родственными понятиями и, как следствие, приобретения более широкого взгляда на предмет. Такой подход требует от студентов определенного доверия и терпения. Но мы надеемся, что, в конце концов, они будут вознаграждены более глубоким интуитивным пониманием предмета в целом.

Теперь расскажем о содержании книги более детально. Цель первых четырех глав — познакомить читателя с алгеброй и анализом квадратных матриц. Таким образом, изучив эти главы, студент должен научиться думать о матрицах как о самостоятельных объектах, а не просто как о наборе чисел, выстроенных по квадрату. В этих главах мы рассматриваем почти исключительно матрицы  $2 \times 2$ , когда самые сложные вычисления сводятся к решению квадратных уравнений. Но каждый раз мы формулируем результаты, имея в виду матрицы более высокого порядка.

Глава 1 начинается с объяснения связи закона умножения матриц  $2 \times 2$  и геометрии прямых линий на плоскости. Далее излагается алгебра матриц  $2 \times 2$ , обсуждается детерминант и его отношение к плоскости и ориентации. Мы вводим определение абстрактного векторного пространства в общем случае, объясняем понятие базиса и замены базиса в одномерном и двумерном векторном пространстве.

В главе 2 мы обсуждаем конформную линейную геометрию на плоскости, т.е. геометрию линий и углов и их связь с определенным видом матрицами  $2 \times 2$ . Кроме того, вводятся понятия собственных значений и собственных векторов, столь важные в квантовой механике. Мы пользуемся этими понятиями, чтобы построить алгоритм вычисления степеней матрицы. В качестве примера, где это применяется, обсуждаются основные свойства цепей Маркова.

Основная задача главы 3 — показать, что система однородных линейных дифференциальных уравнений с постоянными коэффициентами может быть записана в виде  $du/dt = Au$ , где  $A$  — матрица, а  $u$  — вектор, и что решение этой системы имеет вид  $e^{At}u_0$ , где  $u_0$  задается начальными условиями. При этом, конечно, требуется объяснить, что означает экспонента, в показателе которой стоит матрица. Кроме того, описывается качественное поведение решений в неоднородном случае, в том числе обсуждается явление резонанса.

Глава 4 посвящена изучению скалярных произведений и квадратичных форм. В ней имеется много физических приложений, в том числе обсуждение нормальных форм и детальное изложение специальной теории относительности.

В главах 5 и 6 изложены основные положения дифференциального исчисления.

В главе 5 определяется дифференциал отображения одного векторного пространства на другое, обсуждаются его основные свойства, в частности, правило дифференцирования композиции. В качестве физических приложений рассматриваются задача Кеплера и приближение Борна. Мы даем определения частной производной и производной по направлению, определяем линейные дифференциальные формы.

В главе 6 продолжается изложение дифференциального исчисления. Мы представляем в векторной форме теорему о среднем значении, формулу Тейлора и теорему об обратном отображении. Обсуждаем поведение критических точек и множители Лагранжа.

Главы 7 и 8 знакомят читателя с интегральным исчислением.

Глава 7 посвящена изучению линейных дифференциальных форм и их интегралов. Особое внимание уделено замене переменных. Кроме того, обсуждаются и другие одномерные интегралы, например, длина дуги.

Глава 8 посвящена изучению внешних форм второй степени и соответствующих им двумерных интегралов. Вводится понятие внешней производной и подчеркивается инвариантность относительно замены переменных. Доказывается двумерный вариант теоремы Стокса, т. е. теорема Грина. Рассматриваются поверхностные интегралы в трехмерном пространстве.

В главе 9 результаты первых восьми глав используются при построении физической теории — оптики. Вся глава посвящена приложениям, и ее можно опустить без ущерба для понимания последующего материала.

В главе 10 мы возвращаемся назад и доказываем основные свойства конечномерных векторных пространств и их линейных преобразований. Обычно это непосредственное обобщение результатов, полученных в первых четырех главах в двумерном случае. Но при этом дан новый алгоритм редукции матрицы по строкам. Вводятся два новых понятия — дуальное пространство и факторпространство (в начале курса их было трудно определить). В дальнейшем эти понятия окажутся чрезвычайно важными.

Глава 11 посвящена свойствам детерминантов матриц  $n \times n$ . Тема рассмотрена аксиоматически, показаны основные вычислительные алгоритмы.

Главы 12–14 знакомят читателя с математикой формы, т. е. с алгебраической топологией.

В главе 12 мы начинаем изучение электрических цепей. При этом рассматриваются два аспекта. Во-первых, изучаются «провода» цепи, т. е. как соединяются между собой различные ветви. На математическом языке это называется топологией одномерных комплексов. Во-вторых, мы изучаем реакцию цепи в целом, если нам известно поведение отдельных ветвей, в частности изменение мощности и энергии. Мы приводим несколько примеров цепей, интересных с физической точки зрения.

В главе 13 мы продолжаем изучение цепей. Исследуем крайние задачи, связанные с емкостными цепями, и применяем изложенные методы для решения ряда классических задач электростатики, включающих проводники.

В главе 14 мы кратко показываем, как результаты, полученные в главах 12 и 13 в одномерном случае, обобщаются на случай большего числа измерений.

Главы 15–18 посвящены разработке внешнего дифференциального исчисления, рассматриваемого как непрерывная версия дискретной теории комплексов.

В главе 15 излагаются основные сведения о внешнем исчислении: внешняя алгебра,  $k$ -формы, замена переменных, внешняя производная и теорема Стокса.

Глава 16 посвящена электростатике. Мы рассматриваем диэлектрические свойства вакуума как непрерывный аналог емкости цепи, причем эти диэлектрические свойства определяют евклидову геометрию в трехмерном случае. Излагаем основные положения теории потенциалов.

В главе 17 мы продолжаем изучение внешнего дифференциального исчисления. Основные темы: векторные поля и потоки, внутренние произведения и производные Ли. Полученные результаты применены для решения задач магнитостатики.

Глава 18 заканчивает изучение внешнего исчисления углубленным обсуждением общего случая звездного оператора.

Главу 19 можно считать кульминацией всего курса. В ней результаты, полученные в предыдущих главах, применяются для исследования уравнений Максвелла и связанных с ними волновых уравнений.

Главы 20 и 21 никак не связаны с главами 9–19, и их можно читать независимо. Обычно они не включаются в наш годовой курс. А вот главы 1–9, 20 и 21 могут составить самостоятельный краткий курс математики.

Материал главы 20 представляет собой относительно стандартное изложение теории функций комплексного переменного, которое вполне доступно для студентов, подготовленных к чтению этой книги.

В главе 21 обсуждаются простейшие аспекты асимптотик.

В главе 22 показано, как, следуя идеям Бора и Каратеодори, внешнее исчисление можно использовать в классической термодинамике.

Большинство математиков и все физики, результаты которых представлены в этой книге, работали в первом десятилетии двадцатого века. Это значит, что изложенному материалу уже по крайней мере 90 лет. И, тем не менее, он еще не вошел в наши элементарные курсы математики, хотя большая его часть (за исключением, может быть, теории цепей) должна изучаться современными физиками, что они и делают на определенном этапе своей карьеры. Причины этого, в значительной мере, исторические. У. Гамильтону (1805–1865) было очевидно, что вещественных и комплексных чисел недостаточно для глубокого изучения геометрического анализа, что в декартовой геометрии дву- и трехмерного пространства возникает желание рассматривать числа парами и тройками, как самостоятельные объекты с собственными алгебраическими свойствами. Для этого он разработал алгебру кватернионов, которая была очень популярна в Англии в середине девятнадцатого века. Однако, теория имела ряд изъянов: кватернионы более естественны в пространстве четырех, а не трех измерений. Геометрия трех измерений возникала как часть более общей конструкции, а не как самостоятельная теория. Кроме того, кватернионы имели слишком сложную алгебраическую структуру. Например, очень громоздкой была связь между

умножением кватернионов и геометрическими построениями в трехмерном пространстве. (Конечно, первый недостаток сегодня считался бы не очень серьезным, но вместо него возникло бы другое возражение: теория *ограничена* четырьмя измерениями.) В конце концов, *векторная алгебра* в трехмерном пространстве с ее скалярным и векторным произведениями была очищена от теории кватернионов. Ее объединили с необходимыми дифференциальными операциями, в результате чего появился *векторный анализ*, окончательно разработанный Дж. Гиббсом (1839–1903) и обнародованный им в своей знаменитой книге «Элементы векторного анализа», оказавшей сильное влияние на дальнейшее развитие науки.

Итак, векторный анализ с его операциями  $\text{grad}$ ,  $\text{div}$ ,  $\text{curl}$  и т. д. стал стандартным языком изложения геометрических законов физики. Однако, несмотря на то, что векторный анализ прекрасно описывает геометрию трехмерного пространства, он тоже имеет ряд существенных недостатков. Первый — и наименее серьезный — состоит в том, что в нем не просматривается очень важное единство: например, не ясно, что фундаментальная теорема исчисления одна — сейчас ее называют общей теоремой Стокса, а теоремы Грина, Гаусса и классическая теорема Стокса являются ее частными случаями. В рамках векторного анализа мы не можем это увидеть. Более серьезный недостаток состоит в том, что фундаментальные операторы (например,  $\text{grad}$ ,  $\text{div}$ ) предполагают евклидову структуру пространства, т. е. подразумевают трехмерное пространство и ориентацию (например,  $\text{curl}$ ). Это значит, что теория работает только в трехмерном ориентированном евклидовом пространстве. При этом возникает еще одна проблема — при произвольном изменении координат операторы ведут себя неприятно, их выражения в неортогональных координатах очень громоздки. Уже А. Пуанкаре (1854–1912) в своих фундаментальных научных и философских работах, предшествовавших теории относительности, подчеркивал необходимость делать различие между «топологическими» законами геометрии и физики, т. е. законами, которые зависят только от дифференциальной структуры пространства и поэтому инвариантны относительно гладких преобразований, и законами, зависящими от геометрии

ческой структуры, например, от понятия расстояния. Теория относительности оказала огромное влияние на развитие математики главным образом потому, что она поощряла изучение теории многомерных пространств, которая уже существовала в математической литературе, но не считалась важной для изучения геометрии. Кроме того, необходимо было исследовать произвольные изменения координат. Векторный анализ не мог решить эти две задачи и поэтому позднее его дополнили *тензорным анализом*. Но и тензорный анализ с его обилием индексов тоже имеет много серьезных недостатков. Главный изъян состоит в том, что вероятно трудно сказать, какие операции имеют общее геометрическое значение, а какие связаны только с данной системой координат. Таким образом, тензорный анализ очень хорош для вычислений, но трудно сказать, что именно вычисляется. Вся работа, начатая еще Гамильтоном и направленная на создание математического аппарата, в котором объекты имеют геометрический смысл, оказалась напрасной. Для того, чтобы заставить теорию работать, необходимо было ввести аффинную связь — довольно изощренную геометрическую конструкцию. Но даже с такой конструкцией геометрическое значение операций не стало прозрачным. По существу интуитивно ясный векторный анализ так и не был вытеснен тензорным анализом из программ начального университетского обучения.

Математики в основном уже согласились, а физики постепенно привыкают к тому, что самым удобным аппаратом для геометрического анализа является внешнее дифференциальное исчисление Грассмана и Картана. Его преимущество состоит в том, что наряду с простыми и четкими правилами вычисления все объекты имеют прозрачный геометрический смысл. Внешнее исчисление работает при любом числе измерений, хорошо себя ведет при отображениях и изменениях системы координат, обладает существенным единством основных теорем и четко различает «топологические» и «метрические» свойства. В рамках внешнего анализа геометрические законы физики принимают простую и элегантную форму. Чтобы это продемонстрировать, полезно привести таблицу, взятую из книги «Курс математической физики», написанную Тиррингом.

Таблица 1. История записи уравнений Максвелла.

Константы  $c$ ,  $\mu_0$ , и  $\epsilon_0$  положены равными 1.

Однородное уравнение	Неоднородное уравнение
Самая ранняя форма	
$\frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0$ $\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z} = -\dot{B}_x$ $\frac{\partial E_x}{\partial z} - \frac{\partial E_z}{\partial x} = -\dot{B}_y$ $\frac{\partial E_y}{\partial x} - \frac{\partial E_x}{\partial y} = -\dot{B}_z$	$\frac{\partial E_x}{\partial x} + \frac{\partial E_y}{\partial y} + \frac{\partial E_z}{\partial z} = \rho$ $\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z} = j_x + \dot{E}_x$ $\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x} = j_y + \dot{E}_y$ $\frac{\partial B_y}{\partial x} - \frac{\partial B_x}{\partial y} = j_z + \dot{E}_z$
Конец XIX века	
$\nabla \cdot \mathbf{B} = 0$ $\nabla \times \mathbf{E} = -\dot{\mathbf{B}}$	$\nabla \cdot \mathbf{E} = \rho$ $\nabla \times \mathbf{B} = \mathbf{j} + \dot{\mathbf{E}}$
Начало XX века	
$*F^{\beta\alpha}_{,\alpha} = 0$	$F^{\beta\alpha}_{,\alpha} = j^\beta$
Середина XX века	
$dF = 0$	$\delta F = J$

В 1844 году Г. Грассман (1809–1877) опубликовал книгу «Ausdehnungslehre». Математическое сообщество не оценило эту работу и ведущие немецкие математики того времени проигнорировали ее. Фактически, Грассман так никогда и не получил место в университете; всю свою жизнь он преподавал математику в школе. Однако, похоже, что он прожил счастливую и плодотворную жизнь. У Грассмана была большая семья и он считался признанным экспертом по санскритской литературе. В конце жизни он попробовал еще раз выпустить новое издание «Ausdehnungslehre», но оно также не имело успеха. Толь-

ко немногие из математиков того времени оценили его работу. Среди них был К. Мебиус (1825–1908). Тем не менее, в книге «Ausdehnungslehre» впервые появились ключевые в современной математике понятия и большинство алгебраических структур, используемых в этой книге. В ней были определены: векторное пространство, внешняя алгебра, внешнее и внутреннее произведение, теорема Стокса в обобщенной форме.

В настоящее время общепризнано, что Э. Картан (1869–1951) является крупнейшим специалистом по геометрии в XX веке. Его ранние работы, относящиеся к группам Ли и дифференциальным уравнениям в частных производных, имеющие огромное значение для современной математики, тоже сначала не были поняты. Но в 1920 году они стали известны широкому кругу ученых. К этому времени теория групп Ли начала играть ключевую роль в математике и физике, и Г. Вейль заново представил их в своих статьях. В настоящее время работы Картана по основным узлам и связям (известные под общим названием «калибровочные теории») являются основой теории элементарных частиц. В 1922 году Картан опубликовал книгу «Лекции по интегральным инвариантам», где показал, что изобретенное им внешнее дифференциальное исчисление является гибким инструментом не только для геометрии, но и для вариационного исчисления с широким кругом физических приложений. Как уже было сказано, признание пришло не сразу, но сегодня и физики, и математики считают, что внешнее исчисление позволяет наилучшим образом формулировать геометрические законы физики. И нам кажется, что пришла пора заменить им векторное исчисление в программах начального университетского обучения.

Необходимо пояснить, почему в книге так много внимания уделено теории электрических цепей, хотя обычно эта тема не входит в учебные программы по математике. Прежде всего, это оправдано с педагогической точки зрения. Данная тема всегда нравится студентам. Она позволяет на очень приземленном уровне ввести понятия дуального пространства и факторпространства. Иначе эти понятия кажутся студентам слишком абстрактными и поэтому трудно воспринимаются ими. Кроме того, в дискретном, алгебраическом подходе к теории цепей теорема

Стокса появляется, по существу, как естественное определение, что позволяет понять роль оператора  $d$  и теоремы Стокса при формулировке внешнего исчисления. Конечно, есть более глубокие, философские причины, почему мы приняли решение подробно рассказать о теории цепей. Уже почти сто лет, как люди думают, что силы, удерживающие макроскопические тела, имеют электрическую природу. Таким образом, понятие твердого тела и отсюда евклидова геометрия вытекают из электростатики (конечно, в нерелятивистском приближении, когда понятия твердого тела и геометрия Евклида имеют смысл). Физические исследования как предельно малых объектов (элементарных частиц), так и предельно больших (космология) заново ставят фундаментальные вопросы о геометрии пространства и времени. Мы полагаем, что разумно уже в самом начале изучения математики дать студентам представления о связи геометрии и физики. Без сомнения, сторонники компьютеров и ряда современных теорий в физике зададут вопросы о дискретности, противопоставляемой непрерывному характеру пространства и времени (проблема, поднятая Риманом в его диссертации об основах геометрии). Надеемся, что наше обсуждение может быть полезно тем, кто в будущем станет решать эту проблему.

Наш курс математики читается в течение одного года, и поэтому мы вынуждены были опустить некоторые важные темы. Мы не обсуждаем бесконечномерные векторные пространства, в частности, гильбертовы пространства; не даем определения и не изучаем абстрактные дифференцируемые многообразия и их свойства. Наш опыт показал, что эти темы слишком сложны для студентов, и усилия, которые необходимо затратить на их объяснение, лучше использовать по-другому. Конечно, в книге есть места, где нам приходится компенсировать отсутствие этих понятий. Однако, более серьезно то, что в книге нет подробного обсуждения гармонического анализа Фурье, классической механики и теории вероятностей. Мы только слегка касаемся этих тем, не давая подробного изложения. Нас может извинить только то, что каждая из этих тем потребовала бы семестрового курса, а их подробное изложение на современном уровне легко найти. В конце книги приводится список рекомендуемой литературы.

Нам приятно поблагодарить профессора Даниэла Гороффа, тщательно прочитавшего рукопись, за многочисленные исправления и полезные предложения. Мы искренне благодарим Джин Моррис за отличное качество набора, за то, что она внимательно сопровождала производство книги на протяжении восьми лет — с самого начала проекта до его завершения.

## От издателя и редактора русского перевода

Данный курс является примером глубокого синтеза математических и физических знаний и методов. По нашему мнению, хотя он называется «Курс математики для студентов-физиков», с тем же успехом его можно рассматривать как «Курс физики для студентов-математиков».

Первый том содержит изложение основополагающих математических понятий и теорий, таких как дифференциальное и интегральное исчисление, векторные пространства и линейные преобразования, линейные дифференциальные уравнения. На этом этапе в качестве области применения изучались теория колебаний, специальная теория относительности и гауссова оптика.

Во втором томе математический аппарат не только заметно развивается, но и находит важные физические приложения. Прежде всего это электромагнитные явления — от теории электрических цепей (глава 12) до уравнений Максвелла (глава 19). Термодинамике и статистической механике посвящена глава 22.

То обстоятельство, что курс рассчитан всего на один год, но при этом затрагивает самые разные разделы математики, которые на математических специальностях российских университетов преподаются на протяжении большей части срока обучения,

вынуждает авторов использовать местами весьма конспективный стиль изложения. Это выражается в том, что опускаются детали строгих математических доказательств важных результатов (на которые, впрочем, обычно дается ссылка). А некоторые области, как теория вероятностей или алгебраическая топология, вообще представлены только начальными сведениями. Построение курса также не в полной мере соответствует российской университетской программе. Так, ряды Фурье (без которых трудно представить решение многих современных задач физики и других естественных наук) вводятся только в конце главы 20, посвященной комплексному анализу. (Учитывая важную роль гармонических функций в физике, в частности, как потенциалов силовых полей, редактор сделал к той же 20-й главе небольшое добавление о взаимосвязи между гармоническими функциями двух переменных и аналитическими функциями комплексного переменного.)

При переводе и редактировании были исправлены многочисленные ошибки и неточности оригинала. Редакторы (первого и второго томов) добавили к тексту довольно много примечаний. Обновлено и уточнены списки рекомендуемой литературы, даны ссылки на имеющиеся русские переводы. Заново составлен предметный указатель; во втором томе он охватывает материал всей книги.

Мы надеемся, что этот современный курс, весьма отличающийся от отечественных образовательных традиций, будет интересен не только студентам, но и преподавателям.

Надеемся также, что издание будет полезно широкому кругу научных работников разнообразных физических, технических и математических специальностей.

Мы будем благодарны читателям за любые замечания по книге; просим направлять их нам по адресу [publisher@phasis.ru](mailto:publisher@phasis.ru).

Москва, май 2006

## Глава 12

---

# Теория электрических цепей

---

Главы 12–14 можно считать введением в математику формы, т. е. в алгебраическую топологию. В главе 12 мы начинаем изучение электрических цепей. Здесь имеются два аспекта. Во-первых, изучается система проводов цепи с точки зрения связи между собой различных участков. На математическом языке это называется топологией одномерных комплексов. Во-вторых, предполагая, что мы знаем поведение отдельных участков, изучается реакция цепи в целом, в частности, рассматриваются мощность и энергия. Приводятся примеры цепей, интересные с физической точки зрения.

## Введение

Теория электрических цепей является приближением к электромагнитной теории, когда предполагается, что все интересные явления происходят в проводах и других элементах этих цепей. При этом также предполагается, что цепь может быть разложена на различные компоненты с определенным характером поведения, и задача состоит в том, чтобы предсказать поведение всей цепи, если ее компоненты соединяются различными способами.

Основные переменные теории цепей хорошо известны по бытовым приборам: это ток, напряжение и мощность. Фундаментальной единицей электромагнитной теории является заряд

---

электрона. К моменту написания этой книги (до обнаружения кварков) частица, заряд которой был бы долей заряда электрона, неизвестна. На практике единицей заряда считается кулон, который противоположен заряду  $6.24 \times 10^{18}$  электронов. В 1819 году Эрстед заметил, что поток электрического заряда создает силу, действующую на магнитную иголку, причем эта сила пропорциональна скорости потока заряда. Этот эффект измерять гораздо легче, чем электростатические силы, которые требуются для измерения заряда. Поэтому именно *ток*, а не заряд, считается основной переменной теории цепей. Единицей тока является ампер, причем  $1 \text{ ампер} = 1 \text{ кулон/секунду}$ . Ток будем обозначать буквой  $I$ . Итак,

$$I = \frac{dQ}{dt},$$

где  $I$  — ток, измеряемый в амперах,  $Q$  — заряд, измеряемый в кулонах, а  $t$  — время, измеряемое в секундах. Вообще говоря, мы могли бы ожидать, что ток, протекающий по цепи, зависит от точки наблюдения. В теории цепей предполагается, что ток принимает определенное значение на каждом линейном участке (ветви) цепи (но может и меняться со временем). Пусть  $\alpha$  обозначает некую ветвь в цепи, тогда  $I_\alpha(t)$  обозначает ток, протекающий через  $\alpha$  в момент  $t$ .

Проходя вдоль участка цепи, элемент заряда приобретает энергию или теряет ее. Изменение энергии (измеряемой в джоулях) на единицу заряда (измеряемого в кулонах) называется *напряжением*. Это значит, что  $1 \text{ вольт} = 1 \text{ джоуль/кулон}$ . Напряжение будем обозначать буквой  $V$ . Напряжение (разность потенциалов) между двумя концами ветви  $\alpha$  в момент  $t$  обозначается  $V^\alpha(t)$ . Произведение тока на напряжение имеет размерность энергии/время и называется *мощностью*, измеряемой в соответственнно ваттах, т. е.

$$1 \text{ ватт} = 1 \text{ вольт} \times \text{ампер} = 1 \text{ джоуль/секунду}.$$

В теории цепей предполагается, что существует три типа ветвей: *индукторы*, *конденсаторы* и *сопротивления*. Для каждого типа имеется определенное соотношение между напряжением и током. В индукторе напряжение пропорционально скорости изменения

тока, т. е. если  $\alpha$  обозначает индуктор, то

$$L_{\alpha} \frac{dI_{\alpha}}{dt} = V^{\alpha},$$

где постоянная  $L_{\alpha}$  называется *индуктивностью*. Соответствующая единица индуктивности называется *генри*, т. е.

$$1 \text{ генри} = 1 \text{ вольт} / (\text{ампер} / \text{секунду}) = 1 \text{ вольт} \times \text{секунду} / \text{ампер}.$$

В конденсаторе ток пропорционален скорости изменения напряжения. Если  $\alpha$  — конденсатор, то

$$C^{\alpha} \frac{dV^{\alpha}}{dt} = I_{\alpha},$$

где постоянная  $C^{\alpha}$  называется его *емкостью* и измеряется в *фарадах*, т. е.

$$1 \text{ фарад} = 1 \text{ кулон} / \text{вольт}.$$

Для сопротивлений (резисторов) существует функциональное соотношение, связывающее ток и напряжение (но не их производные). В классической теории *линейных пассивных* цепей оно имеет вид  $V = RI$ , где постоянная  $R$ , измеряемая в омах, называется *сопротивлением* ветви. График<sup>1</sup> этого соотношения на плоскости  $I, V$  изображен на рис. 12.1. Это прямая линия, проходящая

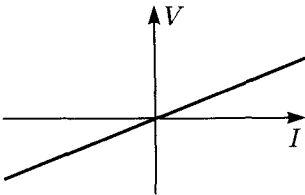


Рис. 12.1. Простой резистор

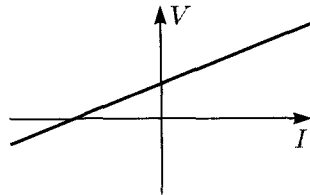


Рис. 12.2. Реальная батарея

через начало координат. Любой прибор, описываемый (неоднородным) линейным уравнением по  $I$  и  $V$ , мы будем считать *линейным сопротивлением*. Например, реальная батарея (рис. 12.2)

<sup>1</sup>называемый вольт-амперной характеристикой. — Прим. ред.

с внутренним сопротивлением  $R^\alpha$  может быть описана уравнением  $V^\alpha = W^\alpha + R^\alpha I_\alpha$ , где постоянная  $W^\alpha$  называется ЭДС (электродвижущей силой) батареи. Идеальная батарея, создающая напряжение  $W^\alpha$  при любом токе, описывается уравнением  $V^\alpha = W^\alpha$ . На рис. 12.3 изображена горизонтальная линия. Идеальный источник тока, создающий ток  $K_\alpha$  независимо от напряжения на его полюсах, описывается уравнением  $I_\alpha = K_\alpha$ , т. е. графиком будет вертикальная линия.

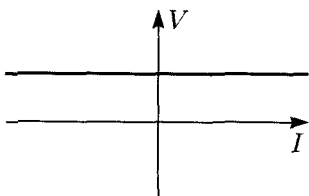


Рис. 12.3. Идеальная батарея

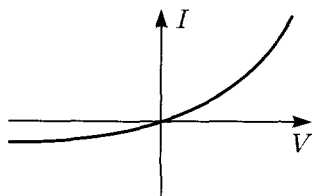


Рис. 12.4

Анализируя цепи, в которых напряжение и токи зависят от времени, мы должны рассматривать источники напряжения и тока, тоже зависящие от времени. В этом случае линейное сопротивление в общем виде описывается уравнением

$$V^\alpha(t) - W^\alpha(t) = R_\alpha(I_\alpha(t) - K_\alpha(t)),$$

где  $W^\alpha(t)$  и  $K_\alpha(t)$  — заданные функции времени.

В более общем случае рассматриваются нелинейные сопротивления, индукторы и конденсаторы. Например, диод или термистор могут считаться *нелинейными сопротивлениями*, описываемыми нелинейной функцией  $I_\alpha = f(V^\alpha)$  типа изображенной на рис. 12.4.

Любое обобщенное сопротивление характеризуется функциональным соотношением  $R(I, V, t) = 0$  с единственным ограничением: в него не входят производные  $I$  и  $V$ .

Обобщая понятие индуктора, можно предположить, что индуктивность зависит от тока  $I$ , и записать

$$V = L(I) \frac{dI}{dt}.$$

Такое, например, может реализоваться в индукторе с железным сердечником, используемом для больших токов. Аналогично можно рассмотреть нелинейный конденсатор

$$I = C(V) \frac{dV}{dt},$$

который получается, если использовать диэлектрик с нелинейным откликом<sup>2</sup>.

Мы еще будем говорить о нелинейных устройствах. Однако важно осознать, что последующая теория, в основном, будет касаться получения уравнений для электрических цепей, а не их решения. Причем эти уравнения будут одинаково применимы как к линейным элементам, так и к нелинейным.

А сейчас предположим, что наша электрическая цепь состоит из  $b$  ветвей трех типов, о которых говорилось выше. Все ветви как-то соединены друг с другом. Мы хотели бы определить токи и напряжения в каждой ветви, т.е. найти  $2b$  неизвестных величин. Для каждой ветви у нас есть только одно уравнение (либо дифференциальное уравнение, либо функциональное соотношение между током и напряжением). Нам надо иметь еще  $b$  уравнений. Эти уравнения известны как *законы Кирхгофа*.

Кирхгоф, будучи студентом на семинаре Нейманна, впервые исследовал проблемы цепей. Его результаты были опубликованы в период с 1845 по 1847 год. Он доказал существование решения проблемы цепей для случая, когда есть только пассивные линейные сопротивления. Занимаясь этой задачей, Кирхгоф первым стал изучать алгебраические свойства формы. Почти пятьдесят лет спустя Пуанкаре создал абстрактную теорию формы, как математическую дисциплину, которую назвал *analysis situs*. Сейчас это называется алгебраической топологией. Природа методов алгебраической топологии наиболее очевидна в случае цепей, состоящих из пассивных линейных сопротивлений. Поэтому сначала

---

<sup>2</sup>Имеется в виду (нетривиальная) зависимость *диэлектрической проницаемости* материала между пластинами конденсатора от напряженности электростатического поля. Строго говоря, при достаточно больших значениях напряженности это явление имеется у всякого диэлектрика, так как наступает его *пробой* (резкий рост диэлектрической проницаемости, а затем переход в состояние проводника). — *Прим. ред.*

мы очень подробно изучим именно эти цепи, прежде чем перейдем к общему случаю. Исследуя такие цепи, мы будем следовать статье Г. Вейля (1923 год), в которой результаты Кирхгофа представлены в форме, прямо показывающей связь с алгебраической топологией.

Законы Кирхгофа, переоткрытые Максвеллом, формулируются очень просто. *Первый закон* вытекает из закона сохранения заряда и состоит в том, что алгебраическая сумма токов, сходящихся в точке разветвления проводников, равна нулю. *Второй закон* утверждает, что существует такая функция  $\Phi$ , называемая электростатическим потенциалом, что напряжение в каждой ветви определяется разностью значений  $\Phi$  на ее концах. Максвелл придумал два метода решения проблемы цепей с сопротивлениями, известные сейчас как метод *круговых токов* и метод *узловых потенциалов*. Мы начнем работать с примерами, иллюстрирующими методы Максвелла, и по ходу дела сформулируем язык алгебраической топологии.

## 12.1. Линейные цепи с резисторами

Резисторы соединяются проводами с пренебрежимо малым сопротивлением. Соединение, изображенное на рис. 12.5, называется *параллельным*. Предположим, что к батарее, создающей постоянное напряжение  $V$ , присоединены несколько резисторов, причем  $i$ -й резистор имеет сопротивление  $R_i$ . Введем величину  $G_i = R_i^{-1}$ , называемую *проводимостью*<sup>3</sup>  $i$ -го резистора. Мы хотим вычислить полный ток, поставляемый батареей, а после того, как процесс установится, — ток в каждой ветви. Соединение верхних входов резисторов обеспечивает на них одинаковый потенциал в стационарном состоянии. Аналогично имеются одинаковые потенциалы на нижних входах резисторов. Следовательно, на каждом резисторе мы имеем одно и то же напряжение, равное напряжению  $V$ , создаваемому батареей. С точки зрения теории цепей мы можем считать, что все верхние и нижние провода стягиваются в точки. Поэтому рис. 12.5 можно заменить на

<sup>3</sup>Единица проводимости, обратная единице сопротивления ом, носит название *сименс*. — Прим. ред.

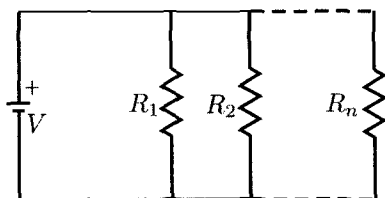


Рис. 12.5

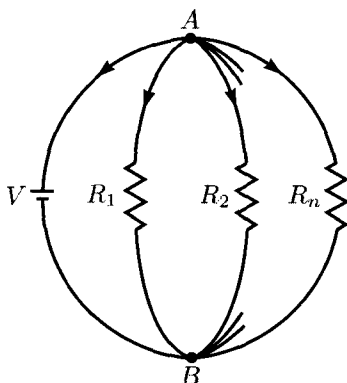


Рис. 12.6

рис. 12.6, где две вершины  $A$  и  $B$  (вверху и внизу) соединяют  $n + 1$  ветвей, на одной из которых находится батарея.

Поскольку ко всем резисторам приложено одно и то же напряжение  $V$ , то из уравнения  $V = RI$  или  $I = GV$  следует, что токи, протекающие через соответствующие резисторы, равны  $G_1V$ ,  $G_2V$ , ...,  $G_nV$ . Согласно первому закону Кирхгофа, ток, протекающий через ветвь с батареей, должен быть равен по величине и противоположен по знаку алгебраической сумме токов в ветвях с резисторами, т. е.

$$-(G_1 + G_2 + \dots + G_n)V.$$

(На рис. 12.6 стрелками обозначены направления токов, причем токи, текущие сверху вниз, считаем положительными, а токи, текущие в обратном направлении, — отрицательными.) Таким образом, мы решили задачу для этой простейшей цепи, т. е. знаем напряжение и ток в каждой ветви. Из полученного результата следует, что полный ток, подаваемый батареей в этом случае, равен току от батареи, если бы в цепи был только один резистор с проводимостью  $G = G_1 + G_2 + \dots + G_n$  и сопротивлением  $R$ , где

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \dots + \frac{1}{R_n}.$$

Итак, если несколько резисторов соединены параллельно, то они эквивалентны одному резистору, обратное сопротивление

которого равно сумме обратных сопротивлений всех отдельных резисторов. Это хорошо известный результат, который преподается в элементарных курсах.

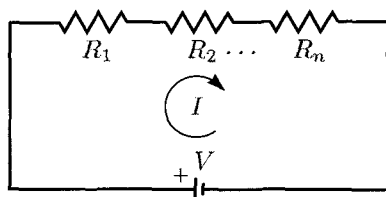


Рис. 12.7

На рис. 12.7 изображена группа резисторов, соединенных проводами *последовательно*. Пусть в цепи имеется батарея, создающая напряжение  $V$ , в результате чего возникает постоянный ток  $I$ . (По первому закону Кирхгофа тот же самый ток  $I$  протекает через каждую ветвь, поскольку в каждом узле алгебраическая сумма токов должна быть равна нулю.) Пусть сопротивления резисторов равны  $R_1, R_2, \dots, R_n$  соответственно. Поскольку через  $i$ -й резистор протекает ток  $I$ , то напряжение на нем равно  $R_i I$ . Из второго закона Кирхгофа следует, что напряжение на батарее должно быть равно сумме напряжений на всех резисторах, т. е.

$$V = (R_1 + R_2 + \dots + R_n)I.$$

Поскольку мы знаем  $V$ , то из этого уравнения получаем ток  $I$ . Таким образом, задача полностью решена — мы знаем ток и напряжение в каждой ветви. Итак, если несколько резисторов соединены последовательно, то они эквивалентны одному резистору, сопротивление которого равно сумме сопротивлений отдельных резисторов. Опять хорошо известный результат.

А теперь давайте рассмотрим более сложную цепь, состоящую из батареи с напряжением  $V$ , соединенной с тремя резисторами, обладающими сопротивлениями  $R_1, R_2$  и  $R_3$ , так, как это показано на рис. 12.8. Точка, из которой выходят два или более проводов, ведущих к различным элементам, называется *узлом* или *вершиной*. Точка  $A$  является узлом, к которому подходят два провода: от батареи и от резистора. К точке  $B$  подходят провода от

трех резисторов. К точке  $C$  подходят два провода от резисторов и один от батареи. Нижний правый угол схемы не считается узлом, потому что он соединяется с точкой  $C$  проводом с нулевым сопротивлением и, следовательно, совпадает с узлом  $C$ . Итак, цепь имеет три узла  $A$ ,  $B$  и  $C$  и четыре ветви: ветвь с батареей, соединяющая узлы  $A$  и  $C$ ; резистор, соединяющий узлы  $A$  и  $B$ ; два резистора, соединяющие узлы  $B$  и  $C$ . Длина и форма проводов, соединяющих все эти точки, не имеют никакого значения. Важны только ветви и узлы, и то, как они соединены друг с другом. Таким образом, на рис. 12.8 и рис. 12.9 изображена одна и та же цепь, в которой есть параллельные и последовательные соединения.

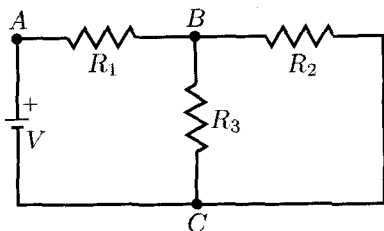


Рис. 12.8

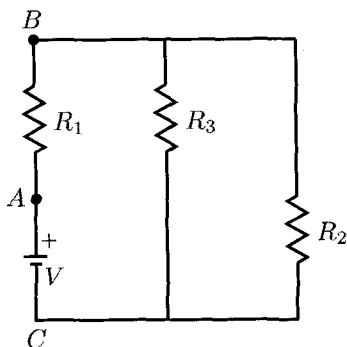


Рис. 12.9

Сопротивления  $R_2$  и  $R_3$  соединены параллельно и, следовательно, с точки зрения всей цепи, эквивалентны одному резистору, обратное сопротивление которого равно  $R_2^{-1} + R_3^{-1}$ , т.е. имеющему сопротивление

$$\frac{R_2 R_3}{R_2 + R_3}.$$

Этот эквивалентный резистор соединен последовательно с сопротивлением  $R_1$ . Это значит, что для батареи есть один резистор с сопротивлением

$$R = R_1 + \frac{R_2 R_3}{R_2 + R_3}.$$

Ток, вытекающий из батареи, равен

$$I = V/R.$$

Такой же ток протекает через резистор  $R_1$ , т. е.  $I_1 = V/R$ . Тогда падение напряжения на резисторе  $R_1$  равняется  $V^1 = I_1 R_1$ . Ток  $I$  делится между резисторами  $R_2$  и  $R_3$  в пропорции, обратной их сопротивлениям, как мы уже видели из обсуждения резисторов, соединенных параллельно. Таким образом, токи, протекающие по резисторам  $R_2$  и  $R_3$ , равны

$$I_2 = \frac{R_3 I}{R_2 + R_3} \quad \text{и} \quad I_3 = \frac{R_2 I}{R_2 + R_3}.$$

Отсюда мы видим, что падение напряжения, одинаковое для резисторов  $R_2$  и  $R_3$ , равно

$$\frac{R_2 R_3}{R_2 + R_3} I.$$

Итак, мы полностью рассчитали цепь, состоящую из резистора  $R_1$ , соединенного последовательно с резисторами  $R_2$  и  $R_3$ , которые соединены параллельно. Такая процедура<sup>4</sup> очень удобна для расчетов простых цепей. Для сложных цепей необходимо разработать более систематический метод. На примере этой простой цепи мы продемонстрируем оба метода Максвелла.

Прежде чем перейти к этому, нарисуем цепь еще раз, изображая ветви (с их ориентацией), узлы и способ их соединения между собой, но не показывая природу отдельных ветвей (рис. 12.10). Итак, у нас есть три узла  $A$ ,  $B$ ,  $C$  и четыре ветви  $\alpha$ ,  $\beta$ ,  $\gamma$  и  $\delta$ . Ветвь  $\alpha$  идет от точки  $A$  к точке  $B$ . Запишем это символически

$$\partial\alpha = B - A.$$

Символ  $\partial$  называется *граничным оператором*. Выписанное соотношение интерпретируется следующим образом: ветвь  $\alpha$  имеет две граничные точки  $B$  и  $A$ ; ветвь выходит из точки  $A$ , поэтому

<sup>4</sup>Вышеописанный анализ параллельных и последовательных соединений называют также методом эквивалентных сопротивлений. — *Прим. ред.*

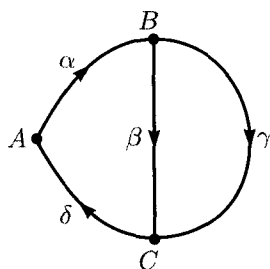


Рис. 12.10

мы пишем ее со знаком минус; ветвь приходит в точку  $B$ , поэтому мы пишем ее со знаком плюс. Граничные соотношения для других ветвей записываются аналогично:

$$\partial\beta = C - B, \quad \partial\gamma = C - B \quad \text{и} \quad \partial\delta = A - C.$$

Под *траекторией* мы будем подразумевать последовательность ветвей, проходимых в их собственном или в обратном направлении так, что конечная точка одного отрезка пути становится началом следующего отрезка. Например, траектория  $\alpha + \beta$  идет из  $A$  в  $B$  и потом из  $B$  в  $C$ . Граничное соотношение для этого пути  $\partial(\alpha + \beta) = B - A + C - B = C - A$ . Траектория  $\beta - \gamma$  идет из  $B$  в  $C$  вдоль  $\beta$ , потом из  $C$  в  $B$  вдоль ветви  $\gamma$ . Граничное соотношение для этого пути  $\partial(\beta - \gamma) = C - B - (B - C) = 0$ . Траектория представляется как последовательность: узел, ветвь, узел, ветвь, ..., ветвь, узел, в которой каждая ветвь ограничена двумя узлами. Траектория последовательно соединяет первую и последнюю точки. Траектория называется *замкнутой*, если первая и последняя точки совпадают. В общем случае мы не предполагаем, что все ветви вдоль траектории разные; мы можем проходить через один и тот же узел или вдоль одной и той же ветви несколько раз. Если все элементы траектории различны (кроме первой и последней точек для замкнутой траектории), то она называется *простой*. Простая замкнутая траектория называется *ячейкой*<sup>5</sup>.

<sup>5</sup>С точки зрения русского читателя естественнее назвать такую траекторию не ячейкой (mesh), а контуром. Термин «ячейка», а чаще «клетка» (cell), более подходит для математической теории комплексов, обсуждаемой в последующих главах. — Прим. ред.

Таким образом,  $M_1 = \alpha + \beta + \delta$  является ячейкой, так же, как и  $M_2 = \gamma - \beta$  (рис. 12.11). В ячейке не важно, что является начальной точкой, т. е. ячейка  $\beta + \delta + \alpha$  совпадает с  $M_1 = \alpha + \beta + \delta$ . Однако направление имеет значение:  $-\gamma + \beta = -M_2$  не то же самое, что  $M_2$ .

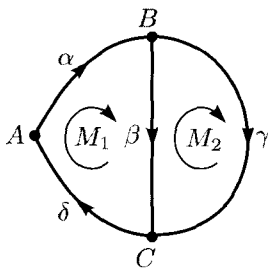


Рис. 12.11

Рассмотрим третью ячейку  $M_3 = \alpha + \gamma + \delta$ . В каком-то смысле эта ячейка не является независимой от других: если мы формально сложим  $M_1$  и  $M_2$ , то при выполнении коммутативного закона для сложения получим

$$M_1 + M_2 = \alpha + \beta + \delta + \gamma - \beta = \alpha + \gamma + \delta = M_3.$$

В следующем параграфе мы подробнее займемся математическим обоснованием таких манипуляций. В рамках метода круговых токов Максвелла выберем набор независимых ячеек (точное определение в следующем параграфе), например,  $M_1$  и  $M_2$ . Введем неизвестные токи  $J_1$  и  $J_2$ , протекающие по этим ячейкам (рис. 12.12). Ток  $J_1$  течет по резисторам  $R_1$ ,  $R_3$  и батарее; ток  $J_2$  течет вниз через резистор  $R_2$  и вверх через резистор  $R_3$ . Если мы знаем значения токов  $J_1$  и  $J_2$ , то сразу же можем вычислить значения токов во всех ветвях. Действительно, поскольку резистор  $R_1$  есть только в ячейке  $M_1$ , то ток через  $M_1$  равен  $J_1$ . Такой же ток протекает по батарее. Поскольку резистор  $R_2$  принадлежит только ячейке  $M_2$ , то ток  $J_2$ , проходящий через эту ячейку, равен току через резистор  $R_2$ . Ветвь, содержащая резистор  $R_3$ , делает положительный вклад в ячейку  $M_1$  и отрицательный вклад в ячейку  $M_2$ . Таким образом, ток, протекающий через резистор  $R_3$ , равен  $J_1 - J_2$ .

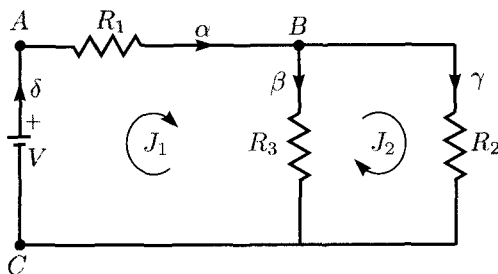


Рис. 12.12

Согласно второму закону Кирхгофа, при обходе любой ячейки изменение потенциала должно быть равно нулю. Например, в случае ячейки  $M_1$  мы имеем падение напряжения  $R_1 J_1$  вдоль ветви  $\alpha$ , падение напряжения  $R_3(J_1 - J_2)$  вдоль ветви  $\beta$  и рост напряжения  $V$  вдоль ветви  $\delta$ , содержащей батарею. Откуда следует

$$R_1 J_1 + R_3(J_1 - J_2) - V = 0.$$

Аналогично для ячейки  $M_2$  получаем

$$R_2 J_2 + R_3(J_2 - J_1) = 0.$$

Итак, у нас есть два уравнения для двух неизвестных  $J_1$  и  $J_2$ . Очевидно, решив эти уравнения, мы получим предыдущий результат.

Таким образом, идея метода круговых токов состоит в выборе независимого набора ячеек, по которым протекают неизвестные токи. Тогда токи в каждой ветви выражаются через токи в ячейках. Затем для каждой ячейки применяется второй закон Кирхгофа, утверждающий, что изменение потенциала при обходе вокруг нее равно нулю. Так получаются уравнения для каждой ячейки, т. е. мы получаем столько уравнений, сколько у нас имеется неизвестных. В каком-то смысле идея метода сводится к следующему: неизвестные токи вводятся таким образом, чтобы автоматически выполнялся первый закон Кирхгофа, а потом используется второй закон. В следующих параграфах мы дадим точное определение «независимых» ячеек и докажем, что этот метод работает, т. е. докажем единственность решения для цепей с резисторами.

А теперь попробуем объяснить суть второго метода Максвелла — метода узловых потенциалов. Поскольку функция электрического потенциала определена с точностью до аддитивной постоянной, выберем потенциал равным нулю в одном из узлов. (Если один из узлов соединен с землей, то удобно именно в нем считать потенциал равным нулю.) Рассмотрим тот же пример, см. рис. 12.13. Пусть потенциал равен нулю в точке  $C$ . Потенциал в точке  $A$ , создаваемый батареей, равен  $V$ . Мы не знаем потенциал только в точке  $B$ . Обозначим этот неизвестный потенциал буквой  $x$ . Применим первый закон Кирхгофа к узлам с неизвестным потенциалом, т. е. в нашем случае к точке  $B$ . Токи, протекающие через узел  $B$ , запишем через сопротивления каждой ветви, входящей в него, и через разность потенциалов соответствующих узлов. Итак, ток, приходящий в точку  $B$  через резистор  $R_1$ , равен  $(V - x)/R_1$ ; ток, приходящий в точку  $B$  через резистор  $R_2$ , равен  $(0 - x)/R_2$ ; ток, приходящий в точку  $B$  через резистор  $R_3$ , равен  $(0 - x)/R_3$ . Согласно первому закону Кирхгофа, алгебраическая сумма токов в каждом узле равна нулю, т. е.

$$(V - x)/R_1 - x/R_2 - x/R_3 = 0.$$

Мы можем решить это уравнение и получить напряжение и токи во всех ветвях. В следующем параграфе мы приведем обоснование этого метода в общем случае.

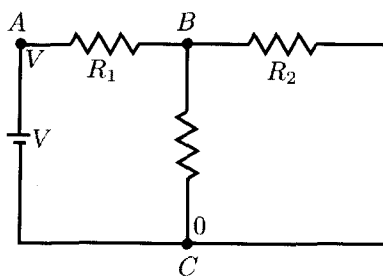


Рис. 12.13

Заметим, что в нашем примере второй метод работает лучше первого, потому что надо найти только одну неизвестную величину, а не две. Но преимущества этого метода по сравнению с анализом последовательных и параллельных соединений не очевидны.

Поэтому, рассчитывая цепь в общем случае, мы советуем сначала проверить, насколько легко ее можно представить в виде параллельных и последовательных соединений. Если это трудно, то, вероятно, не имеет смысла заниматься таким разложением. Тогда посмотрите, сколько получается независимых ячеек и соответственно сколько неизвестных токов. Аналогично, определите, сколько неизвестных узловых потенциалов. Выбирайте тот метод расчета цепи, для которого получается меньше неизвестных. (Число неизвестных часто могут сократить соображения симметрии.) Несмотря на то, что мы рекомендуем делать выбор между двумя методами Максвелла только после тщательного рассмотрения, существует и общее правило: если в цепи мало ячеек и много узлов, то лучше работает первый метод (контурных токов), а если в цепи мало узлов и много ячеек, то лучше воспользоваться вторым методом (узловых потенциалов). Другим важным фактором является тип источников питания. Если ток в цепь подается источником с заданным напряжением, то это уменьшает число узлов с неизвестными потенциалами, что говорит в пользу метода узловых потенциалов. Если же заранее известны токи, то это говорит в пользу метода круговых токов. Чтобы учесть все указанные соображения, обычно хорошо нарисовать две диаграммы цепи, на которых отметить число неизвестных в одном методе и число неизвестных в другом. И только после этого сделать разумный выбор между ними.

До сих пор мы рассматривали цепи, содержащие только резисторы. Оба метода Максвелла применяются также для изучения стационарного (осциллирующего) поведения линейных цепей с индукторами и конденсаторами. (Линейность означает, что все индуктивности и емкости предполагаются постоянными.) Если все генераторы производят ток с одинаковой частотой  $\omega/2\pi$ , т. е. все источники напряжения имеют вид  $Ve^{i\omega t}$ , а все источники тока —  $Ie^{i\omega t}$ , то, как хорошо известно (тривиальное следствие определений)<sup>6</sup>, для индуктора с индуктивностью  $L$  имеем формулу  $V = i\omega LI$ , а для конденсатора с емкостью  $C$  — формулу

---

<sup>6</sup>Предлагается рассматривать вывод нижеследующих формул для импедансов индуктора и конденсатора из определений на с. 577 как несложное упражнение. — *Прим. ред.*

$V = (1/iC\omega)I$ . Теперь можно воспользоваться теми же правилами, что и раньше, но с *комплексными сопротивлениями* или *импедансами*. Однако в этом случае решения существуют не всегда из-за явления резонанса. Например, пусть индуктор с  $L = 1$  соединен с конденсатором, имеющим  $C = 1$ . Если мы к ним присоединим генератор тока с частотой  $\omega = 1$ , то правило сложения сопротивлений дает  $R = i + 1/i = 0$ , т. е. короткое замыкание с бесконечным током. В следующем параграфе мы обсудим некоторые условия, позволяющие избежать такой нереальной ситуации.

## 12.2. Топология одномерных комплексов

Два термина — *ориентированный граф* и *одномерный комплекс* — являются синонимами. Они оба относятся к математической структуре, выражающей соединение различных ветвей в электрическую цепь при условии, что мы не интересуемся природой этих ветвей. Такая структура дает нам возможность изучать законы Кирхгофа в абстрактной форме и видеть, как свойства отдельных ветвей влияют на решение всей задачи.

*Одномерным комплексом* мы называем структуру, состоящую из двух множеств: из множества объектов нулевой размерности, или узлов, обозначаемого  $\{A, B, \dots\} = S_0$ , и множества одномерных объектов, или ветвей, обозначаемых  $\{\alpha, \beta, \dots\} = S_1$ . При этом существует правило, согласно которому каждой ветви соответствуют два различных узла: начальная и конечная точка ветви. Итак, нам дано отображение из  $S_1$  в  $S_0 \times S_0$ .<sup>7</sup> В дальнейшем будем считать, что множества  $S_0$  и  $S_1$  конечны.

Определим термины *траектория*, *ячейка* и т. д., введенные в предыдущем параграфе. Траектория определяет целое число для каждой ветви, т. е. сколько раз проходит данная ветвь для данной траектории с учетом ориентации. Если ветвь проходит от начальной точки к конечной, то ее вклад равен  $+1$ , а если в обратном направлении, то  $-1$ . Таким образом, каждая траектория определяет вектор  $\mathbf{p} = (p_\alpha, p_\beta, \dots)^T$  с целыми координатами, имеющими индексы, соответствующие ветвям. Например,  $p_\alpha$  —

---

<sup>7</sup>  $S_0 \times S_0$  обозначает декартово произведение множества  $S_0$  на себя, т. е.  $S_0 \times S_0$  является набором упорядоченных пар узлов.

это целое число, которое обозначает, сколько раз ветвь  $\alpha$  проходит в положительном направлении, минус сколько раз она проходит в отрицательном направлении. Можно сказать, что распределение тока в цепи определяется вектором  $\mathbf{I} = (I_\alpha, I_\beta, \dots)^T$ , координаты которого имеют индексы, соответствующие ветвям. Например  $I_\alpha$  — это вещественное число, определяющее ток (в амперах), проходящий по ветви  $\alpha$  (если наш одномерный комплекс представляет собой ветви электрической цепи). Это значит, что для любого комплекса можно ввести векторное пространство, состоящее из векторов  $\mathbf{K} = (K_\alpha, K_\beta, \dots)^T$ , компоненты которого имеют индексы, соответствующие ветвям. (Если это специально не оговаривается, то полагаем, что все компоненты являются вещественными числами, так что мы имеем вещественное векторное пространство.) Обозначим это пространство  $C_1$  и назовем его *1-цепью*. Каждую ветвь  $k$  задаем вектором, у которого  $k$ -я компонента равна 1, а остальные равны нулю, т. е.  $\alpha = (1, 0, 0, \dots)^T$ ,  $\beta = (0, 1, 0, \dots)^T$  и т. д.<sup>8</sup>

Аналогично построим вещественное векторное пространство  $C_0$ , состоящее из векторов, компоненты которого имеют индексы, соответствующие узлам, и назовем его *0-цепью*. Узлу  $A$  соответствует вектор,  $A$ -я компонента которого равна 1, а остальные равны нулю, так что теперь узел становится элементом векторного пространства  $C_0$ . В частности,  $A - B$  тоже является элементом векторного пространства  $C_0$ . Заметим, что размерность пространства  $C_1$  равна числу ветвей, а размерность пространства  $C_0$  — числу узлов.

Определим теперь линейное отображение  $\partial$  из  $C_1$  в  $C_0$ , называемое *граничным отображением*. Чтобы это сделать, достаточно задать его значения для каждой ветви, поскольку ветви образуют базис в пространстве  $C_1$ . Каждая ветвь имеет начальную и конечную точки, поэтому по определению  $\partial(\text{ветви}) = (\text{конечный узел}) - (\text{начальный узел})$ . Например если  $\alpha$  — ветвь, идущая из точки  $A$  в точку  $B$ , то  $\partial\alpha = B - A$ .

<sup>8</sup>В этом абзаце и в дальнейшем вектор в пространстве  $\mathbb{R}^n$  обозначается

$(a, b, \dots)^T$  вместо  $\begin{pmatrix} a \\ b \\ \vdots \end{pmatrix}$ , чтобы занимать меньше места.

Давайте разберемся, что означает оператор  $\partial$ . Предположим, что  $\mathbf{K} = (K_\alpha, K_\beta, K_\gamma, \dots)^T$  и  $\partial\mathbf{K} = \mathbf{L}$ , где  $\mathbf{L} = (L_\alpha, L_\beta, L_\gamma, \dots)^T$ . Вычисляя, например,  $L_A$ , мы получим сумму определенных коэффициентов  $\mathbf{K}$ , т. е.

$$L_A = (K_{\delta_1} + \dots + K_{\delta_l}) - (K_{\varepsilon_1} + \dots + K_{\varepsilon_k}),$$

где  $\delta_1, \dots, \delta_l$  — ветви, входящие в узел  $A$ , а  $\varepsilon_1, \dots, \varepsilon_k$  — ветви, выходящие из него. Таким образом, для узла, изображенного на рис. 12.14, мы имеем

$$L_A = K_\rho - K_\eta - K_\alpha.$$

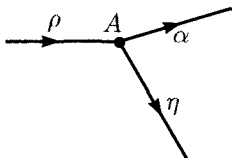


Рис. 12.14

Отсюда видно, что первый закон Кирхгофа имеет очень простую формулировку.

*Первый закон Кирхгофа:* Если  $\mathbf{I}$  является 1-цепью, задающей распределение электрического тока, то

$$\partial\mathbf{I} = 0.$$

Напомним, что простая замкнутая траектория называется *ячейкой*. На рис. 12.15, например, траектория  $\alpha + \beta + \delta$  является ячейкой, которой соответствует вектор  $\mathbf{p}_1 = (1, 1, 0, 1)^T$ ; траектория  $-\beta + \gamma$  — тоже ячейка, и ей соответствует вектор  $\mathbf{p}_2 = (0, -1, 1, 0)^T$ . Сумма этих двух векторов  $\mathbf{p}_3 = \mathbf{p}_1 + \mathbf{p}_2 = (1, 0, 1, 1)^T$  соответствует ячейке  $\alpha + \gamma + \delta$ . Очевидно, что каждая из этих ячеек не имеет границ, т. е.  $\partial\mathbf{p}_1 = \partial\mathbf{p}_2 = \partial\mathbf{p}_3 = 0$ .

Для любого одномерного комплекса подпространство  $C_1$ , состоящее из 1-цепей  $\mathbf{K}$ , удовлетворяющих условию  $\partial\mathbf{K} = 0$ , будем

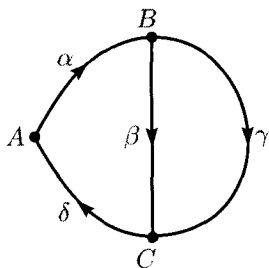


Рис. 12.15

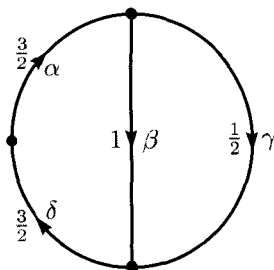


Рис. 12.16

обозначать  $Z_1$ . Запишем это символически:  $Z_1 = \ker \partial \subset C_1$ . Элементы подпространства  $Z_1$  назовем *циклами*. Каждая ячейка является циклом, но не каждый цикл является ячейкой. Например, вектор

$$\mathbf{I}_1 = \begin{pmatrix} 3/2 \\ 1 \\ 1/2 \\ 3/2 \end{pmatrix}$$

удовлетворяет условию  $\partial \mathbf{I}_1 = 0$ , но он не описывает ячейку, потому что его элементы не все равны 1,  $-1$  или 0. Этот вектор представляет на самом деле набор токов, удовлетворяющий первому закону Кирхгофа. По существу, легче всего убедиться, что  $\mathbf{I}_1$  является циклом, написав у каждой ветви соответствующий ток и проверив, что алгебраическая сумма токов в каждом узле равна нулю (см. рис. 12.16).

Или, по-другому, запишем вектор в виде  $\mathbf{I}_1 = \frac{3}{2}\mathbf{p}_1 + \frac{1}{2}\mathbf{p}_2$ . Поскольку  $\mathbf{p}_1$  и  $\mathbf{p}_2$  являются элементами пространства  $Z_1$ , то и  $\mathbf{I}_1$  принадлежит этому же пространству. В этой простой цепи две ячейки  $\mathbf{p}_1$  и  $\mathbf{p}_2$  образуют базис пространства  $Z_1$ .

Определив подпространство  $Z_1 \subset C_1$  как *ядро* граничного отображения  $\partial$ , рассмотрим теперь *образ*  $\partial$ . Пусть  $B_0$  — подпространство  $C_0$ , являющееся образом  $\partial$ . Символически это можно записать как

$$B_0 = \partial C_1 \subset C_0.$$

Подпространство  $B_0$  будем называть *пространством грани*.

Смысл пространства  $B_0$  может стать яснее, если посмотреть на рис. 12.17. Граница ветви  $\alpha$ , равная  $A - B = (1, -1, 0)^T$ , принадлежит пространству  $C_0$  и лежит в подпространстве  $B_0$ . Аналогично граница  $\beta$  равна  $B - C = (0, 1, -1)^T$ . Сумма этих двух векторов  $A - C = (1, 0, -1)^T$  тоже является элементом пространства  $C_0$  и границей траектории  $\alpha + \beta$ .

Однако, не каждый элемент подпространства  $B_0$  реализуется как граница траектории. Например,  $(2, -1, -1)^T$  является элементом  $B_0$ , но не соответствует никакой траектории.

Если мы рассмотрим те элементы пространства  $C_0$ , которые не лежат в подпространстве  $B_0$ , то увидим, что они не образуют подпространства. Рассмотрим цепь, изображенную на рис. 12.17. Вектора  $A = (1, 0, 0)^T$  и  $B = (0, 1, 0)^T$  не являются элементами пространства  $B_0$ , но их разность  $A - B = (1, -1, 0)^T$  лежит в этом подпространстве. Мы можем образовать факторпространство  $H_0 = C_0/B_0$ , элементы которого суть классы эквивалентности элементов  $C_0$ , разность которых лежит в  $B_0$ . Векторы  $A = (1, 0, 0)^T$  и  $B = (0, 1, 0)^T$  в пространстве  $C_0$  соответствуют тому же вектору в  $H_0$ , потому что их разность равна вектору  $(1, -1, 0)^T$ , лежащему в  $B_0$ . Обозначим этот класс эквивалентности  $\bar{A}$  или  $\overline{(1, 0, 0)}^T$ . Аналогично векторы  $(2, 0, 0)^T$  и  $(0, 1, 1)^T$  принадлежат одному и тому же классу эквивалентности  $\overline{(2, 0, 0)}^T$ , потому что их разность  $(2, -1, -1)^T$  лежит в  $B_0$ . Для цепи, изображенной на рис. 12.17, каждый вектор из пространства  $C_0$  лежит в том же классе эквивалентности, что и вектор типа  $(\lambda, 0, 0)^T$ , где  $\lambda$  — вещественное число. Поэтому факторпространство  $H_0 = C_0/B_0$  является одномерным векторным пространством, изоморфным пространству вещественных чисел.

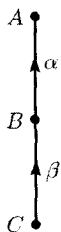


Рис. 12.17

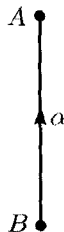


Рис. 12.18



На рис. 12.18 изображен пример комплекса, для которого факторпространство  $H_0$  является двумерным. Векторы  $(1, 0, 0, 0)^T$  и  $(0, 0, 1, 0)^T$  не принадлежат одному классу эквивалентности, потому что их разность  $(1, 0, -1, 0)^T$  не является границей никакого элемента  $C_1$ . В этом случае класс эквивалентности вектора  $(1, 0, 0, 0)^T$  и класс эквивалентности вектора  $(0, 0, 1, 0)^T$  образуют базис двумерного векторного пространства  $H_0$ . Заметим, что если бы к комплексу добавить ветвь  $\gamma$ , соединяющую  $A$  и  $C$ , то  $(1, 0, -1, 0)^T$  стал бы элементом пространства  $B_0$ , а факторпространство  $H_0$  стало бы одномерным.

### Смысл $H_0$ и $Z_1$

Первая задача, стоящая перед нами — дать геометрическую интерпретацию пространств  $H_0$  и  $Z_1$ , что поможет лучше понять метод круговых токов. Мы хотим доказать следующие два факта (точные формулировки дадим в процессе обсуждения):

- (i)  $\dim H_0$  равна числу связных компонент комплекса;
- (ii) можно найти базис  $Z_1$ , состоящий из ячеек.

Начнем с (i). Что такое связные компоненты? Мы говорим, что траектория соединяет узел  $A$  с узлом  $D$ , если  $A$  является начальной точкой первого отрезка пути, а  $D$  — конечная точка последнего отрезка. Если  $\mathbf{P}$  — 1-цепь с целыми коэффициентами, соответствующая этой траектории, то очевидно, что  $\partial \mathbf{P} = D - A$ , потому что значение  $\partial \mathbf{P}$  можно вычислить, складывая границы отдельных ветвей. При этом исчезают все промежуточные узлы. Комплекс называется *связным*, если все его точки могут быть соединены друг с другом. В этом случае произвольную 0-цепь<sup>9</sup>

$$\mathbf{L} = \sum_{N=A,B,C,\dots} L_N N = (L_A, L_B, \dots)$$

можно записать в виде

$$\mathbf{L} = \sum L_N \{(A + (N - A))\} = \sum L_N A + \sum_{N \neq A} L_N (N - A),$$

<sup>9</sup>Напомним, что под  $N$  здесь в выражении  $L_N N$  понимается вектор — набор чисел, занумерованный узлами, у которого  $N$ -я компонента равна 1, а остальные 0. — Прим. ред.

или, что то же самое,

$$\mathbf{L} = (L_A + L_B + \dots, 0, 0, \dots)^T + L_B(-1, 1, 0, 0, \dots)^T + L_C(-1, 0, 1, 0, \dots)^T + \dots$$

Для связного комплекса  $N - A$  является границей для всех  $N$  ( $\neq A$ ) и

$$\mathbf{L} - \left( \sum_N L_N \right) A = \sum_{N \neq A} L_N (N - A)$$

является элементом  $B_0$ . Таким образом *каждая* 0-цепь  $\mathbf{L}$  содержится в классе эквивалентности некоторого кратного  $A$ , а каждый элемент факторпространства  $H_0 = C_0/B_0$  является кратным  $\bar{A}$  — класса эквивалентности  $A$ . С другой стороны, само  $A$  не является границей, поэтому  $\bar{A} \neq 0$ . Итак, мы доказали, что для любого связанного комплекса размерность  $H_0$  равна 1.

Пусть  $A$  — узел произвольного комплекса. Рассмотрим все ветви, для которых  $A$  является граничной точкой, начальной или конечной. Точку  $A$  свяжем со всеми узлами, находящимися на другом конце ветвей. Теперь добавим все остальные ветви, исходящие из этих узлов, и все новые точки на концах этих ветвей. Продолжаем эту процедуру до тех пор, пока не останется ни узлов, ни ветвей, которые можно было бы еще присоединить. В результате получим набор узлов  $S_0(A)$  и набор ветвей  $S_1(A)$ . Причем граничные точки любой ветви из  $S_1(A)$  принадлежат набору  $S_0(A)$ . Таким образом,  $(S_0(A), S_1(A))$  образует *подкомплекс* начального комплекса. Если он не совпадает со всем комплексом, то можно выбрать узел  $B$ , не лежащий в  $S_0(A)$ , и повторить всю процедуру. Действуя таким образом, мы получим семейства непересекающихся множеств: узлов  $S_0(A)$ ,  $S_0(B)$  и т. д. и ветвей  $S_1(A)$ ,  $S_1(B)$  и т. д. При этом связные подкомплексы  $S_0(A)$ ,  $S_1(A)$ ,  $S_0(B)$ ,  $S_1(B)$  и т. д. не соединены друг с другом. Мы получим соответствующие векторные пространства  $C_0(A)$  и т. д. и разложения в прямую сумму<sup>10</sup>  $C_0 = C_0(A) \oplus C_0(B) \oplus \dots$  и  $C_1 =$

<sup>10</sup>Прямая сумма  $V_1 \oplus V_2$  двух векторных пространств  $V_1$  и  $V_2$  есть декартово произведение  $V_1 \times V_2$ , превращенное в векторное пространство по естественным правилам:  $(v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2)$  и  $r(v_1, v_2) = (rv_1, rv_2)$ .

$C_1(A) \oplus C_1(B) \oplus \dots$ , причем  $\partial C_1(A) \subset C_0(A)$ ,  $\partial C_1(B) \subset C_0(B)$  и т. д. Отсюда мы видим, что  $H_0 = H_0(A) \oplus H_0(B) + \dots$ , поэтому размерность  $H_0$  равняется числу слагаемых и числу связных компонент.

Теперь рассмотрим (ii). Ячейка — это простой замкнутый путь. Следовательно, каждая ячейка определяет элемент  $\mathbf{M}$  в  $C_1$  с координатами, равными либо  $+1$ , либо  $-1$ , либо  $0$ , а это значит, что  $\partial \mathbf{M} = 0$ . Мы говорим, что множество ячеек независимо, если линейно независимы соответствующие элементы  $C_1$ . (Это и есть строгая формулировка понятия независимости, которое мы использовали в предыдущем параграфе.) А сейчас мы хотим показать, что можно найти такое семейство ячеек, что соответствующие элементы  $\mathbf{M}_1, \dots, \mathbf{M}_z$  образуют базис в пространстве  $Z_1$ . В процессе доказательства будет показана процедура нахождения такого семейства ячеек. Заметим, что в силу сказанного при обсуждении пункта (i) достаточно работать со связным комплексом. Если комплекс не связан, то процедура применяется для каждой связной компоненты отдельно. Поскольку компоненты не зависят друг от друга, то мы получим результат для всего комплекса.

Связный комплекс, не содержащий ячеек, называется *деревом*. В дереве имеется по крайней мере один узел, который является граничной точкой только для одной ветви. (Мы исключаем тривиальный случай, когда в комплексе есть только один узел и нет ветвей.) Чтобы доказать этот факт, начнем с любого узла. Если к узлу подходит несколько ветвей, то выберем одну и двинемся вдоль нее к противоположному концу. Если этот новый узел не обладает требуемым свойством, то двинемся вдоль другой ветви (не той, вдоль которой пришли) к следующему узлу. Продолжим эту процедуру. Поскольку в комплексе нет ячеек, мы никогда не

Если  $U$  и  $W$  являются подпространствами векторного пространства  $V$  такими, что каждый элемент  $v \in V$  может быть записан единственным образом в виде  $v = u + w$ , где  $u \in U$ ,  $v \in W$ , то  $v \mapsto (u, w)$  дает изоморфизм  $V$  с прямой суммой  $U$  и  $W$ . В этом случае мы пишем  $V = U \oplus W$ , причем  $\dim V = \dim U + \dim W$ . [Обычно говорят, что  $V$  является *внутренней* прямой суммой подпространств  $U$  и  $W$ , в отличие от описанной вначале *внешней* суммы независимых векторных пространств, а также используют для внутренней прямой суммы другое обозначение операции, например,  $V = U \dot{+} W$ . — *Ред.*]

вернемся в исходный узел. Так как в комплексе конечное число узлов, то в конце концов мы придем в узел, который будет границей только одной ветви.

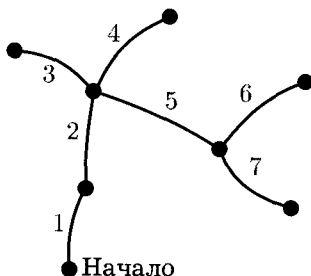


Рис. 12.19

Предположим, что у нас есть дерево и мы начали с узла, который является концом только одной ветви. Начиная с этого узла, мы построим все дерево, добавляя по одной ветви (рис. 12.19). Поскольку нет ячеек, то каждый раз, добавляя ветвь, мы добавляем узел. Таким образом, у дерева число узлов на единицу больше числа ветвей: если  $b$  обозначает число ветвей, а  $n$  — число узлов, то

$$n = b + 1.$$

Мы уже доказали, что  $\dim H_0 = 1$  для любого связного комплекса. Ввиду  $H_0 = C_0/B_0$  имеем  $\dim H_0 = \dim C_0 - \dim B_0$ , откуда следует, что  $\dim B_0 = \dim C_0 - \dim H_0 = n - 1 = b$ . Поэтому  $B_0$ , являясь образом отображения  $\partial$ , имеет ту же самую размерность, что и пространство  $C_1$ , где действует  $\partial$ . Мы знаем, что для любого линейного отображения  $T : V \rightarrow W$  выполняется условие  $\dim(\text{Im } T) + \dim(\text{Ker } T) = \dim V$ . Применяя его к отображению  $\partial$ , мы получаем, что  $\dim(\text{Ker } \partial) = \dim C_1 - \dim B_0 = b - b = 0$ , следовательно,  $\text{Ker } \partial = \{0\}$ , т.е.  $Z_1 = \{0\}$ . Таким образом, мы доказали (ii) для случая дерева: если нет ячеек, то нет нетривиальных циклов.

Предположим, что у нас есть связный комплекс, построенный, как только что было показано, начиная с одного произвольного узла. Каждый раз добавляя ветвь, мы можем добавить или не добавить новый узел. В конце концов, поскольку комплекс связан,

когда мы включим все ветви, мы добавим и все узлы. Итак, в общем случае для связного комплекса мы имеем

$$n \leq b + 1.$$

Поскольку для связного комплекса  $\dim H_0 = 1$ , а  $\dim C_0 - \dim B_0 = \dim H_0$ , находим  $n - \dim B_0 = 1$  и  $\dim B_0 = n - 1$ . Поскольку  $B_0$  является образом  $\partial$  и  $Z_1$  является ядром  $\partial$ , то  $\dim B_0 + \dim Z_1 = \dim C_1 = b$ , или  $(n - 1) + \dim Z_1 = b$ . Поэтому для любого связного комплекса

$$\dim Z_1 = b + 1 - n.$$

Чтобы доказать (ii), мы должны найти  $b + 1 - n$  независимых ячеек, которые образуют базис в пространстве  $Z_1$ . Для этого сначала разделим множество всех ветвей  $S_1$  на *максимальное дерево*  $T$ , которое содержит все узлы нашего комплекса<sup>11</sup>, и множество ветвей  $\bar{T}$ , которые не принадлежат дереву. Делать будем следующим образом: выберем любую ветвь, являющуюся частью ячейки, пусть она будет элементом подмножества  $\bar{T}$ . Все оставшиеся ветви по-прежнему соединяют все узлы. Повторим эту процедуру, пока не останется ни одной ячейки. Тогда ветви, *не* помещенные в подмножество  $\bar{T}$ , образуют максимальное дерево  $T$ . Такая процедура не обеспечивает единственность максимального дерева<sup>12</sup>. На рис. 12.20 изображены два различных способа, которыми данный комплекс может быть сведен к максимальному дереву (сплошные линии) путем удаления ветвей (пунктирные линии), являющихся частями ячеек. Поскольку  $T$  образует дерево, связывающее  $n$  узлов, то оно содержит  $n - 1$  ветвь. Подмножество  $\bar{T}$  содержит  $b - (n - 1) = b + 1 - n$  ветвей. Концы каждой ветви из  $\bar{T}$  соединены ветвями из множества  $T$ , поскольку  $T$  соединяет все

<sup>11</sup>Дерево максимально в том смысле, что при добавлении к нему любой новой ветви первоначального комплекса подкомплекс перестает быть деревом. Кстати, как легко видеть, дерево будет максимальным как раз в том и только том случае, если оно соединяет все узлы данного комплекса, что дает альтернативный способ построения такого дерева: не выбрасывать по ветви (как описано ниже), а добавлять по узлу. — *Прим. ред.*

<sup>12</sup>Максимальное дерево единственно только при отсутствии в комплексе ячеек, за исключением ветвей с совпадающими началом и концом. — *Прим. ред.*

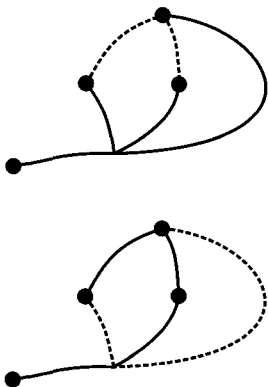


Рис. 12.20. Два различных максимальных дерева для одного комплекса

узлы. Более того, поскольку  $T$  не содержит ячеек, то связывающая траектория единственна. Каждая ветвь из  $\bar{T}$  в комбинации с этой единственной траекторией в  $T$ , соединяющей ее концы, образует ячейку. Поскольку число ветвей в  $\bar{T}$  равно  $b + 1 - n$ , что совпадает с размерностью  $Z_1$ , нам остается только доказать, что построенные нами ячейки линейно независимы.

Для доказательства линейной независимости  $i$ -ю ветвь в  $\bar{T}$  обозначим  $\alpha_i$ . Пусть  $\mathbf{M}_i$  обозначает ячейку, содержащую ветвь  $\alpha_i$ . Рассмотрим линейную комбинацию ячеек  $\sum c_i \mathbf{M}_i$ . Поскольку ветвь  $\alpha_i$  присутствует *только* в ячейке  $\mathbf{M}_i$  с коэффициентом  $+1$ , то коэффициент при  $\alpha_i$  в сумме должен быть  $c_i$ . Поэтому  $\sum c_i \mathbf{M}_i = 0$  выполняется только если *все*  $c_i = 0$ . Отсюда следует, что ячейки  $\mathbf{M}_i$  линейно независимы.

## Деревья и проекции

Заметим, что выбор максимального дерева  $T$  определяет проекцию  $\rho_T$  пространства  $C_1$  на подпространство  $Z_1$  по правилу:

$$\begin{aligned} \text{если } \alpha_i \in T, & \quad \text{то } \rho_T(\alpha_i) = 0; \\ \text{если } \alpha_i \in \bar{T}, & \quad \text{то } \rho_T(\alpha_i) = \mathbf{M}_i. \end{aligned}$$

Например, для максимального дерева, состоящего из  $\alpha$  и  $\beta$  (см. рис. 12.21), проекционный оператор  $\rho_T$  есть  $\rho_T(\alpha) = 0$ ,  $\rho_T(\beta) = 0$ ,

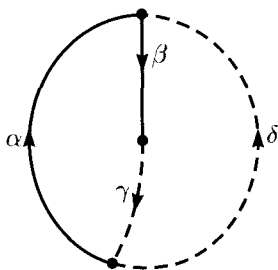


Рис. 12.21

$\rho_T(\gamma) = \alpha + \beta + \gamma$ ,  $\rho_T(\delta) = -\alpha + \delta$ . (Заметим, что каждая ячейка  $\mathbf{M}_i$  включает в себя ветвь  $\alpha_i \in \bar{T}$  с коэффициентом  $+1$ , а не  $-1$ .) Проекция  $\rho_T$  может быть записана в виде матрицы

$$\begin{pmatrix} 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Если мы имеем дело с несвязным комплексом, то (ii) доказывается выбором максимального дерева для каждой связной компоненты и построением проекции  $\rho_T$ .

## Метод круговых токов

Теперь становится понятной важность метода круговых токов, разработанного Максвеллом. Выбирая семейство независимых ячеек и приписывая каждой ячейке *круговой ток*  $J_i$ , мы образуем элемент  $J_i \mathbf{M}_i + \dots + J_m \mathbf{M}_m$  пространства  $Z_1$ . Таким образом можно получить наиболее общее выражение для токов, удовлетворяющих первому закону Кирхгофа, т. е. наиболее общий вид элемента пространства  $Z_1$ .

Ранее мы уже пользовались уравнением

$$1 - \dim Z_1 = n - b \quad (12.1)$$

для связных комплексов. Применяя его к каждой компоненте произвольного комплекса и складывая, получаем

$$\dim H_0 - \dim Z_1 = \dim C_0 - \dim C_1. \quad (12.2)$$

Поскольку  $\dim H_0 = \dim C_0 - \dim B_0$  и  $\dim B_0 = \dim C_1 - \dim Z_1$ , уравнение (12.2) есть прямое следствие определений различных пространств, связанных с  $\partial$ .

В рамках этого метода Максвелла мы считаем, что круговой ток определяет токи в каждой ветви. Это можно сформулировать математически строго: рассмотрим пространство круговых токов как еще одну копию пространства  $Z_1$ ; назовем его  $H_1$ . Это пространство круговых токов, будучи копией  $Z_1$ , рассматривается как абстрактное векторное пространство, не являющееся подпространством  $C_1$ . А теперь рассмотрим процедуру нахождения токов в ветвях, определяемых круговыми токами, как линейное отображение  $\sigma : H_1 \rightarrow C_1$ , которое идентифицирует  $H_1$  с подпространством  $Z_1$  пространства  $C_1$ . Например, в цепи, изображенной на рис. 12.21, пространство  $H_1$  двумерно. Отображение  $\sigma$  приписывает ячейке  $\mathbf{M}_1$  токи в ветвях  $\alpha + \beta + \gamma$ , так что можно написать

$$\sigma(\mathbf{M}_1) = \alpha + \beta + \gamma$$

и аналогично

$$\sigma(\mathbf{M}_2) = \gamma - \beta.$$

Итак, если мы используем  $\mathbf{M}_1$  и  $\mathbf{M}_2$  в качестве базиса абстрактного пространства  $H_1$  (пространства токов), а  $\alpha, \beta, \gamma, \delta$  — в качестве базиса пространства  $C_1$ , то матрица  $\sigma$ , связывающая эти базисы, имеет вид

$$\begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Таким образом, мы видим, что метод круговых токов сводит задачу обращения линейного отображения из пространства  $H_1$  в его дуальное пространство, обозначаемое  $H^1$ .

Возвращаясь к уравнению (12.2), можно теперь написать

$$\boxed{\dim H_0 - \dim H_1 = \dim C_0 - \dim C_1.} \quad (12.3)$$

Второй закон Кирхгофа утверждает, что напряжение в каждой ветви может быть получено как разность потенциалов от

потенциальной функции, определенной на узлах. Поскольку имеет смысл только разность потенциалов, мы можем положить потенциал на одном узле каждой связной компоненты равным нулю. Тогда потенциалы всех остальных узлов определяют напряжение на ветвях. Например, рассмотрим цепь, изображенную на рис. 12.22. Если положить потенциалы  $\Phi^A = \Phi^D = 0$ , то напряжения ветвей будут равны:  $V^\alpha = \Phi^C$ ,  $V^\beta = \Phi^B$ ,  $V^\delta = \Phi^C - \Phi^B$ ,  $V^\gamma = \Phi^E$ . Поскольку число узлов равно  $\dim C_0$ , а число связных компонент равно  $\dim H_0$ , существует только  $\dim C_0 - \dim H_0$  независимых способов определения напряжений в ветвях, совместимых со вторым законом Кирхгофа. Это значит, что, грубо говоря, число условий, налагаемых вторым законом Кирхгофа, равно

$$\dim C_1 - (\dim C_0 - \dim H_0) = \dim H_1.$$

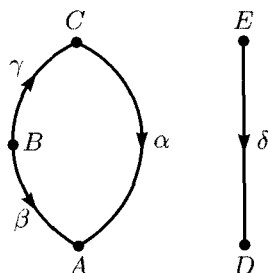


Рис. 12.22

Поскольку число независимых распределений токов в ветвях, совместимых с первым законом Кирхгофа, равно  $\dim H_1$  (числу линейно независимых ячеек), то первый закон Кирхгофа налагает  $\dim C_1 - \dim H_1$  условий. Оба закона вместе налагают  $\dim C_1 = b$  условий. (Законы независимы, потому что один относится к токам, а второй к напряжениям.) Связь между током и напряжением в каждой ветви дает еще  $b$  уравнений. Таким образом мы имеем  $2b$  уравнений для  $2b$  неизвестных токов и напряжений. Теорема Кирхгофа для линейной цепи с резисторами утверждает, что эти уравнения могут быть разрешены. Чтобы доказать эту теорему, мы должны обратиться ко второму закону Кирхгофа.

### 12.3. Коцепи и $d$ -оператор

Прежде всего отметим, что напряжение в каком-то смысле двойственно току. Произведение напряжения  $V_\gamma$  на ток  $I_\gamma$ , который течет по ветви  $\gamma$ , дает мощность, выделенную в этой ветви. Поэтому если мы введем *вектор напряжений*  $\mathbf{V} = (V^\alpha, V^\beta, \dots)$ , то он должен лежать в пространстве, дуальном пространству  $C_1$ . Тогда введем это дуальное пространство и назовем его пространством *1-коцепей*. Обозначим его  $C^1$ . Таким образом,  $C^1$  — это пространство линейных функций в  $C_1$ . Аналогично введем пространство  $C^0$  линейных функций на  $C_0$ , и назовем его пространством *0-коцепей*. А теперь давайте введем несколько обозначений, которые на первый взгляд могут показаться странными, но потом окажутся весьма продуктивными в смысле обобщения. Пусть даны вектора  $\mathbf{K} = (K_\alpha, K_\beta, \dots)^T \in C_1$  и  $\mathbf{W} = (W^\alpha, W^\beta, \dots) \in C^1$ . Значение линейной функции  $\mathbf{W}$  на векторе  $\mathbf{K}$  будем обозначать  $\int_{\mathbf{K}} \mathbf{W}$ . Тогда

$$\int_{\mathbf{K}} \mathbf{W} = W^\alpha K_\alpha + W^\beta K_\beta + \dots$$

Аналогично значение 0-коцепи  $f = (f^A, f^B, \dots)^T \in C^0$  на 0-цепи  $\mathbf{c} = (c_A, c_B, \dots)^T$  будем обозначать  $\int_{\mathbf{c}} f$ . Тогда

$$\int_{\mathbf{c}} f = f^A c_A + f^B c_B + \dots$$

Следующее обозначение связано с отображением  $d$ . Граничное отображение  $d$  является линейным отображением из  $C_1$  в  $C_0$ . Его сопряженное будет линейным отображением из  $C^0$  в  $C^1$ . Это сопряженное отображение будем обозначать  $d$  и назовем его *кограничным оператором*. Итак, если  $f$  есть 0-коцепь и  $\mathbf{K}$  есть 1-цепь, то значение  $df$  на  $\mathbf{K}$  равно  $f$ , вычисленному на  $d\mathbf{K}$ . В терминах уже введенных обозначений, мы можем написать

$$\int_{\mathbf{K}} df = \int_{d\mathbf{K}} f. \quad (12.4)$$

(Наши обозначения выбраны таким образом, чтобы это уравнение было похоже на фундаментальную теорему дифференциального исчисления. Конечно, в уравнении нет никаких интегралов; это вычисление линейных функций от векторов в векторном пространстве, т.е. конечных сумм. Однако, если представить, что  $f$  — дифференцируемая функция на прямой,  $df$  — как  $f'(x)dx$ , а  $\mathbf{K}$  — как интервал  $A \leq x \leq B$ , то левая часть уравнения (12.4) может быть интерпретирована как интеграл  $\int_K df = \int_A^B f'(x) dx = f(B) - f(A)$ , а правую часть этого уравнения именно так и надо понимать. А теперь мы просим читателя на время забыть эту аналогию. Мы вернемся к ней в параграфе 15.2.)

Если  $\alpha$  обозначает ветвь и  $\partial\alpha = B - A$ , то уравнение (12.4) принимает вид  $\int_\alpha df = f(B) - f(A)$ . Второй закон Кирхгофа утверждает, что для узлов существует такая функция  $\Phi$ , т.е. 0-коцепь, что  $V^\alpha = \Phi(A) - \Phi(B)$  для любой ветви  $\alpha$  такой, что  $\partial\alpha = B - A$ . Тогда этот закон Кирхгофа можно записать в виде

$$\mathbf{V} = -d\Phi.$$

(Заметим, что если уравнение  $\int_K \mathbf{V} = -\int_{\partial K} \Phi$  справедливо для всех ветвей, то оно справедливо для всех 1-цепей, поскольку ветви образуют базис пространства 1-цепей. Тогда  $\mathbf{V}$  и  $-d\Phi$  принимают одно и то же значение на всех 1-цепях, т.е.  $\mathbf{V} = -d\Phi$ .)

Из двух законов Кирхгофа следует *теорема Теллегена*. Предположим, что для данной электрической цепи  $\mathbf{I} \in C_1$  обозначает распределение токов по ветвям, которое удовлетворяет первому закону Кирхгофа  $\partial\mathbf{I} = 0$ . Предположим также, что для той же цепи распределение напряжений  $\mathbf{V}$  удовлетворяет второму закону Кирхгофа  $\mathbf{V} = -d\Phi$ . Тогда полная мощность, выделенная во всех ветвях, равна

$$P = \sum V^\alpha I_\alpha = \int_{\mathbf{I}} \mathbf{V} = - \int_{\mathbf{I}} d\Phi.$$

Однако, по определению оператора  $d$ ,

$$P = - \int_{\mathbf{I}} d\Phi = - \int_{\partial\mathbf{I}} \Phi = 0,$$

поскольку  $\partial \mathbf{I} = 0$ . Этот результат показывает, что в электрической цепи сохраняется энергия — батареи и генераторы подают ровно столько энергии, сколько рассеивается в резисторах. Важно отметить, что никакие предположения о природе элементов цепи не делаются.

Для рассмотрения токов оказалось полезным ввести два подпространства — ядро и образ граничного отображения  $\partial$ : подпространство циклов  $Z_1 = \text{Ker } \partial \subset C_1$  и подпространство границ  $B_0 = \text{Im } \partial \subset C_0$ . Давайте сделаем то же самое для кограничного оператора  $d$ .

Пусть  $Z^0 \subset C^0$  обозначает ядро оператора  $d$ . Тогда  $Z^0$  является подпространством потенциальных функций на узлах, обладающих тем свойством, что все напряжения на ветвях равны нулю. Для связного комплекса все напряжения будут равны нулю тогда и только тогда, когда потенциал во всех узлах одинаков. В общем случае  $Z^0$  есть подпространство потенциалов, постоянных на каждой связной компоненте. Сложение элемента из  $Z^0$  с элементом из  $C^0$  не имеет физического смысла — при этом изменяется только произвольный уровень, от которого отсчитывается каждый потенциал.

Пусть  $\Phi$  является элементом пространства  $Z^0$ , так что  $d\Phi = 0$ . Пусть  $\partial \mathbf{K}$  обозначает произвольный элемент пространства границ  $B_0$ . Тогда

$$\int_{\partial \mathbf{K}} \Phi = \int_{\mathbf{K}} d\Phi = 0.$$

Другими словами, элемент из  $Z^0$ , действуя на элемент из  $B_0$ , дает нуль. И обратно, предположим, что  $\int_{\partial \mathbf{K}} \Phi = 0$  для всех  $\mathbf{K}$ . Тогда  $\int_{\mathbf{K}} d\Phi = 0$  для всех  $\mathbf{K}$ . Следовательно,  $d\Phi = 0$ , т. е. это нулевая функция в пространстве  $C^1$  и  $\Phi \in Z^0$ . Именно поэтому говорят, что  $Z^0$  есть пространство, *аннулирующее*  $B_0$ .

Рассмотрим факторпространство  $H_0 = C_0/B_0$ , элементы которого являются классами эквивалентности элементов пространства  $C_0$ , разность которых лежит в  $B_0$ . Пусть  $c_1$  и  $c_2$  являются элементами пространства  $C_0$ , разность которых  $c_1 - c_2$  лежит в  $B_0$ . Тогда для любого  $\Phi \in Z^0$  имеем

$$\int_{c_1} \Phi - \int_{c_2} \Phi = \int_{c_1 - c_2} \Phi = 0.$$

Следовательно,  $\int_{c_1} \Phi = \int_{c_2} \Phi$ , если  $c_1 \equiv c_2 \pmod{B_0}$ . Итак, элемент  $Z^0$  принимает одно и то же значение на любом представителе данного класса эквивалентности. Поэтому  $Z^0$  может рассматриваться как пространство линейных функций на  $H_0$ .

Связь пространства  $Z^0$  с  $C_0$ ,  $B_0$  и  $H_0$  является примером общего принципа, с которым мы встретились: *ядро отображения, сопряженного к линейному преобразованию, является аннулирующим подпространством образа этого преобразования*. Оно является также *дуальным* пространством *факторпространства* по образу преобразования. В нашем случае мы имеем преобразование  $\partial$ , действующее на  $C_1$ ; сопряженное преобразование  $d$ , действующее в  $C^0$ , и ядро сопряжения  $Z^0$ , являющееся аннулирующим пространством  $B_0 = \text{Im } \partial$  и дуальным к пространству  $H_0 = C_0/B_0$ .

А теперь рассмотрим факторпространство  $P^0 = C^0/Z^0$ . Поскольку  $Z^0$  состоит из функций, постоянных на каждой связной компоненте, элементы пространства  $P^0$  являются классом эквивалентности функций, отличающихся на каждой связной компоненте только постоянной. Итак, если мы к вектору из пространства  $C^0$  добавим физически несущественную постоянную, возможно, разную для каждой связной компоненты, то результирующий вектор остается тем же элементом пространства  $P^0$ . Когда мы прибавляем произвольную аддитивную константу к каждой связной компоненте, то именно  $P^0$  соответствует пространству потенциалов. Для данного вектора напряжений  $\mathbf{V}$ , удовлетворяющего второму закону Кирхгофа, соответствующий потенциал  $\Phi$ , удовлетворяющий условию  $\mathbf{V} = -d\Phi$ , определен с точностью до элемента пространства  $P^0$ , которое определяет  $\Phi$ .

## 12.4. Базисы и дуальные базисы

Чтобы выбрать базис для  $P^0$ , удобно в каждой связной компоненте *заземлить* один узел. После этого выбрать  $\dim C^0 - \dim Z^0$  базисных векторов, для которых потенциал равен единице в одном незаземленном узле и нулю во всех остальных узлах. Например, в цепи, изображенной на рис. 12.23, узлы  $A$  и  $D$  заземлены, а вектор  $(\Phi^B, \Phi^C, \Phi^E)$  представляет произвольный элемент  $P^0$ .

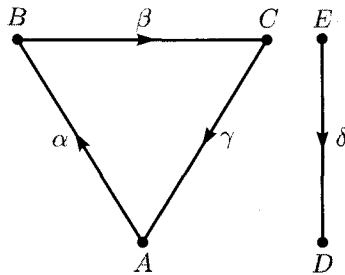


Рис. 12.23

Отображение  $d : C^0 \rightarrow C^1$  индуцирует отображение  $[d] : P^0 \rightarrow C^1$ , которое будем называть *ограниченным кограничным отображением*. Матричное представление  $[d]$  получается из матричного представления  $d$ , если убрать столбцы, соответствующие заземленным узлам. Так, для цепи, изображенной на рис. 12.23, мы убираем первый столбец ( $A$ ) и четвертый ( $D$ ), т. е.

$$d = \begin{pmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}, \quad \text{а} \quad [d] = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Поскольку  $Z^0$  — аннулирующее пространство для  $B_0$ , любые два элемента  $C^0$ , отличающиеся на элемент  $Z^0$ , дают одинаковый результат при действии на элемент  $B_0$ . Так факторпространство  $P^0 = C^0/Z^0$  может быть идентифицировано с дуальным пространством  $B_0$ . Сопряженным к отображению  $[d] : P^0 \rightarrow C^1$  является ограниченный граничный оператор  $[\partial] : C_1 \rightarrow B_0$ . Для  $B_0$  удобно пользоваться базисом, дуальным к базису, выбранному для  $P^0$ . Тогда матричное представление  $[\partial]$  есть просто транспонированное матричное представление  $[d]$ . По отношению к этому базису вектор в пространстве  $B_0$  есть вектор в пространстве  $C_0$ , из которого убраны компоненты, соответствующие заземленным узлам. Например, для цепи, изображенной на рис. 12.23, где заземлены узлы  $A$  и  $D$ , мы имеем

$$\partial\alpha = B - A.$$

В векторных обозначениях это выглядит так:

$$\partial \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{и} \quad [\partial] \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix};$$

кроме того,

$$\partial = \begin{pmatrix} -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \text{но} \quad [\partial] = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = [d]^T.$$

Пусть  $B^1$  обозначает подпространство  $C^1$ , являющееся образом  $d$ . С физической точки зрения  $B^1$  — это пространство векторов напряжений на ветвях, подчиняющихся второму закону Кирхгофа. Более того, согласно теореме Теллегена, доказанной нами ранее,  $B^1$  является аннулирующим пространством для  $Z_1$ : любое распределение напряжений, удовлетворяющее второму закону Кирхгофа, при применении к любому распределению токов, подчиняющихся первому закону Кирхгофа, дает нуль. Если мы образуем факторпространство  $H^1 = C^1/B^1$ , то его можно идентифицировать с дуальным пространством  $H_1$  (известным также как  $Z_1$ ). Доказательство этого утверждения очень простое. Пусть  $\mathbf{V}_1$  и  $\mathbf{V}_2$  — вектора в пространстве  $C^1$ , а разность между ними лежит в пространстве  $B^1$ , так что  $\mathbf{V}_1 - \mathbf{V}_2 = -d\Phi$  для некоторой функции  $\Phi$ . Тогда для произвольного вектора тока  $\mathbf{I} \in Z_1$ , удовлетворяющего условию  $\partial \mathbf{I} = 0$ , имеем

$$\int_{\mathbf{I}} \mathbf{V}_1 - \int_{\mathbf{I}} \mathbf{V}_2 = \int_{\mathbf{I}} (\mathbf{V}_1 - \mathbf{V}_2) = - \int_{\mathbf{I}} d\Phi = - \int_{\partial \mathbf{I}} \Phi = 0,$$

откуда следует, что  $\int_{\mathbf{I}} \mathbf{V}_1 = \int_{\mathbf{I}} \mathbf{V}_2$ . Таким образом,  $\mathbf{V}_1$  и  $\mathbf{V}_2$ , которые соответствуют одному и тому же элементу  $H^1$ , совпадают как элементы дуального пространства  $H_1$ .

Теперь подведем итог соотношениям между пространствами  $C^0$ ,  $C^1$ ,  $Z^1$ ,  $B^1$  и  $Z_1$ .

- Преобразование  $d$  действует в пространстве  $C^0$  и отображает  $C^0 \rightarrow C^1$ .
- Его сопряжение  $\partial$  действует в пространстве  $C_1$ , отображая  $C_1 \rightarrow C_0$ .
- Ядро  $\partial$  есть  $Z_1$ , его аннулятором является  $B^1 = \text{Im } d$ , а его дуальное пространство есть  $H^1 = C^1/B^1$ .

## 12.5. Методы Максвелла

Мы рассмотрели процедуру построения базиса ячеек в пространстве  $Z_1$  путем выбора максимального дерева. Пусть  $\sigma$  обозначает отображение, которое определяет каждый вектор в  $Z_1 = H_1$  как элемент пространства  $C_1$ . Теперь выберем базис для  $H^1$ , который будет дуален базису, выбранному для  $H_1$ . Пусть  $s$  обозначает проекцию пространства  $C^1$  на факторпространство  $H^1 = C^1/B^1$ . Тогда  $s$  и  $\sigma$  — взаимно сопряженные преобразования, и матричное представление одного будет транспонированным представлением другого.

Матрицы  $\sigma$  и  $s$  построить очень просто. Рассмотрим цепь, изображенную на рис. 12.24, с максимальным деревом, состоящим из  $\alpha$  и  $\gamma$ . Тогда базис пространства  $Z_1$  имеет вид

$$\mathbf{M}_1 = \alpha + \beta + \gamma = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{M}_2 = -\gamma + \delta = \begin{pmatrix} 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

Следовательно, матрица преобразования  $\sigma$  есть

$$\begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \\ 0 & +1 \end{pmatrix}.$$

Заметим, что

$$\sigma \left( \begin{pmatrix} J_1 \\ J_2 \end{pmatrix} \right) = \begin{pmatrix} J_1 \\ J_1 \\ J_1 - J_2 \\ J_2 \end{pmatrix},$$

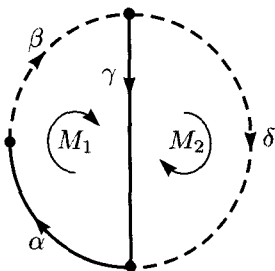


Рис. 12.24

т. е.  $\sigma$  определяет правило, по которому токи в ветвях вычисляются через круговые токи.

Матрица преобразования  $s$  — транспонированная матрица  $\sigma$ :

$$s = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

Заметим, что

$$s \begin{pmatrix} v^\alpha \\ v^\beta \\ v^\gamma \\ v^\delta \end{pmatrix} = \begin{pmatrix} v^\alpha + v^\beta + v^\gamma \\ -v^\gamma + v^\delta \end{pmatrix},$$

откуда следует, что компоненты вектора  $s\mathbf{V}$  равны напряжениям на ветвях, просуммированным по всем ячейкам, выбранным в качестве базиса  $H_1$ .

Полученные соотношения сводятся к следующему:

$$\boxed{\begin{array}{l} P^0 \xrightarrow{[d]} C^1 \xrightarrow{s} H^1, \\ B_0 \xleftarrow{[\partial]} C_1 \xleftarrow{\sigma} H_1. \end{array}}$$

Векторные пространства, расположенные одно над другим, дуальны друг к другу:  $C^1$  и  $C_1$ ;  $P^0$  и  $B_0$ ;  $H^1$  и  $H_1$ . Более того,  $[d]$  и  $[\partial]$  сопряжены, так же, как  $s$  и  $\sigma$ . Оба отображения  $[d]$  и  $\sigma$  инъективны (имеют нулевое ядро), в то время как  $[\partial]$  и  $s$  сюръективны (все векторное пространство является их образом). Наконец,

$\text{Im } \sigma = \text{Ker } [\partial]$  и  $\text{Im } [d] = \text{Ker } s$ . Учитывая это, законы Кирхгофа можно записать в симметричной форме.

Первый закон Кирхгофа: $\mathbf{I} = \sigma \mathbf{J}, \quad \mathbf{J} \in H_1;$
Второй закон Кирхгофа: $\mathbf{V} = -[d]\Phi, \quad \Phi \in P^0.$

Теперь мы можем обсудить теорему Кирхгофа для цепей с резисторами. Предположим, что наша электрическая цепь состоит из резисторов, вольт-амперная характеристика которых для каждой ветви  $\alpha$  является прямой линией на плоскости  $(I_\alpha, V^\alpha)$ . Сделаем реалистичное (и упрощающее) предположение, что эта прямая не параллельна ни одной из осей. Это значит, что любой источник напряжения соединен с некоторым резистором последовательно, а любой источник тока соединен с сопротивлением параллельно, т. е. нет источников, подающих постоянное напряжение при любом токе, и нет источников тока, создающих постоянный ток независимо от напряжения. Таким образом, мы предполагаем, что любая ветвь  $\alpha$  с  $\partial\alpha = B - A$  имеет вид, изображенный на рис. 12.25.<sup>13</sup>

Напряжение на резисторе равно  $V^\alpha - W^\alpha$ , а ток, протекающий по нему равен  $I_\alpha - K_\alpha$ . Тогда вольт-амперная характеристика системы определяется формулой

$$(V^\alpha - W^\alpha) = z_\alpha(I_\alpha - K_\alpha).$$

Здесь при  $z_\alpha \neq 0$  заданы значения  $W^\alpha$  и  $K_\alpha$  (причем любое из них может быть равно нулю<sup>14</sup>). Если в цепи есть только резисторы,

<sup>13</sup>Показано широко используемое в элементарных курсах физики разложение *реального* источника питания на простой резистор и *идеальный* источник напряжения (с нулевым внутренним сопротивлением) или тока (с «бесконечно большими» сопротивлением  $R_\alpha$  и напряжением  $U^\alpha = R_\alpha K_\alpha$ ). — *Прим. ред.*

<sup>14</sup>Авторы, вероятно, имеют в виду, что для создания вышеописанной линейной вольт-амперной характеристики вида  $V^\alpha = z_\alpha(I_\alpha - K_\alpha) = z_\alpha I_\alpha - W^\alpha$  достаточно добавить к резистору источник только одного (любого) из этих двух типов. Принято называть  $W^\alpha$  *напряжением холостого хода* реального источника, а  $K_\alpha$  — его *током короткого замыкания*. Отметим, что ветвь с двумя источниками, удовлетворяющими соотношению  $W^\alpha = z_\alpha K_\alpha$  (при соблюдении указанной полярности), ведет себя с точки зрения остальной цепи как простой резистор. — *Прим. ред.*

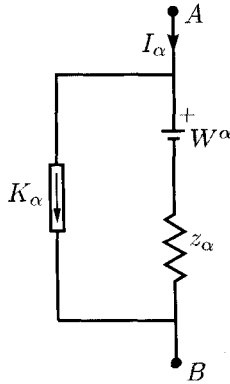


Рис. 12.25

то  $z_\alpha$  — положительное вещественное число. Мы можем также рассмотреть стационарное состояние линейной цепи, содержащей конденсаторы и индукторы. В этом случае  $z_\alpha$  — комплексное число, зависящее от частоты. Величины  $I_\alpha$  и  $V^\alpha$  неизвестны. Пусть  $Z : C_1 \rightarrow C^1$  — линейное преобразование, матрица которого в базисе, состоящем из всех ветвей, диагональна, и на диагонали стоят соответствующие  $z_\alpha$ . Тогда предыдущее уравнение можно переписать в форме

$$\mathbf{V} - \mathbf{W} = Z(\mathbf{I} - \mathbf{K}).$$

Комбинируя это соотношение с законами Кирхгофа, получаем уравнения

$$\mathbf{V} - \mathbf{W} = Z(\mathbf{I} - \mathbf{K}), \quad \mathbf{I} = \sigma \mathbf{J}, \quad \mathbf{V} = -[d]\Phi$$

для неизвестных  $\mathbf{J} \in H_1$  и  $\Phi \in P^0$ , при заданных  $\mathbf{W}$ ,  $\mathbf{K}$  и  $Z$ . Эти уравнения мы будем пытаться решать одним из двух методов Максвелла, описанных ниже.

### Метод круговых токов

(i) Запишем  $\mathbf{I} = \sigma \mathbf{J}$ , чтобы удовлетворялся первый закон Кирхгофа. После этого применим  $s$  к уравнению  $\mathbf{V} - \mathbf{W} = Z(\sigma \mathbf{J} - \mathbf{K})$ . Поскольку из второго закона Кирхгофа  $s\mathbf{V} = s(-[d]\Phi) = 0$ , мы

получаем  $-s\mathbf{W} = sZ\sigma\mathbf{J} - sZ\mathbf{K}$  или  $(sZ\sigma)\mathbf{J} = s(Z\mathbf{K} - \mathbf{W})$ . Правая часть этого равенства известна, как и линейное преобразование  $sZ\sigma$ . Если преобразование  $sZ\sigma$  обратимо, то

$$\mathbf{J} = (sZ\sigma)^{-1}s(Z\mathbf{K} - \mathbf{W}), \quad (12.5)$$

из которого могут быть получены все токи и напряжения, так как  $\mathbf{I} = \sigma\mathbf{J}$ ,  $\mathbf{V} = \mathbf{W} + Z(\mathbf{I} - \mathbf{K})$ . В этом состоит метод круговых токов. Он зависит от обратимости преобразования  $sZ\sigma$ .

### Метод узловых потенциалов

(ii) Запишем  $\mathbf{V} = -[d]\Phi$ , чтобы обеспечить выполнение второго закона Кирхгофа. После этого обратим преобразование  $Z$ , которое задается диагональной матрицей, и получим  $\mathbf{I} - \mathbf{K} = Z^{-1}(-[d]\Phi - \mathbf{W})$ . К обеим частям этого равенства применим  $[\partial]$ . Поскольку, в силу первого закона Кирхгофа,  $[\partial]\mathbf{I} = 0$ , мы получаем

$$-[\partial]\mathbf{K} = -[\partial]Z^{-1}[d]\Phi - [\partial]Z^{-1}\mathbf{W}$$

или

$$([\partial]Z^{-1}[d])\Phi = [\partial](\mathbf{K} - Z^{-1}\mathbf{W}).$$

Если преобразование  $[\partial]Z^{-1}[d]$  обратимо, мы получаем

$$\Phi = ([\partial]Z^{-1}[d])^{-1}[\partial](\mathbf{K} - Z^{-1}\mathbf{W}), \quad (12.6)$$

после чего все токи и напряжения находятся из уравнения

$$\mathbf{V} = -[d]\Phi, \quad \mathbf{I} = \mathbf{K} + Z^{-1}(\mathbf{V} - \mathbf{W}). \quad (12.7)$$

В этом состоит метод узловых потенциалов. Он зависит от обратимости преобразования  $[\partial]Z^{-1}[d]$ .

## 12.6. Матричные выражения для операторов

Уравнения (12.5)–(12.7) дают нам суть методов Максвелла. Чтобы можно было с ними работать, надо получить матричные выражения для операторов. Поэтому необходимо выбрать соответствующие базисы и базисы, дуальные к ним. В этом параграфе мы еще раз посмотрим, как это делается и приведем несколько

примеров. *Доказательство* того, что эти методы всегда работают в случае цепей с резисторами, мы дадим в начале параграфа 12.7.

Базис для пространств  $B_0$ ,  $H_0$ ,  $Z^0$  и  $P^0$  определяется выбором заземленного узла в каждой связной компоненте цепи. Например, на рис. 12.26 можно сделать заземление в узле  $A$  для одной связной компоненты и в узле  $E$  для другой. Тогда базис пространства  $B_0$  состоит из границ траекторий, соединяющих заземленные узлы с другими узлами той же компоненты. В этом случае базисные векторы имеют вид

$$\mathbf{b}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{b}_3 = \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{b}_4 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \\ 1 \end{pmatrix}.$$

Чтобы построить базис факторпространства  $H_0 = C_0/B_0$ , надо базис пространства  $B_0$  расширить до базиса пространства  $C_0$ , введя базисные вектора, у которых на месте какого-либо одного заземленного узла стоит 1, а все остальные элементы равны нулю. После этого для каждого из этих векторов берется класс эквивалентности. В нашем примере это дает

$$\mathbf{h}_1 = \overline{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}, \quad \mathbf{h}_2 = \overline{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}.$$

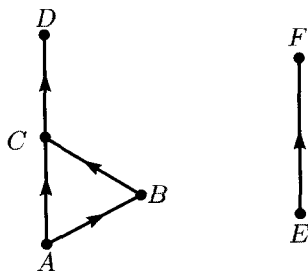


Рис. 12.26

В пространстве  $Z^0$ , являющимся ядром  $d$  и дуальным к  $H_0$ , базисными элементами будут потенциалы с постоянным значением 1 поочередно на каждой связной компоненте. В нашем примере мы имеем

$$\mathbf{z}^1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{z}^2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},$$

которые, очевидно, дуальны  $\mathbf{h}_1$  и  $\mathbf{h}_2$ , приведенным выше. Чтобы выбрать базис факторпространства  $P^0 = C^0/Z^0$ , мы просто положим последовательно потенциалы на каждом незаземленном узле равными 1 и для каждого возьмем класс эквивалентности. В нашем примере мы получим

$$\mathbf{p}^1 = \overline{\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}, \quad \mathbf{p}^2 = \overline{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}, \quad \mathbf{p}^3 = \overline{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}}, \quad \mathbf{p}^4 = \overline{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}}.$$

Очевидно, что эти классы эквивалентности дуальны построенному выше базису  $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4\}$  в пространстве  $B_0$ . Поэтому матричное представление  $[d] : P^0 \rightarrow C^0$  является транспонированным к матричному представлению  $[\partial] : C_0 \rightarrow B_0$ . *Столбцы*  $[d]$  и *строки*  $[\partial]$  отвечают незаземленным узлам, которые соответствуют базисным векторам.

Теперь рассмотрим пространства  $Z_1$ ,  $B^1$  и  $H^1$ , которые связаны с ядром  $\partial$  и образом  $d$ . Для полноты картины рассмотрим также факторпространство  $G_1 = C_1/Z_1$ , дуальное к пространству  $B^1$ . Во всех этих пространствах выбор базиса определяется выбором максимального дерева. На рис. 12.27 изображено такое дерево с ветвями  $\alpha$ ,  $\beta$  и  $\varepsilon$ .

Как мы уже видели, базисные элементы пространства  $Z_1$  — это ячейки, связанные с ветвями, *не* принадлежащими макси-

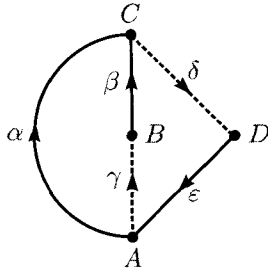


Рис. 12.27

мальному дереву. В нашем примере это будут

$$\mathbf{m}_1 = \begin{pmatrix} -1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{и} \quad \mathbf{m}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix},$$

связанные соответственно с ветвями  $\gamma$  и  $\delta$ . Для факторпространства  $G_1 = C_1/Z_1$ , которое представляет нарушение первого закона Кирхгофа, мы можем образовать классы эквивалентности векторов, имеющих 1 на месте соответствующей ветви максимального дерева и нули в остальных местах. В нашем случае мы имеем

$$\mathbf{g}_1 = \overline{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}, \quad \mathbf{g}_2 = \overline{\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}, \quad \mathbf{g}_3 = \overline{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}}.$$

Поскольку эти вектора имеют 0 на месте каждой ветви, не принадлежащей дереву, то ни одна линейная комбинация из них не может быть ячейкой. Это значит, что они являются независимыми элементами пространства  $G_1$ .

Для подпространства  $B^1$ , образованного кограницами — напряжениями ветвей, подчиняющимися второму закону Кирхгофа, мы можем построить базис, приписывая напряжение, равное 1 одной из ветвей максимального дерева, и нулю всем остальным его ветвям. Такая процедура приписывает потенциал каждому узлу (с точностью до константы), и эти потенциалы можно

использовать для определения напряжений на ветвях, не принадлежащих дереву. Или, другими словами, поскольку каждая ветвь, не принадлежащая максимальному дереву, связана с единственной ячейкой, содержащей эту ветвь и другие ветви, которые принадлежат дереву, то каждую такую ячейку можно использовать для определения напряжений на таких ветвях. Например на рис. 12.27 эта процедура приводит к базисным векторам, связанным соответственно с ветвями  $\alpha$ ,  $\beta$  и  $\varepsilon$ :

$$\mathbf{b}^1 = (1 \ 0 \ 1 \ -1 \ 0),$$

$$\mathbf{b}^2 = (0 \ 1 \ -1 \ 0 \ 0),$$

$$\mathbf{b}^3 = (0 \ 0 \ 0 \ -1 \ 1).$$

Эти вектора, аннулирующие ячейки  $\mathbf{m}_1$  и  $\mathbf{m}_2$ , дуальны базису  $\mathbf{g}_1$ ,  $\mathbf{g}_2$ ,  $\mathbf{g}_3$  пространства  $G_1$ . Они не совпадают со столбцами  $[d]$ , но они задают то же самое подпространство пространства  $C^1$ .

Итак, базисные вектора пространства  $B^1$  могут ассоциироваться с ветвями, принадлежащими максимальному дереву. Точно так же базисные вектора  $H^1 = C^1/B^1$  ассоциируются с ветвями, не принадлежащими дереву. Они являются классами эквивалентности векторов напряжений, которые имеют 1 на месте одной ветви, не принадлежащей дереву, и нули на всех остальных местах. Таким образом, в нашем примере базисные вектора пространства  $H_1$  имеют вид

$$\mathbf{h}^1 = \overline{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}}, \quad \mathbf{h}^2 = \overline{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}.$$

Они соответствуют *единичным нарушениям* второго закона Кирхгофа в ячейках 1 и 2 соответственно. Очевидно, что  $\mathbf{h}^1$  и  $\mathbf{h}^2$  дуальны ячейкам  $\mathbf{m}_1$  и  $\mathbf{m}_2$ . Вследствие этого матрица преобразования  $s : C^1 \rightarrow H^1$  является транспонированным по отношению к матрице  $\sigma : Z_1 \rightarrow C_1$ .

Иногда удобнее пользоваться другим методом выбора базисных векторов в пространстве  $B^1$  (напряжений, подчиняющихся

второму закону Кирхгофа). В этом случае потенциал выбирается равным 1 последовательно в каждом незаземленном узле и нулю во всех остальных узлах. На получившиеся потенциалы действуем оператором  $d$ . (Или, что то же самое, можно взять потенциалы равными  $-1$  последовательно в каждом незаземленном узле и записать напряжения на ветвях). В нашем примере заземлен узел  $A$ . Базисные вектора, связанные с узлами  $B$ ,  $C$  и  $D$  соответственно, имеют вид:

$$\mathbf{a}^1 = (0 \ -1 \ 1 \ 0 \ 0),$$

$$\mathbf{a}^2 = (1 \ 1 \ 0 \ -1 \ 0),$$

$$\mathbf{a}^3 = (0 \ 0 \ 0 \ 1 \ -1).$$

В качестве дуального базиса в пространстве  $G_1 = C_1/Z_1$  мы должны выбрать распределения токов, для которых единичный ток протекает из заземленного узла вдоль дерева поочередно к каждому незаземленному узлу. В нашем примере мы получаем базисные вектора

$$\mathbf{k}_1 = \overline{\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}, \quad \mathbf{k}_2 = \overline{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}}, \quad \mathbf{k}_3 = \overline{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}},$$

соответствующие единичным нарушениям первого закона Кирхгофа в незаземленных узлах  $B$ ,  $C$  и  $D$  соответственно.

Давайте применим оба метода Максвелла к цепи, изображенной на рис. 12.28. Ее топология уже знакома нам по рис. 12.24. Сначала мы произвольно выбираем заземленный узел, пусть это будет узел  $A$ . Тогда  $\Phi^A = 0$ . Напряжения в ветвях выражаются через потенциалы в других узлах следующим образом:

$$V^\alpha = -\Phi^B, \quad V^\beta = \Phi^B - \Phi^C, \quad V^\gamma = \Phi^C, \quad V^\delta = \Phi^C.$$

Поэтому для того, чтобы получить  $V = -[d]\Phi$ , мы находим

$$[d] = \begin{pmatrix} +1 & 0 \\ -1 & +1 \\ 0 & -1 \\ 0 & -1 \end{pmatrix}.$$

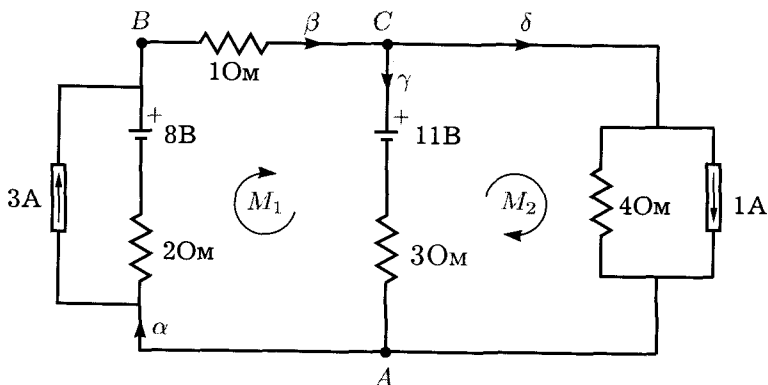


Рис. 12.28

Чтобы это проверить, построим матрицу  $\partial$ . Поскольку

$$\partial\alpha = B - A, \quad \partial\beta = C - B, \quad \partial\gamma = A - C, \quad \partial\delta = A - C,$$

мы имеем:

$$\partial = \begin{pmatrix} -1 & 0 & +1 & +1 \\ +1 & -1 & 0 & 0 \\ 0 & +1 & -1 & -1 \end{pmatrix}.$$

Чтобы получить  $[\partial]$ , уберем из матрицы  $\partial$  строку, соответствующую заземленному узлу A. Итак, убирая первую строку, мы получаем

$$[\partial] = \begin{pmatrix} +1 & -1 & 0 & 0 \\ 0 & +1 & -1 & -1 \end{pmatrix}.$$

Как и ожидалось, матрица  $[\partial]$  является транспонированной матрицей  $[d]$ .

Если выбрать ячейки  $M_1$  и  $M_2$ , как показано на рис. 12.28, то токи в ветвях выражаются через токи в ячейках  $J_1$  и  $J_2$  следующим образом:

$$I_\alpha = J_1, \quad I_\beta = J_1, \quad I_\gamma = J_1 - J_2, \quad I_\delta = J_2.$$

Поскольку  $\mathbf{I} = \sigma \mathbf{J}$ , мы имеем

$$\sigma = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Для проверки построим матрицу  $s$ , выражая падение напряжения  $\mathcal{E}$  в каждой ячейке через напряжения на ветвях.

$$\text{В ячейке 1: } \mathcal{E}_1 = V^\alpha + V^\beta + V^\gamma.$$

$$\text{В ячейке 2: } \mathcal{E}_2 = -V^\gamma + V^\delta.$$

Поскольку  $\mathcal{E} = s\mathbf{V}$ , мы имеем:

$$s = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}.$$

Как и ожидалось, матрица  $s$  является транспонированной к матрице  $\sigma$ .

Далее мы должны записать вектора  $\mathbf{W}$  и  $\mathbf{K}$ , которые представляют источники напряжения и тока в четырех ветвях. Договариваемся считать напряжение батареи  $\mathbf{W}$  положительным, если направление убывания потенциала (от положительного полюса к отрицательному, противоположное направлению ЭДС) совпадает с направлением прохода ветви, указанным стрелкой, в противном случае — отрицательным. Для цепи, изображенной на рис. 12.28, мы имеем:

$$W^\alpha = -8, \quad W^\beta = 0, \quad W^\gamma = +11, \quad W^\delta = 0.$$

Аналогично, источник тока дает положительный вклад, если его направление совпадает с направлением прохода ветви по стрелке, и отрицательным, если наоборот. Следовательно,

$$K_\alpha = +3, \quad K_\beta = 0, \quad K_\gamma = 0, \quad K_\delta = +1.$$

Теперь мы можем ввести вектора для описания источников в цепи:

$$\mathbf{W} = \begin{pmatrix} -8 \\ 0 \\ 11 \\ 0 \end{pmatrix} \quad \text{и} \quad \mathbf{K} = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Наконец, запишем матрицу  $Z$ , описывающую резисторы. Поскольку

$$R_\alpha = 2, \quad R_\beta = 1, \quad R_\gamma = 3, \quad R_\delta = 4,$$

мы имеем:

$$Z = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Чтобы применить метод круговых токов, запишем уравнение

$$\mathbf{J} = (sZ\sigma)^{-1}s(Z\mathbf{K} - \mathbf{W}).$$

Перемножение матриц дает

$$sZ\sigma = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 6 & -3 \\ -3 & 7 \end{pmatrix},$$

следовательно,

$$(sZ\sigma)^{-1} = \frac{1}{33} \begin{pmatrix} 7 & 3 \\ 3 & 6 \end{pmatrix}.$$

Кроме того, получаем

$$Z\mathbf{K} - \mathbf{W} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} -8 \\ 0 \\ 11 \\ 0 \end{pmatrix} = \begin{pmatrix} 14 \\ 0 \\ -11 \\ 4 \end{pmatrix},$$

откуда

$$s(Z\mathbf{K} - \mathbf{W}) = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 14 \\ 0 \\ -11 \\ 4 \end{pmatrix} = \begin{pmatrix} 3 \\ 15 \end{pmatrix}.$$

И, наконец,

$$\mathbf{J} = \frac{1}{33} \begin{pmatrix} 7 & 3 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} 3 \\ 15 \end{pmatrix} = \frac{1}{33} \begin{pmatrix} 66 \\ 99 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix},$$

откуда получаем  $J_1 = 2$  и  $J_2 = 3$ . Итак,

$$\mathbf{I} = \sigma\mathbf{J} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ -1 \\ 3 \end{pmatrix}.$$

Если внимательно проследить за процедурой вычисления, то можно убедиться, что для любой цепи справедливы следующие утверждения.

1. В матрице  $sZ\sigma$  каждый диагональный элемент равен сумме сопротивлений ветвей, образующих ячейку. В нашем случае ячейка  $M_1$ , состоящая из ветвей  $\alpha$ ,  $\beta$  и  $\gamma$ , имеет полное сопротивление, равное 6 омам — левый верхний элемент матрицы.
2. В матрице  $sZ\sigma$  каждый недиагональный элемент равен сопротивлению в ветви, общей для двух ячеек (с точностью до знака  $\pm$ ). В рассматриваемой цепи ячейки 1 и 2 имеют общее сопротивление, равное 3 омам, а знак — возникает из-за того, что ветвь  $\gamma$  в ячейках  $M_1$  и  $M_2$  проходит в разных направлениях.
3. Элементы матрицы  $(sZ\sigma)J$  равны падению напряжения, которое возникло бы при обходе каждой ячейки при отсутствии источников<sup>15</sup>. Например, если  $J_1 = 2$  и  $J_3 = 3$ , то при обходе ячейки 1 возникает падение напряжения в 3 вольта.
4. Элементы матрицы  $ZK - W$  равны скачкам напряжения<sup>16</sup> в ветвях, если бы все токи в ячейках были равны нулю. Например, если  $J_1 = 0$ , то в ветви  $\alpha$  будет ток в 3 ампера, протекающий через резистор с сопротивлением в 2 ома. Тогда скачок напряжения в этой ветви будет равен  $(2 \cdot 3 + 8) = 14$  вольтам.
5. Элементы матрицы  $s(ZK - W)$  являются скачками напряжения при обходе различных ячеек, если бы все токи в ветвях были равны нулю. Например, если  $J_1 = J_2 = 0$ , то при обходе ячейки 1 полный скачок напряжения будет равен 3 вольтам.

<sup>15</sup>и тех же токах. Это падение напряжения в замкнутом контуре (короткое замыкание) можно интерпретировать как (катастрофический) результат мгновенного отключения ЭДС источника. — *Прим. ред.*

<sup>16</sup>Этим (не слишком корректно звучащим) термином обозначается возрастание потенциала (voltage rise) вдоль участка цепи, в отличие от убывания (voltage drop), называемого падением напряжения. — *Прим. ред.*

6. Поэтому уравнение  $(sZ\sigma)\mathbf{J} = s(Z\mathbf{K} - \mathbf{W})$  означает, что для каждой ячейки суммарное падение напряжения в ячейках, создаваемое токами, протекающими через резисторы, равно сумме скачков напряжения, которые получились бы при условии, что все токи в ячейках равны нулю. Тогда полное падение напряжения при обходе каждой ячейки при имеющихся токах в ячейках и источниках равно нулю, что и требуется вторым законом Кирхгофа.

Чтобы воспользоваться методом узловых потенциалов, рассмотрим уравнение

$$\Phi = ([\partial]Z^{-1}[d])^{-1}[\partial](\mathbf{K} - Z^{-1}\mathbf{W}).$$

Перемножение матриц дает

$$\begin{aligned} [\partial]Z^{-1}[d] &= \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 3/2 & -1 \\ -1 & 19/12 \end{pmatrix}. \end{aligned}$$

Откуда

$$([\partial]Z^{-1}[d])^{-1} = \frac{24}{33} \begin{pmatrix} 19/12 & 1 \\ 1 & 3/2 \end{pmatrix}.$$

Кроме того, получаем

$$\mathbf{K} - Z^{-1}\mathbf{W} = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix} - \begin{pmatrix} 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/4 \end{pmatrix} \begin{pmatrix} -8 \\ 0 \\ 11 \\ 0 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \\ -11/3 \\ 1 \end{pmatrix},$$

следовательно,

$$[\partial](\mathbf{K} - Z^{-1}\mathbf{W}) = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 7 \\ 0 \\ -11/3 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 8/3 \end{pmatrix}.$$

Тогда

$$\Phi = \frac{24}{33} \begin{pmatrix} 19/12 & 1 \\ 1 & 3/2 \end{pmatrix} \begin{pmatrix} 7 \\ 8/3 \end{pmatrix} = \frac{24}{33} \begin{pmatrix} 165/12 \\ 11 \end{pmatrix} = \begin{pmatrix} 10 \\ 8 \end{pmatrix},$$

откуда получаем неизвестные потенциалы на узлах  $\Phi^B = 10$ ,  $\Phi^C = 8$ .

Здесь также, внимательно проследив процедуру вычисления, можно увидеть, что для любой цепи справедливы следующие утверждения.

1. Диагональные элементы матрицы  $[\partial]Z^{-1}[d]$  равны сумме обратных сопротивлений (проводимостей) всех резисторов, подходящих к узлу. В нашем примере к узлу  $B$  подсоединены два резистора с сопротивлениями 1 и 2 ома. Поэтому левый верхний элемент равен  $1 + 1/2 = 3/2$ .
2. Недиагональные элементы матрицы  $[\partial]Z^{-1}[d]$  равны (с точностью до знака  $\pm$ )<sup>17</sup> обратному сопротивлению в ветви, соединяющей два соответствующих узла. Поскольку у нас узлы  $B$  и  $C$  соединены резистором с сопротивлением в 1 ом, то недиагональный элемент матрицы равен  $-1$ .
3. Компоненты вектора  $[\partial]Z^{-1}[d]\Phi$  представляют суммарный ток, который вытекал бы из каждого узла, если убрать из цепи источники. Так, в вышеуказанном примере, если нет источников,  $\Phi^A = 0$  и  $\Phi^B = 8$ , то ток в 5 ампер будет течь от узла  $B$  к узлу  $A$  через резистор с сопротивлением в 2 ома, и ток в 2 ампера будет течь от узла  $B$  к узлу  $C$  через резистор с сопротивлением 1 ом, создавая полный ток в 7 ампер. Можно проверить, что первая компонента вектора  $[\partial]Z^{-1}[d]\Phi$  равна 7.
4. Компонентами вектора  $\mathbf{K} - Z^{-1}\mathbf{W}$  являются токи в ветвях, которые возникли бы при условии, что все узловые потенциалы обратились в нуль. В нашем случае, если  $\Phi^A$  и  $\Phi^B$

<sup>17</sup>Выбор знака, как и в методе круговых токов (пункт 2 на с. 623), зависит от направления прохода ветви. Точную формулировку оставим читателю в качестве упражнения. — *Прим. ред.*

приравнять нулю, то в ветви  $\alpha$  через резистор с сопротивлением в 2 ома протекал бы ток в 4 ампера, а полный ток был бы равен  $4 + 3 = 7$  амперам. Первая компонента тока  $\mathbf{K} - \mathbf{Z}^{-1}\mathbf{W}$  действительно равна 7.

5. Компонентами вектора  $[\partial](\mathbf{K} - \mathbf{Z}^{-1}\mathbf{W})$  являются токи, которые втекали бы в каждый узел при условии, что все узловые потенциалы равны нулю. Итак, если  $\Phi^A = \Phi^B = \Phi^C = 0$ , то в узел  $B$  будет втекать ток в 7 ампер по ветви  $\alpha$ . Не будет никакого тока в ветви  $\beta$ , текущего к узлу  $C$ . Поэтому суммарный ток в узле  $B$  будет равен 7 амперам.
6. Уравнение  $[\partial]\mathbf{Z}^{-1}[d]\Phi = [\partial](\mathbf{K} - \mathbf{Z}^{-1}\mathbf{W})$  показывает, что сумма токов, вытекающих из каждого узла по ветвям через резисторы в силу имеющейся разности потенциалов на узлах, равна сумме токов, втекающих в этот узел, если бы все узловые потенциалы были равны нулю. Поэтому полный ток, втекающий в каждый узел в силу имеющихся разности потенциалов и источников, равен нулю, что соответствует первому закону Кирхгофа.

Если теперь сравнить шесть утверждений, касающихся метода узловых потенциалов, с шестью утверждениями относительно метода круговых токов, то можно увидеть их двойственность. Замените термин *узел* на *ячейку*, *ток* на *напряжение* и *сопротивление* на *проводимость* — и факты замечательно совпадут.

В рассмотренном примере есть одна особенность, которая не выполняется в общем случае: размер матрицы, которую надо обращать, был одинаковым для двух методов, поскольку там было две независимых ячейки и два незаземленных узла. Однако число независимых ячеек не всегда равно числу незаземленных узлов, поэтому матрицы могут быть разного размера.

## 12.7. Теорема Кирхгофа

Возникает вопрос: будет ли метод круговых токов работать для любой цепи? Эта проблема сводится к обратимости отображения  $s\mathbf{Z}\sigma$ . Поскольку  $s\mathbf{Z}\sigma$  является отображением из пространства  $H_1$

(пространство круговых токов) на пространство  $H^1$ , дуальное к нему (пространство напряжений при обходе ячеек), то оба эти пространства имеют одинаковую размерность. Поэтому нам надо только проверить, что  $sZ\sigma$  — инъективное отображение, т. е. что его ядро равно нулю. Рассмотрим произвольный ненулевой элемент  $\mathbf{J}$  в пространстве  $H_1$ . Его образ  $(sZ\sigma)\mathbf{J}$  является элементом дуального пространства  $H^1$ . Нам надо показать, что этот элемент *не* равен нулю. Сделаем это, просто вычислив его значение на элементе  $\mathbf{J}$ . Поскольку  $s$  является сопряжением  $\sigma$ , то мы имеем  $[(sZ\sigma)\mathbf{J}](\mathbf{J}) = [Z\sigma\mathbf{J}](\sigma\mathbf{J})$ . Однако  $\sigma\mathbf{J}$  есть просто вектор тока  $\mathbf{I}$  в ветви, а  $Z\sigma\mathbf{J} = Z\mathbf{I}$ . Следовательно,

$$[(sZ\sigma)\mathbf{J}]\mathbf{J} = (Z\mathbf{I})\mathbf{I} = \int_{\mathbf{I}} Z\mathbf{I} = \sum_{\gamma} z_{\gamma} I_{\gamma}^2,$$

где сумма берется по всем ветвям. Если у матрицы  $Z$  все диагональные элементы положительны, то  $\int_{\mathbf{I}} Z\mathbf{I} > 0$ , кроме случая, когда  $\mathbf{I} = 0$ . Поскольку  $\sigma$  инъективно, то условие  $\mathbf{I} = \sigma\mathbf{J} = 0$  означает, что  $\mathbf{J} = 0$ , т. е. все токи в ветвях равны нулю, только если равны нулю все токи в ячейках. Таким образом, для ненулевого  $\mathbf{J}$  мы имеем  $[(sZ\sigma)\mathbf{J}]\mathbf{J} > 0$ . Отсюда можно заключить, что  $(sZ\sigma)\mathbf{J}$  не может быть нулевым элементом пространства  $H^1$ , если  $\mathbf{J} \neq 0$ . Поэтому  $sZ\sigma$ , имея нулевое ядро, является обратимым отображением. Следовательно, мы доказали, что метод круговых токов работает для любых цепей, содержащих только резисторы. Аналогичные аргументы показывают, что в этом случае  $[\partial]Z^{-1}[d]$  также обратимо. А это значит, что работает и метод узловых потенциалов.

В более общем случае, когда в цепи имеются резисторы, конденсаторы и индукторы, изложенная процедура также дает интересную информацию. Если в цепи  $b$  ветвей, то  $Z$  задается диагональной матрицей  $b \times b$ , с матричными элементами (импедансами ветвей), зависящими от частоты  $\omega/2\pi$ . Если каждая ветвь содержит только один конденсатор или один индуктор, то импеданс будет равен либо  $i\omega L$ , либо  $-i/\omega C$ . Если в цепи есть  $m$  независимых ячеек, то  $sZ\sigma$  будет матрицей  $m \times m$  с матричными элементами, зависящими от  $\omega$ . Определитель этой матрицы приводится

к полиному  $D(\omega^2)$  степени не выше  $m$ . Оператор  $sZ\sigma$  не будет обратимым при условии  $D(\omega^2) = 0$ . Это значит, что существует не более  $m$  значений  $\omega^2$ , для которых  $sZ\sigma$  необратим. Такие частоты  $\omega/2\pi$  называются *резонансными частотами* системы. Для каждой из этих резонансных частот уравнение  $(sZ\sigma)\mathbf{J} = 0$  имеет ненулевое решение, несмотря на нулевую правую часть, указывающую на отсутствие источников, т. е. токи могут течь по системе даже без источников напряжения или тока. Решения этого уравнения определяют *нормальные моды* системы. Аналогичные соображения, примененные к матрице  $[\partial]Z^{-1}[d]$  размера  $(b-m) \times (b-m)$ , показывают, что может быть не более  $b-m$  резонансных частот. Поэтому для нахождения резонансных частот цепи мы строим матрицу оператора  $sZ\sigma$  или  $[\partial]Z^{-1}[d]$  (у которой меньше размер) и вычисляем частоты  $(\omega^2)$ , для которых определитель матрицы равен нулю.

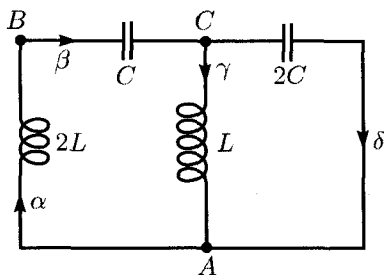


Рис. 12.29

В качестве примера рассмотрим цепь, изображенную на рис. 12.29. Эта цепь имеет ту же топологию, что и цепь на рис. 12.28. Поэтому

$$s = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

В этом случае матрица  $Z$  содержит импедансы различных компонент:

$$z_\alpha = 2i\omega L, \quad z_\beta = -i/\omega C, \quad z_\gamma = i\omega L, \quad z_\delta = -i/2\omega C.$$

Тогда

$$Z = i \begin{pmatrix} 2\omega L & 0 & 0 & 0 \\ 0 & -(\omega C)^{-1} & 0 & 0 \\ 0 & 0 & \omega L & 0 \\ 0 & 0 & 0 & -(2\omega C)^{-1} \end{pmatrix}.$$

Перемножая матрицы, получаем

$$sZ\sigma = i \begin{pmatrix} 3\omega L - (\omega C)^{-1} & -\omega L \\ -\omega L & \omega L - (2\omega C)^{-1} \end{pmatrix}.$$

Определитель этой матрицы равен

$$\text{Det}(sZ\sigma) = -(3\omega L - (\omega C)^{-1})(\omega L - (2\omega C)^{-1}) + \omega^2 L^2,$$

или

$$\text{Det}(sZ\sigma) = -2\omega^2 L^2 + 5L/2C - 1/2\omega^2 C^2.$$

Полагая  $\text{Det}(sZ\sigma) = 0$ , получаем уравнение, квадратичное по  $\omega^2$ :

$$D(\omega^2) = 2L^2\omega^4 - 5L\omega^2/2C + (2C^2)^{-1} = 0.$$

Оно легко разлагается на множители:

$$(L\omega^2 - C^{-1})(2L\omega^2 - (2C)^{-1}) = 0.$$

Таким образом, резонансные частоты этой цепи равны

$$\omega^2 = 1/LC \quad \text{и} \quad \omega^2 = 1/4LC.$$

Чтобы найти соответствующие нормальные моды, для каждой резонансной частоты решим уравнение  $(sZ\sigma)\mathbf{J} = 0$ . Например, для  $\omega = \sqrt{1/LC}$

$$sZ\sigma = i \begin{pmatrix} 2\sqrt{L/C} & -\sqrt{L/C} \\ -\sqrt{L/C} & \frac{1}{2}\sqrt{L/C} \end{pmatrix}.$$

Решением уравнения  $(sZ\sigma)\mathbf{J} = 0$  является  $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ . Это значит, что ток в ячейке 2 в два раза больше тока в ячейке 1. Аналогично для  $\omega = \sqrt{1/4LC}$  имеем

$$sZ\sigma = i \begin{pmatrix} -\frac{1}{2}\sqrt{L/C} & -\frac{1}{2}\sqrt{L/C} \\ -\frac{1}{2}\sqrt{L/C} & -\frac{1}{2}\sqrt{L/C} \end{pmatrix},$$

так что уравнение  $(sZ\sigma)\mathbf{J} = 0$  имеет решение  $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ . Это значит, что в нормальной моде с частотой  $\omega/2\pi$  токи по ячейкам  $J_1$  и  $J_2$  имеют одинаковую амплитуду, но противоположные фазы.

## 12.8. Цепи переменного тока в стационарном режиме и фильтры

Понятие импеданса, использовавшееся в предыдущем параграфе для определения резонансных частот цепей, содержащих индукторы и конденсаторы, может быть использовано для анализа цепей переменного тока в установившемся режиме. Предположим, что все источники тока и напряжения имеют одинаковую частоту  $\omega/2\pi$ , так что токи в ветвях образуют вектор

$$\mathbf{I} = \begin{pmatrix} I_\alpha \\ I_\beta \\ \vdots \end{pmatrix} e^{i\omega t},$$

а напряжения на ветвях образуют вектор

$$\mathbf{V} = \begin{pmatrix} V^\alpha \\ V^\beta \\ \vdots \end{pmatrix} e^{i\omega t}.$$

Здесь компоненты  $I_\alpha, I_\beta, \dots$  и  $V^\alpha, V^\beta, \dots$  являются *комплексными* величинами; вещественная часть каждой компоненты представляет истинный ток или напряжение. (Вспомним, что  $e^{i\omega t} = \cos \omega t + i \sin \omega t$ .) Источники напряжения и тока представляются соответственно векторами

$$\mathbf{W} = \begin{pmatrix} W^\alpha \\ W^\beta \\ \vdots \end{pmatrix} e^{i\omega t} \quad \text{и} \quad \mathbf{K} = \begin{pmatrix} K_\alpha \\ K_\beta \\ \vdots \end{pmatrix} e^{i\omega t}.$$

Для каждой ветви по-прежнему выполняется соотношение

$$V^\alpha - W^\alpha = z_\alpha(I_\alpha - K_\alpha),$$

где  $z_\alpha$  является диагональной матрицей, элементы которой — комплексные импедансы в различных ветвях. Формализм, разработанный нами для цепей постоянного тока, переносится практически без изменения на цепи переменного тока в стационарном режиме. Необходимо сделать только несколько изменений.

1. Физические напряжения и токи в ветвях определяются *вещественной частью* соответствующей компоненты комплексных векторов  $Ie^{i\omega t}$  или  $Ve^{i\omega t}$ .
2. Амплитудные значения напряжений и токов в ветвях определяются модулем комплексной величины, которая представляет напряжение или ток, т. е. если в ветви  $\alpha$  ток представляется  $I_\alpha$ , то его амплитудное значение равно

$$|I_\alpha| = \sqrt{(\operatorname{Re} I_\alpha)^2 + (\operatorname{Im} I_\alpha)^2}.$$

3. Средняя мощность, выделенная в ветви  $\alpha$ , определяется средним значением произведения *вещественных частей* комплексных величин  $V^\alpha$  и  $I_\alpha$ , т. е. средняя мощность  $P$  выражается формулой

$$\begin{aligned} P_\alpha &= \text{Среднее от } [\operatorname{Re}(V^\alpha e^{i\omega t}) \operatorname{Re}(I_\alpha e^{i\omega t})] \\ &= \frac{1}{4} \text{Среднее от } [(V^\alpha e^{i\omega t} + V^{\alpha*} e^{-i\omega t})(I_\alpha e^{i\omega t} + I_\alpha^* e^{-i\omega t})] \\ &= \frac{1}{4} \text{Среднее от } (V^\alpha I_\alpha e^{2i\omega t} + V^\alpha I_\alpha^* + V^{\alpha*} I_\alpha + V^{\alpha*} I_\alpha^* e^{-2i\omega t}), \end{aligned}$$

где знак  $*$  обозначает комплексное сопряжение.

Поскольку члены, содержащие  $e^{2i\omega t}$  и  $e^{-2i\omega t}$ , при усреднении по времени обращаются в нуль, мы имеем

$$P_\alpha = \frac{1}{4} (V^{\alpha*} I_\alpha + V^\alpha I_\alpha^*) = \frac{1}{2} \operatorname{Re}(V^\alpha I_\alpha^*).$$

В случае, когда ветвь  $\alpha$  состоит только из резистора, т. е.  $V^\alpha = RI_\alpha$ , это выражение упрощается, и мы получаем

$$P_\alpha = \frac{1}{2} RI_\alpha^2.$$

Если ветвь состоит только из конденсатора или индуктора, то ее импеданс  $z_\alpha$  чисто мнимый, поэтому средняя выделенная мощность в этой ветви равна нулю.

В качестве простого примера цепи в стационарном режиме рассмотрим так называемый низкочастотный фильтр, изображенный на рис. 12.30. Будем считать, что в цепи есть две ветви. Базисом для  $Z_1$  является ячейка  $\alpha + \beta$ , так что  $\sigma = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ ,

$s = \begin{pmatrix} 1 & 1 \end{pmatrix}$ . Поскольку  $Z = \begin{pmatrix} R & 0 \\ 0 & -i/\omega C \end{pmatrix}$ , то сразу же получаем

$$sZ\sigma = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} R & 0 \\ 0 & -i/\omega C \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = R - i/\omega C.$$

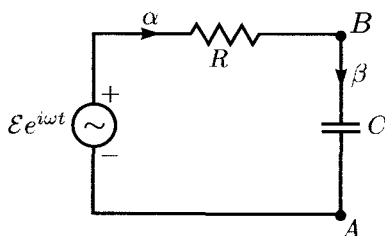


Рис. 12.30

Напряжение на ветви  $\alpha$  равно  $W^\alpha = -\mathcal{E}e^{i\omega t}$  (полярность указана на рисунке). Тогда ток в ячейке равен

$$J = (sZ\sigma)^{-1}(-sW) = \frac{1}{R - i/\omega C} \mathcal{E}e^{i\omega t}.$$

Амплитудное значение этого тока равно модулю этого выражения, т. е.

$$|J| = \left| \frac{\mathcal{E}}{R - i/\omega C} \right| = \frac{\mathcal{E}}{\sqrt{R^2 + (\omega^2 C^2)^{-1}}},$$

а амплитудное значение напряжения в ветви  $\beta$  равно

$$|V^\beta| = |z_\beta I_\beta| = \left| \frac{-i\mathcal{E}/\omega C}{R - i/\omega C} \right| = \frac{\mathcal{E}}{\sqrt{(\omega RC)^2 + 1}}.$$

Это значение уменьшается от  $\mathcal{E}$  для  $\omega = 0$  до нуля при  $\omega \rightarrow \infty$ . Отсюда и происходит название *низкочастотный* фильтр<sup>18</sup>.

Эту цепь можно слегка усложнить, присоединив последовательно к конденсатору индуктор. В этом случае импеданс ветви  $\beta$  равняется  $z_L = i(\omega L - 1/\omega C)$ , а ток в ячейке равен<sup>19</sup>

$$J = \mathcal{E} / (R + i(\omega L - (\omega C)^{-1})).$$

Легко видеть, что этот ток достигает максимальной амплитуды при  $L = 1/\omega^2 C$  или  $\omega = \sqrt{1/LC}$ . Эта частота называется *резонансной* частотой цепи, изображенной на рис. 12.31.

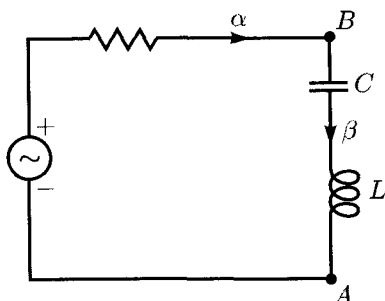


Рис. 12.31

На рис. 12.32 изображен более сложный пример цепи. Теперь базис  $Z$  состоит из двух ячеек  $\mathbf{M}_1 = \alpha + \beta$  и  $\mathbf{M}_2 = -\beta + \gamma$ , поэтому

$$\sigma = \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \quad \text{и} \quad s = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Поскольку

$$Z = \begin{pmatrix} R & 0 & 0 \\ 0 & i\omega L & 0 \\ 0 & 0 & -i/\omega C \end{pmatrix},$$

<sup>18</sup>Описанное поведение снимаемого с ветви  $\beta$  напряжения — убывание с ростом частоты — дает возможность отделить низкую частоту сигнала (например, звука) от модулированной им высокой частоты электромагнитных волн. Данный принцип лежит в основе работы всей радиосвязи. — Прим. ред.

<sup>19</sup>Здесь и далее для упрощения записи опускается фазовый множитель  $e^{i\omega t}$ . — Прим. ред.

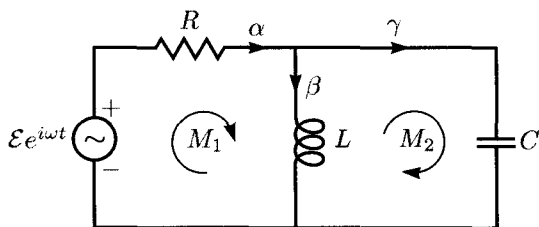


Рис. 12.32

мы имеем

$$sZ\sigma = \begin{pmatrix} R & i\omega L & 0 \\ 0 & -i\omega L & -i/\omega C \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} R + i\omega L & -i\omega L \\ -i\omega L & i\omega L - i/\omega C \end{pmatrix}.$$

Нетрудно получить обратную матрицу

$$(sZ\sigma)^{-1} = \frac{1}{R(i\omega L - i/\omega C) + LC^{-1}} \begin{pmatrix} i\omega L - i/\omega C & i\omega L \\ i\omega L & R + i\omega L \end{pmatrix}.$$

Чтобы определить токи в ячейках, запишем вектор

$$\mathbf{W} = \begin{pmatrix} -\mathcal{E} \\ 0 \\ 0 \end{pmatrix}$$

и вычислим

$$J = (sZ\sigma)^{-1}(-sW) = \frac{\mathcal{E}}{R(i\omega L - i/\omega C) + LC^{-1}} \begin{pmatrix} i\omega L - i/\omega C \\ i\omega L \end{pmatrix}.$$

Интересным является решение для резонансной частоты  $\omega/2\pi$ , где  $\omega L = 1/\omega C$  и  $J = \frac{\mathcal{E}}{L/C} \begin{pmatrix} 0 \\ i\omega L \end{pmatrix}$ . В этом случае в ячейке  $M_1$  нет тока, и в резисторе не выделяется мощность.

Наша техника позволяет анализировать цепи переменного тока произвольной сложности так же легко, как и цепи постоянного тока той же топологии. В качестве примера рассмотрим двухступенчатый фильтр (рис. 12.33).

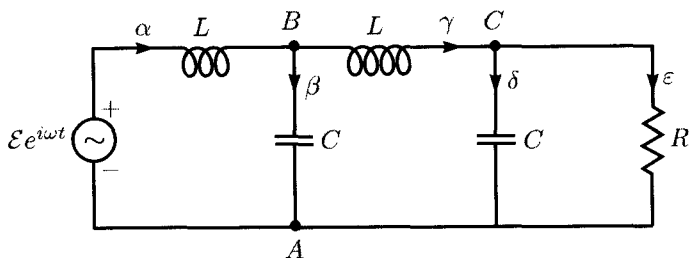


Рис. 12.33

Поскольку эта цепь имеет три ячейки и только два незаземленных узла, то проще воспользоваться методом узловых потенциалов. Пусть узел  $A$  заземлен. Тогда

$$[\partial] = \begin{pmatrix} 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 \end{pmatrix} \quad \text{и} \quad [d] = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ -1 & 1 \\ 0 & -1 \\ 0 & -1 \end{pmatrix}.$$

В этом случае

$$\begin{aligned} [\partial]Z^{-1}[d] &= \begin{pmatrix} 1 & -1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} -i/\omega L & 0 \\ i\omega C & 0 \\ i/\omega L & -i/\omega L \\ 0 & i\omega C \\ 0 & -1/R \end{pmatrix} \\ &= \begin{pmatrix} i\omega C - 2i/\omega L & i/\omega L \\ i/\omega L & i\omega C - (i/\omega L) + 1/R \end{pmatrix}. \end{aligned}$$

Чтобы обратить эту матрицу, воспользуемся формулой

$$\Phi = ([\partial]Z^{-1}[d])^{-1}[\partial]Z^{-1}\mathbf{W},$$

и определим потенциалы узлов  $B$  и  $C$ . Ограничимся решением задачи для резонансной частоты, когда  $\omega C = 1/\omega L$ . В этом случае обращаемая матрица имеет вид

$$[\partial]Z^{-1}[d] = \begin{pmatrix} -i\omega C & i\omega C \\ i\omega C & 1/R \end{pmatrix}.$$

Заметим, что для этой частоты

$$[\partial]Z^{-1}\mathbf{W} = \begin{pmatrix} -1/i\omega L \\ 0 \end{pmatrix} = \begin{pmatrix} i\mathcal{E}\omega C \\ 0 \end{pmatrix}.$$

Тогда

$$\begin{pmatrix} \phi^B \\ \phi^C \end{pmatrix} = \frac{1}{(-i\omega C/R) - \omega^2 C^2} \begin{pmatrix} 1/R & -i\omega C \\ -i\omega C & -i\omega C \end{pmatrix} \begin{pmatrix} i\omega C \\ 0 \end{pmatrix} \mathcal{E},$$

или

$$\begin{pmatrix} \phi^B \\ \phi^C \end{pmatrix} = \frac{\mathcal{E}}{(-i\omega C/R) - \omega^2 C^2} \begin{pmatrix} i\omega C/R \\ \omega^2 C^2 \end{pmatrix}.$$

## Резюме

### А. 1-комплексы

Для заданной схемы, представляющей 1-комплекс, Вы должны знать, как построить максимальное дерево и связанные с ним базисные ячейки пространства  $Z_1$  и как написать матрицы для  $\sigma$ ,  $s$ ,  $\partial$  и  $d$ .

Вы должны знать определения подпространств  $B_0$ ,  $H_0$ ,  $Z_1$ ,  $H^0$ ,  $P^0$  и  $B^1$  и уметь строить для них базисы.

### В. Цепи с резисторами

Вы должны уметь писать соотношение между  $V$  и  $I$  для ветви, содержащей батарею, источник тока и резистор.

Для расчета электрических цепей Вы должны знать и уметь пользоваться методами Максвелла: методом круговых токов и методом узловых потенциалов.

### С. Цепи переменного тока

Вы должны уметь пользоваться методами Максвелла для нахождения стационарных решений и нормальных мод сетей, содержащих индукторы, конденсаторы, синусоидальные источники тока и напряжений.

## Задачи

12.1, 2. Для цепей, изображенных на рис. 12.34 и 12.35, сделайте следующее:

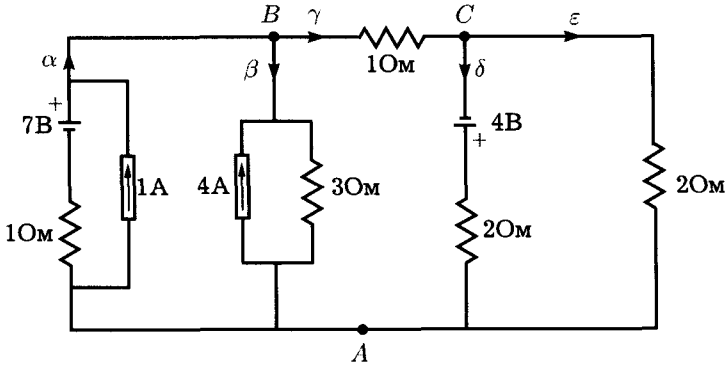


Рис. 12.34

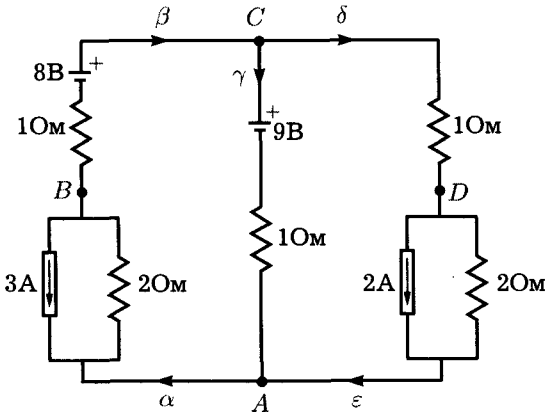


Рис. 12.35

- (а) Найдите множество независимых ячеек и напишите элемент  $C_1$ , соответствующий каждой ячейке. (Можно воспользоваться обозначениями типа  $M_1 = \alpha + \beta - \gamma$ .) Проверьте, что число независимых ячеек  $m$  удовлетворяет условию  $m = b + 1 - n$ .

- (b) Выразите ток в каждой ветви через токи в ячейках и найдите напряжения на ветвях.
- (c) Применяя второй закон Кирхгофа к каждой ячейке, постройте систему уравнений для токов в ячейках. Решите эту систему.
- (d) Пусть узел  $A$  заземлен. Выразите напряжения и токи через остающиеся неизвестные узловые потенциалы. Применяя первый закон Кирхгофа к каждому узлу, получите систему уравнений для узловых потенциалов и решите ее.

12.3, 4. Для комплексов, изображенных на рис. 12.36 и 12.37, сделайте следующее:

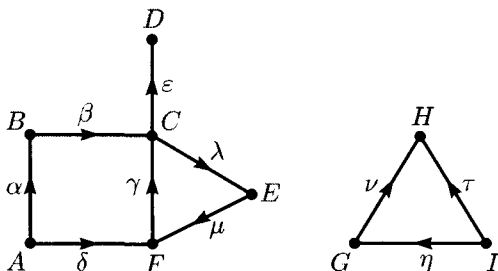


Рис. 12.36

- (a) Определите размерность пространств  $C_1$ ,  $Z_1$ ,  $C_0$  и  $H_0$ . Проверьте, что  $\dim H_0 - \dim Z_1 = \dim C_0 - \dim C_1$ .
- (b) Для каждой связной компоненты найдите дерево  $T$ , которое соединяет все узлы. После этого постройте базис пространства  $Z_1$  и запишите явно проекцию  $\rho_T$  в виде

$$\begin{aligned}\rho_T(\alpha) &= 0, \\ \rho_T(\beta) &= \beta + \gamma - \delta, \quad \text{и т. д.}\end{aligned}$$

- (c) Определите, сколько необходимо иметь уравнений для токов в ячейках и для узловых потенциалов, чтобы рассчитать цепь с данной топологией.

12.5. Для каждой из трех ветвей, изображенных на рис. 12.38, найдите соотношение между  $V$  и  $I$ . В каждом случае постройте ветвь с тем же соотношением, содержащую только подходящие батарею и резистор (но не источник тока), а потом — источник тока и резистор (но без батареи).

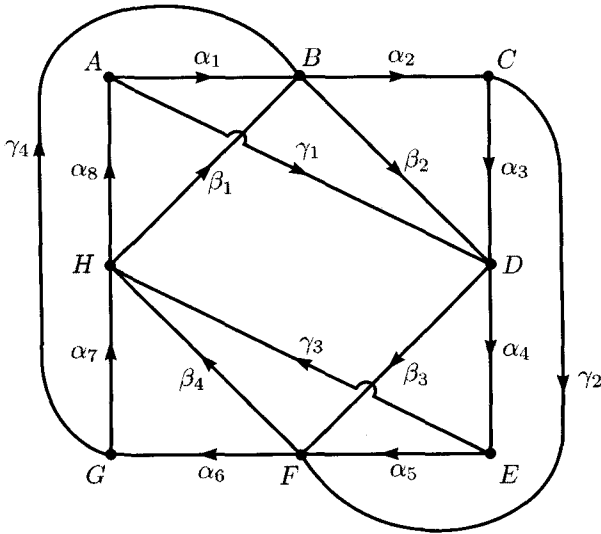


Рис. 12.37

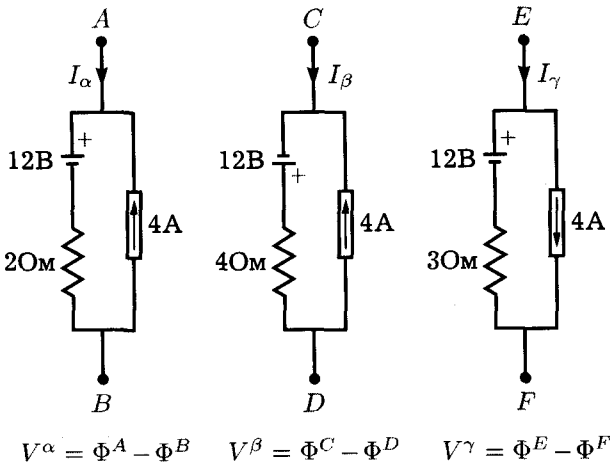


Рис. 12.38

12.6. Рассмотрите комплекс, изображенный на рис. 12.39.

- (а) Напишите матрицы  $\partial$  и  $d$ , соответствующие порядку ветвей  $\alpha, \beta, \gamma, \delta, \varepsilon, \phi$ .

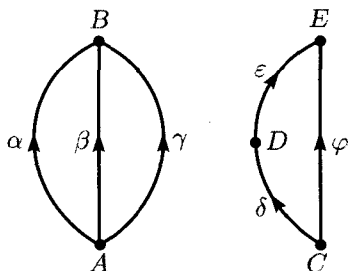


Рис. 12.39

- (b) Покажите, что вектора  $\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$  и  $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$  образуют базис про-

странства  $H^0$ . Заземлите узлы  $A$  и  $C$  и покажите, что век-

тора  $\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ ,  $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$  и  $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$  образуют базис пространства  $P^0 =$

$C^0/Z^0$ . В этом базисе постройте матрицу оператора  $[d] : P^0 \rightarrow C^1$ .

- (c) Для каждой компоненты выберите максимальное дерево с ветвями  $\alpha$ ,  $\delta$  и  $\epsilon$ . Три базисными элементами пространства  $Z_1$  будут ячейки, образованные остальными ветвями  $\beta$ ,  $\gamma$  и  $\phi$  с ветвями дерева (ячейку выбирайте так, чтобы  $\beta$ ,  $\gamma$  или  $\delta$  имели коэффициент  $+1$ ). В этом базисе напишите матрицу  $\sigma : Z_1 \rightarrow C_1$ .
- (d) Покажите, что столбцы матрицы  $[d]$  определяют базис  $B^1$ . Свяжите вектора факторпространства  $H^1 = C^1/B^1$  с тремя ветвями, не принадлежащими максимальным деревьям, получив

$$\overline{\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}}, \quad \overline{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}} \quad \text{и} \quad \overline{\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}}.$$

- Покажите, что эти три вектора образуют базис пространства  $H^1$ , дуальный к базису, выбранному для  $Z_1$ .
- (е) В этом базисе постройте матрицу  $s : C^1 \rightarrow H^1$ . Убедитесь, что матрица  $s$  есть транспонированная матрица к  $\sigma$ .
- (ф) Покажите, что границы траекторий, соединяющих заземленные узлы с незаземленными, образуют базис пространства  $B_0$ , дуальный к базису пространства  $P^0$ .
- (г) Покажите, что классы эквивалентности заземленных узлов

$$\overline{\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}} \quad \text{и} \quad \overline{\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}}$$

образуют базис пространства  $H_0 = C^0/B^0$ , дуальный к базису пространства  $Z^0$ .

- 12.7. Для цепи, изображенной на рис. 12.34 (задача 12.1), постройте  $[d]$ ,  $[\partial]$ ,  $\sigma$ ,  $s$ ,  $Z$ ,  $\mathbf{K}$  и  $\mathbf{W}$ . Составьте уравнения для токов в ячейках и уравнения для узловых потенциалов, используя эти матрицы и вектора. Проверьте, что получатся в итоге те же уравнения. Примите, что узел  $A$  заземлен, а ветви  $\beta$  и  $\delta$  образуют максимальное дерево.
- 12.8. Сделайте то же самое для цепи, изображенной на рис. 12.35 (задача 12.2), считая, что узел  $A$  заземлен, а ветви  $\alpha$ ,  $\beta$  и  $\delta$  образуют максимальное дерево.
- 12.9. Воспользуйтесь методом круговых токов и найдите токи в ветвях  $I_\alpha$ ,  $I_\beta$ ,  $I_\gamma$  для цепи, изображенной на рис. 12.40. Используйте ячейки, определенные выбором  $\beta$  в качестве максимального дерева.
- 12.10. Рассмотрите цепь на рис. 12.41. Пусть матрица преобразования  $s$  имеет вид

$$s = \begin{pmatrix} -1 & +1 & 0 & +1 \\ 0 & 0 & +1 & -1 \end{pmatrix}.$$

Найдите в явном виде базис пространства  $Z_1$ , который определяется этой матрицей  $s$ . Какое максимальное дерево приведет к этому базису? Пусть базис в пространстве  $P^0$  выбран так:

$$\mathbf{p}^1 = \overline{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}, \quad \mathbf{p}^2 = \overline{\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}}.$$

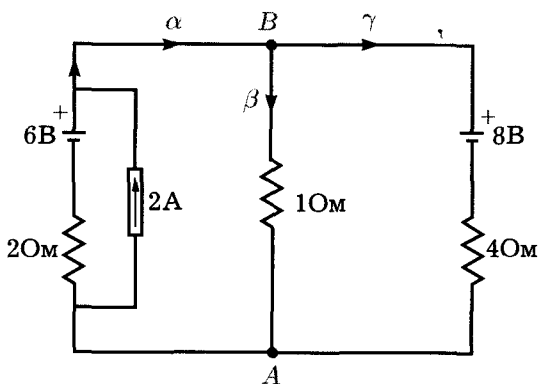


Рис. 12.40

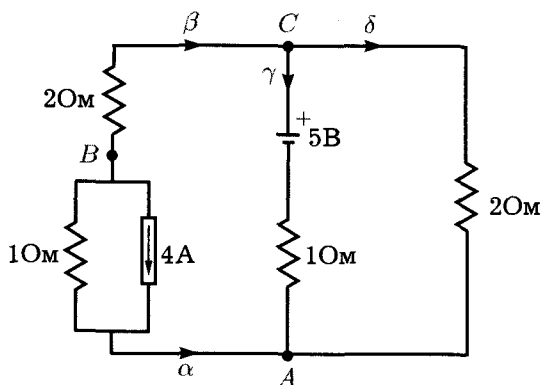


Рис. 12.41

Выразите вектор  $\begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$  через базисные вектора  $\mathbf{p}^1$  и  $\mathbf{p}^2$ . Напишите матрицы, представляющие  $[\partial]$  и  $[d]$ . Для данной цепи воспользуйтесь методом узловых потенциалов и вычислите  $\Phi^B$  и  $\Phi^C$ .

- 12.11. Рассмотрите цепь, изображенную на рис. 12.42. Для этой цепи напишите матрицы  $\partial$  и  $d$ . Пусть ветви  $\alpha$  и  $\gamma$  образуют максимальное дерево. Напишите базис пространства  $Z_1$ , соответствующий этому максимальному дереву, и в этом базисе постройте матрицу  $\sigma$ . Воспользуйтесь методом круговых токов и получите токи в ветвях этой цепи.

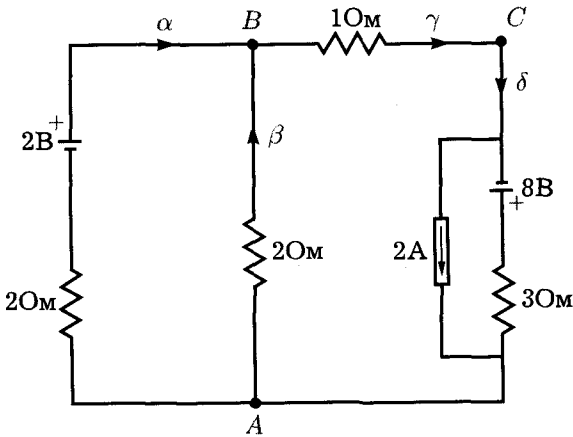


Рис. 12.42

- 12.12. (а) Используя метод круговых токов, определите частоты нормальных мод для цепи, изображенной на рис. 12.43, т.е. определите значения  $\omega/2\pi$ , для которых матрица  $sZ\sigma$  вырождена. Для этих частот найдите решения уравнения

$$(sZ\sigma)\mathbf{J} = 0.$$

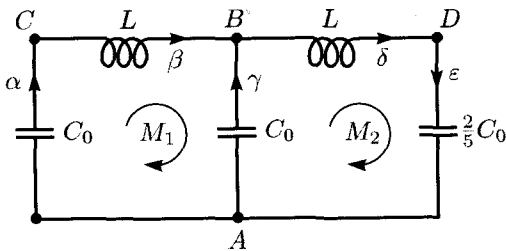


Рис. 12.43

- (б) Рассчитайте эту цепь методом узловых потенциалов. Если узел  $A$  заземлен, то все равно придется обращаться матрицу  $3 \times 3$ . Почему существуют только две нормальные моды?
- (с) Соединив последовательно ветви  $\alpha$  с  $\beta$  и  $\delta$  с  $\epsilon$  и убрав узлы  $C$  и  $D$ , преобразуйте эту цепь в цепь, содержащую только

два узла. Рассчитайте цепь методом узловых потенциалов. На этот раз остается работать с матрицей  $1 \times 1$ !

- 12.13. (a) Используя метод круговых токов, определите частоты нормальных мод для цепи, изображенной на рис. 12.44. Для каждой моды выразите частоту  $\omega/2\pi$  через величину  $\omega_0 = 1/\sqrt{L_0 C_0}$  и получите отношение токов в ячейках  $J_2/J_1$ .

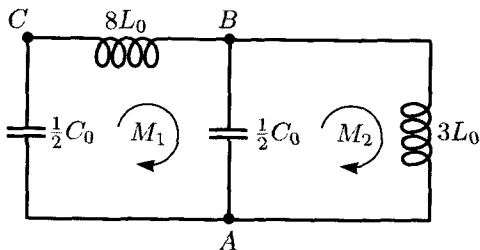


Рис. 12.44

- (b) Для той же цепи методом узловых потенциалов определите частоты нормальных мод.
- 12.14. (a) Для цепи, изображенной на рис. 12.45, постройте комплексную матрицу  $sZ\sigma$ .

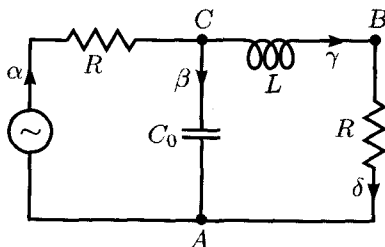


Рис. 12.45

- (b) Предположим, что генератор в ветви  $\alpha$  подает напряжение, равное вещественной части  $\mathcal{E}_0 e^{i\omega t}$ . Получите выражения для стационарных напряжений  $V^\beta$  и  $V^\delta$ .
- 12.15. Вернемся к упражнению 12.13 и предположим, что последовательно с конденсатором между узлами  $A$  и  $B$  соединен источник, создающий ток  $K = \text{Re}(K_0 e^{i\omega_0 t})$ , где  $\omega_0 = 1/\sqrt{L_0 C_0}$ . Получите выражения для токов в ячейках  $J_1$  и  $J_2$ .

## Глава 13

---

# Метод ортогональных проекций

---

В этой главе мы продолжаем изучать электрические цепи. Исследуем краевые задачи, возникающие в цепях с конденсаторами, и используем изложенные методы для решения ряда классических задач в электростатике при наличии проводников.

### 13.1. Метод ортогональных проекций Вейля

Рассмотрим еще два метода, используемых для расчета электрических цепей: метод ортогональных проекций Вейля и оригинальный метод Кирхгофа. Оба метода интересны и их стоит изучить, хоть они и ограничены только цепями с сопротивлениями, т. е. случаем, когда все импедансы  $z_\alpha$  — положительные числа,  $z_\alpha = r_\alpha > 0$ .

Начнем с метода Вейля. Уравнение, связывающее векторы напряжений на ветвях  $\mathbf{V}$  и токов в ветвях  $\mathbf{I}$ , имеет вид

$$\mathbf{V} - \mathbf{W} = Z(\mathbf{I} - \mathbf{K}).$$

Перепишем его в виде

$$\mathbf{V} = \mathbf{W} + Z(\mathbf{I} - \mathbf{K})$$

или

$$\mathbf{V} = Z(Z^{-1}\mathbf{W} + \mathbf{I} - \mathbf{K}).$$

Матрицу  $Z$  используем для определения положительно определенного скалярного произведения в пространстве  $C_1$  токов,

протекающих по ветвям:

$$(\mathbf{I}, \mathbf{I}')_Z = \int_{\mathbf{I}} Z \mathbf{I}' = r_\alpha I_\alpha I'_\alpha + r_\beta I_\beta I'_\beta + \dots$$

Предположим, что у нас есть произвольное распределение токов по ветвям  $\mathbf{I}'$ , удовлетворяющее первому закону Кирхгофа  $\partial \mathbf{I}' = 0$ , и произвольное распределение напряжений  $\mathbf{V}$ , удовлетворяющее второму закону Кирхгофа  $\mathbf{V} = -d\Phi$ . Нам известно, что  $\int_{\mathbf{I}'} \mathbf{V} = 0$ . Поскольку  $\mathbf{V} = Z(Z^{-1}\mathbf{W} + \mathbf{I} - \mathbf{K})$ , можно написать

$$\int_{\mathbf{I}'} Z(Z^{-1}\mathbf{W} + \mathbf{I} - \mathbf{K}) = 0.$$

Если это соотношение выразить через скалярное произведение, введенное выше, то его можно переписать в виде

$$(\mathbf{I}', Z^{-1}\mathbf{W} + \mathbf{I} - \mathbf{K})_Z = 0.$$

Это значит, что вектор  $(\mathbf{K} - Z^{-1}\mathbf{W}) - \mathbf{I}$  ортогонален всем циклам.

Теперь мы можем задачу о цепи с сопротивлениями сформулировать по-другому. Нам дано пространство  $C_1$ , образованное токами в ветвях. Внутри него лежит подпространство циклов  $Z_1$ . Для сети с заданными векторами источников  $\mathbf{W}$  и  $\mathbf{K}$  можно построить вектор  $\mathbf{K} - Z^{-1}\mathbf{W}$ , который описывает токи в ветвях, существующие при условии, что все узловые потенциалы равны нулю. В общем случае этот вектор не является элементом подпространства  $Z_1$ . Теперь мы хотели бы получить такой вектор  $\mathbf{I} \in Z_1$ , чтобы вектор  $(\mathbf{K} - Z^{-1}\mathbf{W}) - \mathbf{I}$  был ортогонален всем циклам. Итак, мы свели расчет цепи с резисторами к геометрической задаче, изображенной на рис. 13.1. Для данного вектора  $\mathbf{K} - Z^{-1}\mathbf{W}$  в пространстве  $C_1$  надо вычислить его ортогональную проекцию на подпространство  $Z_1$  относительно скалярного произведения, определенного матрицей  $Z$ .

Тогда вектор  $\mathbf{K} - Z^{-1}\mathbf{W}$  записывается в форме

$$\mathbf{K} - Z^{-1}\mathbf{W} = \mathbf{I} + (\mathbf{K} - Z^{-1}\mathbf{W} - \mathbf{I}),$$

где  $\mathbf{I}$  принадлежит пространству  $Z_1$ , а  $\mathbf{K} - Z^{-1}\mathbf{W} - \mathbf{I} = -Z^{-1}\mathbf{V}$  ортогонален  $Z_1$ . Пусть  $\pi$  обозначает линейное преобразование,

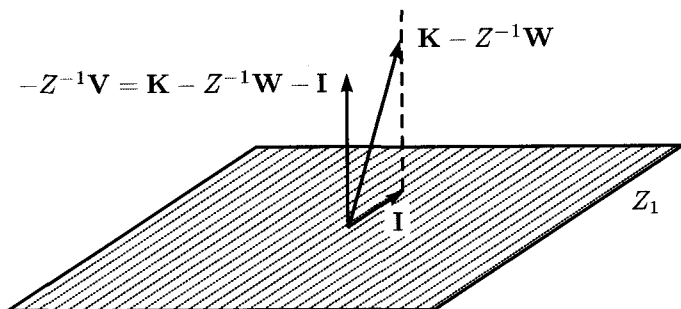


Рис. 13.1

которое проектирует  $C_1$  ортогонально на  $Z_1$ . В этом случае решением задачи о цепи является

$$\begin{aligned} \mathbf{I} &= \pi(\mathbf{K} - \mathbf{Z}^{-1}\mathbf{W}), \\ \mathbf{V} &= \mathbf{Z}(\pi - 1)(\mathbf{K} - \mathbf{Z}^{-1}\mathbf{W}). \end{aligned}$$

Геометрически очевидно, что существует единственный проекционный оператор  $\pi$ , а следовательно, существует единственное решение задачи для данной цепи.

Одним из способов получения ортогональной проекции из  $C_1$  в  $Z_1$  основан на выборе ортонормированного базиса. Пусть  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\}$  — ортонормированный базис пространства  $Z_1$ . Тогда если  $\mathbf{u}$  — элемент пространства  $C_1$ , то его проекция на подпространство  $Z_1$  дается формулой

$$\pi \mathbf{u} = (\mathbf{e}_1, \mathbf{u})\mathbf{e}_1 + (\mathbf{e}_2, \mathbf{u})\mathbf{e}_2 + \dots + (\mathbf{e}_m, \mathbf{u})\mathbf{e}_m.$$

Процедура вычисления матрицы  $\pi$  довольно утомительна. Сначала мы должны выбрать базис подпространства  $Z_1$ , например, семейство независимых ячеек. После этого применяем процедуру Грама–Шмидта, чтобы этот базис преобразовать в ортонормированный относительно скалярного произведения, определенного матрицей  $\mathbf{Z}$ . Наконец, мы должны брать по очереди каждый базисный вектор пространства  $C_1$  и вычислять его проекцию на пространство  $Z_1$ , беря сумму проекций на векторы ортонормированного базиса  $Z_1$ . Каждая такая проекция определяет столбец матрицы  $\pi$ .

В качестве иллюстрации метода Вейля рассмотрим цепь, изображенную на рис. 13.2. Для этой цепи пространство  $C_1$  двумерно. В ней есть только одна ячейка  $\alpha + \beta$ , поэтому подпространство  $Z_1$  одномерно и натянуто на вектор  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ . Матрица  $Z$  имеет вид  $\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ .

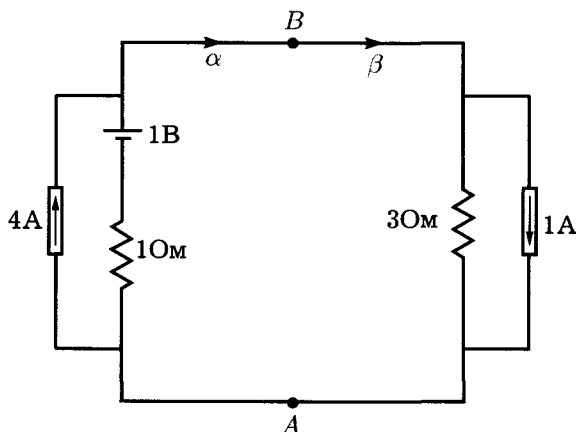


Рис. 13.2

Чтобы вычислить проекционную матрицу  $\pi$ , мы должны по очереди спроектировать два базисных вектора  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  и  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  на подпространство  $Z_1$ . Сначала отнормируем базисный элемент  $Z_1$ : длина вектора  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  равна  $\left( \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) = 4$ , поэтому мы делим вектор  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  на  $\sqrt{4} = 2$ , в результате получая вектор  $\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$ , длина которого равна 1. Проекция вектора  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  равна

$$\left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \right) \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 1/4 \\ 1/4 \end{pmatrix},$$

а проекция вектора  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  равна

$$\left( \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \right) \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} = \begin{pmatrix} 3/4 \\ 3/4 \end{pmatrix}.$$

Отсюда матрица  $\pi = \begin{pmatrix} 1/4 & 3/4 \\ 1/4 & 3/4 \end{pmatrix}$ . Для цепи, изображенной на рис. 13.2, получаются вектора  $\mathbf{K} = \begin{pmatrix} 4 \\ 1 \end{pmatrix}$  и  $\mathbf{W} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ . Поэтому  $\mathbf{K} - Z^{-1}\mathbf{W} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} - \begin{pmatrix} -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \end{pmatrix}$ . Применяя матрицу  $\pi$  к этому вектору, получаем

$$\mathbf{I} = \pi(\mathbf{K} - Z^{-1}\mathbf{W}) = \begin{pmatrix} 1/4 & 3/4 \\ 1/4 & 3/4 \end{pmatrix} \begin{pmatrix} 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Тогда решение имеет вид  $I_\alpha = I_\beta = 2$ . Заметим, что вектор  $\mathbf{K} - Z^{-1}\mathbf{W} - \mathbf{I} = \begin{pmatrix} 5 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}$  ортогонален  $\mathbf{I}$ , как и должно быть. Расчет этой цепи графически изображен на рис. 13.3. Поскольку введенное скалярное произведение не является обычным евклидовым, то подпространство  $Z_1$  и его ортогональное дополнение представляются не перпендикулярными прямыми.

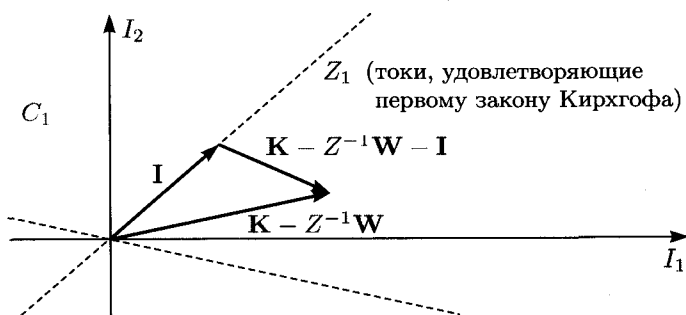


Рис. 13.3

Понятно, как найти явное выражение проекционного оператора Вейля  $\pi$ , используя решение методом круговых токов. Вспомним, что в этом методе  $\mathbf{I} = \sigma \mathbf{J}$ , где  $\mathbf{J} = (sZ\sigma)^{-1}s(Z\mathbf{K} - \mathbf{W})$ .

Тогда  $\mathbf{I} = \sigma(sZ\sigma)^{-1}sZ(\mathbf{K} - Z^{-1}\mathbf{W})$ . Сравнивая этот результат с формулой Вейля  $\mathbf{I} = \pi(\mathbf{K} - Z^{-1}\mathbf{W})$ , получаем, что

$$\pi = \sigma(sZ\sigma)^{-1}sZ.$$

Легко проверить, что  $\pi^2 = \pi$ , как и должно быть для любого проекционного оператора<sup>1</sup>.

## 13.2. Метод Кирхгофа

Рассмотрим метод, разработанный самим Кирхгофом. И хотя он был придуман за много десятилетий до метода Вейля, мы обращаемся к нему сейчас, потому что он дает явную формулу для проекционного оператора  $\pi$  через деревья в цепях.

Предположим, что у нас есть связный комплекс и пусть  $T$  обозначает максимальное дерево. Вспомним, что для каждого такого дерева мы определили проекционный оператор  $\rho_T$ :

$$\rho_T(\alpha) = \begin{cases} 0, & \text{если } \alpha \in T, \\ \mathbf{M}_\alpha, & \text{если } \alpha \notin T, \end{cases}$$

где  $\mathbf{M}_\alpha$  есть цикл, соответствующий единственной ячейке, содержащей ветвь  $\alpha$ , а все остальные ветви принадлежат  $T$ . (Цикл следует выбирать так, чтобы он содержал  $+\alpha$ , а не  $-\alpha$ .) Например, на рис. 13.4 одна ветвь  $\beta$  образует максимальное дерево. Тогда

$$\rho_T(\alpha) = \alpha - \beta, \quad \rho_T(\beta) = 0, \quad \rho_T(\gamma) = \gamma + \beta.$$

Поэтому матрица, представляющая эту проекцию  $\rho_T$ , имеет вид  $\begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ . Заметим, что все диагональные элементы  $\rho_T$  равны либо  $+1$ , либо  $0$ .

Любой такой оператор  $\rho_T$  является проекционным оператором на пространство  $Z_1$ : его образ есть  $Z_1$  и  $\rho_T(\mathbf{M}) = \mathbf{M}$  для

<sup>1</sup>Математическое название этого свойства проекции — *идемпотентность*. Каждый идемпотентный оператор  $\pi$  в конечномерном векторном пространстве может быть геометрически охарактеризован как проекция на образ  $\text{Im } \pi$  вдоль ядра  $\text{Ker } \pi$ . — *Прим. ред.*

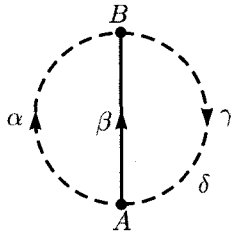


Рис. 13.4

$M \in Z_1$ . Однако он не является *ортогональным* проекционным оператором, потому что его ядро не ортогонально подпространству  $Z_1$  (по отношению к скалярному произведению, определенному матрицей  $Z$ ). Другими словами,  $\rho_T$  — не самосопряженный оператор, потому что в общем случае для двух ветвей  $\alpha$  и  $\beta$

$$(\rho_T \alpha, \beta)_Z \neq (\alpha, \rho_T \beta)_Z.$$

Итак,  $\rho_T$  для отдельно взятого максимального дерева  $T$  — не самосопряженный оператор, но в комплексе есть много различных максимальных деревьев. Открытие Кирхгофа состояло в нахождении процедуры получения самосопряженного оператора: различные проекционные операторы усредняются с определенным весом. Предположим, что для каждого максимального дерева  $T$  мы имеем вещественное число  $\lambda_T$ , причем  $0 \leq \lambda_T \leq 1$  и  $\sum_T \lambda_T = 1$ , где сумма берется по всем деревьям. Тогда оператор

$$\rho_\lambda = \sum_T \lambda_T \rho_T$$

тоже является проекцией на пространство  $Z_1$ : его образ есть  $Z_1$ , и он отображает любой вектор в пространстве  $Z_1$  в самого себя. Вообще говоря,  $\rho_\lambda$  — не ортогональная проекция, но существует такой набор  $\lambda_T$ , что ортогональность имеет место. Рецепт Кирхгофа состоит в том, что для каждого максимального дерева  $T$  определяем

$$Q_T = \prod_{\beta \notin T} r_\beta,$$

т. е. произведение сопротивлений тех ветвей, которые *не* принадлежат  $T$ . Весовые множители  $\lambda_T$  выбираются пропорциональными  $Q_T$  так, чтобы их сумма равнялась единице, т. е.

$$\lambda_T = Q_T / R, \quad \text{где} \quad R = \sum_T Q_T.$$

Таким образом, мы получили явный вид проекционного оператора Вейля  $\pi$ , а именно

$$\pi = R^{-1} \sum_T Q_T \rho_T.$$

Теперь мы должны показать, что этот оператор самосопряженный; другими словами, что самосопряжен оператор  $R\pi = \sum_T Q_T \rho_T$ . Поскольку ветви образуют базис в пространстве  $C_1$ , то достаточно показать, что для любой пары ветвей  $\alpha$  и  $\beta$  мы имеем

$$\sum_T Q_T (\rho_T \alpha, \beta)_Z = \sum_T Q_T (\alpha, \rho_T \beta)_Z.$$

Для фиксированных  $\alpha$  и  $\beta$  мы суммируем по всем максимальным деревьям. Имеются три случая, изображенных на рис. 13.5.

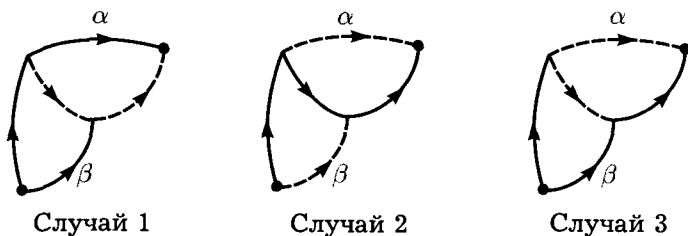


Рис. 13.5

**Случай 1.** Ветви  $\alpha$  и  $\beta$  принадлежат дереву. В этом случае  $\rho_T(\alpha) = 0$  и  $\rho_T(\beta) = 0$ . Поэтому от дерева нет никакого вклада в обе части уравнения.

**Случай 2.** Ветви  $\alpha$  и  $\beta$  не принадлежат дереву. В этом случае  $\rho_T(\alpha)$  является ячейкой, которая не содержит  $\beta$ , поэтому  $(\rho_T(\alpha), \beta)_Z = 0$ . Аналогично  $(\alpha, \rho_T(\beta))_Z = 0$ , поэтому опять от дерева нет никакого вклада.

**Случай 3.** Ветвь  $\alpha$  не принадлежит дереву, а  $\beta$  принадлежит. В этом случае если  $\beta$  входит в ячейку  $M_\alpha$ , образованную  $\alpha$  и другими ветвями дерева (иначе ситуация такая же, как в случае 2), то мы имеем

$$(\rho_T \alpha, \beta)_Z = \pm r_\beta,$$

где берется знак  $+$ , если  $\alpha$  и  $\beta$  входят в  $M_\alpha$  с одинаковыми знаками, и знак  $-$ , если эти ветви входят в ячейку в противоположными знаками. Например на рис. 13.6(а) мы имеем  $\rho_T(\alpha) = \alpha - \gamma - \beta + \delta$  и  $(\rho_T(\alpha), \beta)_Z = -r_\beta$ . Всегда существует единственное максимальное дерево  $T'$ , образованное заменой ветви  $\beta$  на  $\alpha$  в дереве  $T$ , для которого  $\rho_{T'}(\beta) = \pm \rho_T(\alpha)$ , так что  $(\alpha, \rho_{T'}(\beta))_Z = \pm r_\alpha$ . Такое дерево  $T'$  показано на рис. 13.6(б).

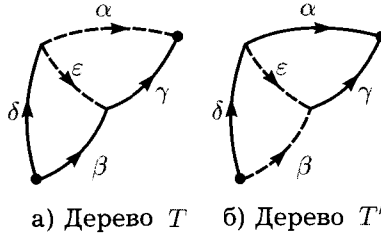


Рис. 13.6

Если из дерева  $T$  убрать ветвь  $\beta$ , то останутся те же ветви, что и при удалении из  $T'$  ветви  $\alpha$ . Значит, для этих двух деревьев справедливо равенство  $Q_T r_\beta = Q_{T'} r_\alpha$ , и оба выражения  $Q_T r_\beta$  и  $Q_{T'} r_\alpha$  являются произведениями сопротивлений остающихся ветвей. Соотношения между  $T$  и  $T'$  симметричны относительно перестановки ветвей  $\alpha$  и  $\beta$ . Тогда

$$Q_T(\rho_T \alpha, \beta)_Z = Q_{T'}(\alpha, \rho_{T'} \beta)_Z.$$

Каждому дереву  $T$ , для которого левая часть равенства не равна нулю, соответствует дерево  $T'$ , удовлетворяющее этому равенству, и наоборот. Таким образом, если просуммировать по всем  $T$ , то получим желаемый результат

$$\sum_T Q_T(\rho_T \alpha, \beta)_Z = \sum_T Q_T(\alpha, \rho_T \beta)_Z,$$

где добавлены некоторые нулевые слагаемые. Итак, доказана формула Кирхгофа

$$\pi = R^{-1} \sum_T Q_T \rho_T,$$

где

$$Q_T = \prod_{\alpha \notin T} r_\alpha, \quad R = \sum_T Q_T,$$

и суммирование ведется по всем максимальным деревьям.

В качестве иллюстрации метода Кирхгофа рассмотрим цепь, изображенную на рис. 13.7 и имеющую ту же топологию, что и цепь на рис. 13.4. В цепи есть три максимальных дерева: ветвь  $\alpha$  (дерево  $T_1$ ), ветвь  $\beta$  (дерево  $T_2$ ) и ветвь  $\gamma$  (дерево  $T_3$ ).

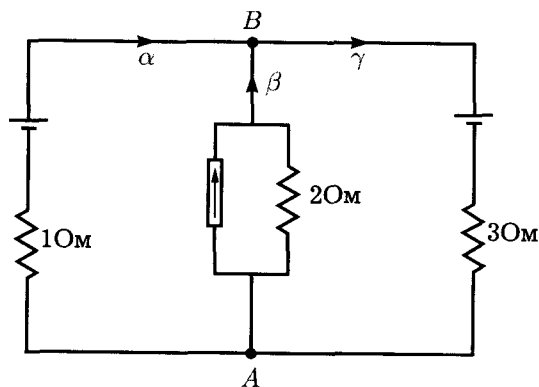


Рис. 13.7

Для дерева  $T_1$

$$\rho_{T_1} = \begin{pmatrix} 0 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{и} \quad Q_{T_1} = 2 \cdot 3 = 6.$$

Для дерева  $T_2$

$$\rho_{T_2} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{и} \quad Q_{T_2} = 1 \cdot 3 = 3.$$

Для дерева  $T_3$

$$\rho_{T_3} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{и} \quad Q_{T_3} = 1 \cdot 2 = 2.$$

Тогда

$$R = Q_{T_1} + Q_{T_2} + Q_{T_3} = 6 + 3 + 2 = 11.$$

Следовательно,

$$\pi = \frac{1}{11} [6\rho_{T_1} + 3\rho_{T_2} + 2\rho_{T_3}],$$

или

$$\pi = \frac{1}{11} \begin{pmatrix} 5 & -6 & 6 \\ -3 & 8 & 3 \\ 2 & 2 & 9 \end{pmatrix}.$$

Эта матрица проектирует на подпространство  $Z_1$ , натянутое на

векторы ячеек  $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$  и  $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ . Можно проверить, что

$$\pi \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \text{и} \quad \pi \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Матрица  $1 - \pi$  равна

$$1 - \pi = \frac{1}{11} \begin{pmatrix} 6 & 6 & -6 \\ 3 & 3 & -3 \\ -2 & -2 & 2 \end{pmatrix}.$$

Она проектирует на ортогональное дополнение к  $Z_1$ . Ее образ, являющийся ядром  $\pi$ , есть одномерное пространство, определяемое

вектором  $\begin{pmatrix} 6 \\ 3 \\ -2 \end{pmatrix}$ . Этот вектор ортогонален обеим ячейкам  $\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$

и  $\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ . Заметим, что

$$Z \begin{pmatrix} 6 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 6 \\ 3 \\ -2 \end{pmatrix} = \begin{pmatrix} 6 \\ 6 \\ -6 \end{pmatrix}.$$

Как и следовало ожидать, вектор  $\begin{pmatrix} 6 \\ 6 \\ -6 \end{pmatrix}$  задает напряжения в ветвях, что соответствует второму закону Кирхгофа.

### 13.3. Теорема взаимности Грина

Изучение цепей с резисторами мы закончим мощным результатом, называемым теоремой взаимности Грина. С этой теоремой мы встретимся еще раз при изучении цепей, содержащих емкости, и в электростатике. Чтобы доказать эту теорему, представим себе цепь, в которой все ветви, кроме двух, содержат только резисторы без источников тока и напряжения, а в двух ветвях  $\alpha$  и  $\beta$ , кроме резисторов, могут быть такие источники тока или напряжения. При определенном выборе источников в ветвях  $\alpha$  и  $\beta$  мы получим решение с токами  $\mathbf{I}$ , удовлетворяющими первому закону Кирхгофа, и напряжениями  $\mathbf{V}$ , удовлетворяющими второму закону Кирхгофа. Выбрав другие источники, мы получим другое решение, для которого токи  $\hat{\mathbf{I}}$  и напряжения  $\hat{\mathbf{V}}$  также удовлетворяют двум законам Кирхгофа. (Типичная ситуация, к которой применима теорема Грина, показана на рис. 13.8.) Поскольку оба вектора токов  $\mathbf{I}$  и  $\hat{\mathbf{I}}$  удовлетворяют первому закону Кирхгофа, то они лежат в пространстве  $Z_1$ . Аналогично, поскольку оба вектора напряжений  $\mathbf{V}$  и  $\hat{\mathbf{V}}$  удовлетворяют второму закону Кирхгофа,

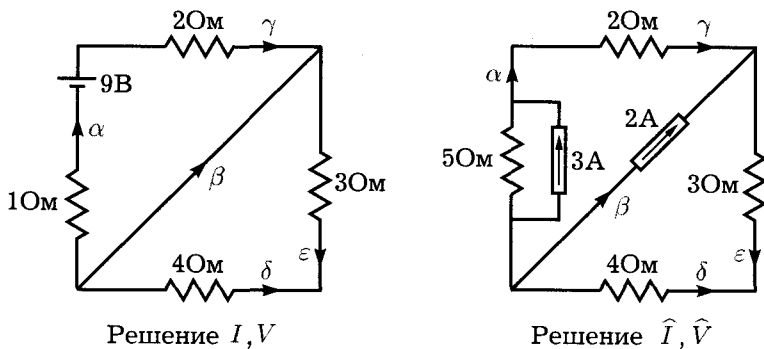


Рис. 13.8

они лежат в пространстве  $B^1$ . Мы знаем, что любой элемент пространства  $B^1$ , действуя на любой элемент пространства  $Z_1$ , дает нуль. Отсюда можно сделать вывод, что

$$\int_I \hat{\mathbf{V}} = 0 = \int_I \mathbf{V}.$$

Тогда, если записать сумму по ветвям, мы получим

$$\sum_{\text{по всем ветвям}} \hat{V}^\gamma I_\gamma = \sum_{\text{по всем ветвям}} V^\gamma \hat{I}_\gamma.$$

Из сумм выделим слагаемые, соответствующие ветвям  $\alpha$  и  $\beta$ :

$$\hat{V}^\alpha I_\alpha + \hat{V}^\beta I_\beta + \sum_{\text{по остальным ветвям}} \hat{V}^\gamma I_\gamma = V^\alpha \hat{I}_\alpha + V^\beta \hat{I}_\beta + \sum_{\text{по остальным ветвям}} V^\gamma \hat{I}_\gamma.$$

Мы предположили, что все ветви, кроме  $\alpha$  и  $\beta$ , имеют только резисторы. Поэтому  $\hat{V}^\gamma = r_\gamma \hat{I}_\gamma$  и  $V^\gamma = r_\gamma I_\gamma$ . Тогда

$$\sum \hat{V}^\gamma I_\gamma = \sum r_\gamma \hat{I}_\gamma I_\gamma = \sum V^\gamma \hat{I}_\gamma,$$

и мы можем сократить суммы по другим ветвям в обеих частях уравнения. В результате получаем

$$\boxed{\hat{V}^\alpha I_\alpha + \hat{V}^\beta I_\beta = V^\alpha \hat{I}_\alpha + V^\beta \hat{I}_\beta},$$

что и называется *теоремой взаимности Грина*.

С помощью теоремы взаимности можно получить ряд удивительных свойств пассивных цепей, содержащих резисторы, т. е. цепей, в которых нет никаких источников. Взяв такую цепь, мы можем добавлять к ней новую ветвь двумя способами, как это показано на рис. 13.9. Начав с цепи, содержащей только резисторы (рис. 13.9(а)), мы можем присоединить новую ветвь  $\alpha$  между двумя узлами (рис. 13.9(б)). Такая «спайка» не создает новых узлов. Эта новая ветвь может быть как коротким замыканием, так и содержать источники тока, источники напряжения или резистор. Новую ветвь можно подключить и так, как показано на рис. 13.9(в), т. е. мы можем «перекусить» провод и вставить в разрыв новую ветвь, создав новый узел.

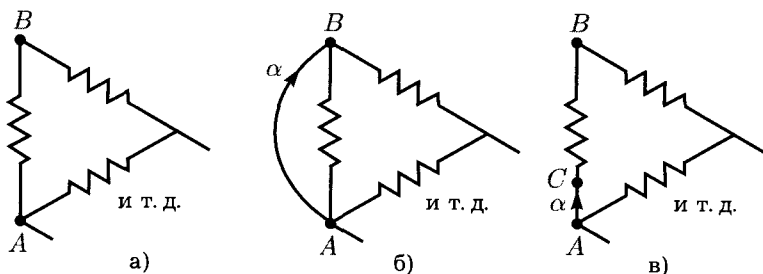


Рис. 13.9. (а) Исходная цепь. (б) Новая ветвь без создания новых узлов (спайка). (в) Новая ветвь с новым узлом (вставка)

Предположим, что у нас есть линейная цепь, показанная на рис. 13.10, и к ней присоединены две новые ветви: ветвь  $\alpha$  между узлами A и B, а ветвь  $\beta$  между узлами C и D. Несущественно, как мы это сделали: припаяв новый провод или вставив в старый. Из теоремы взаимности можно сделать следующие выводы.

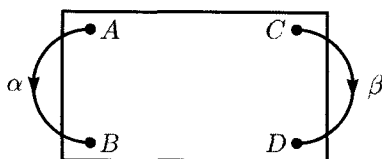


Рис. 13.10

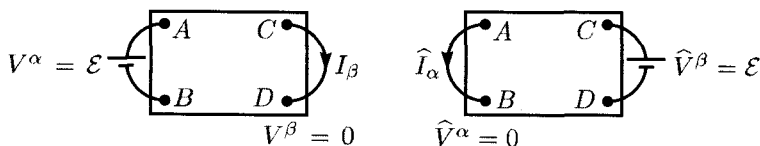


Рис. 13.11

**Случай 1.** Вставим батарею, создающую напряжение  $\mathcal{E}$  в ветвь  $\alpha$ , и измерим ток  $I$  в ветви  $\beta$ . После этого ту же батарею вставим в ветвь  $\beta$  и измерим ток в ветви  $\alpha$  (см. рис. 13.11). Поскольку  $V^\beta = 0$  и  $\hat{V}^\alpha = 0$  (обе ветви соединены накоротко),

из теоремы взаимности следует, что

$$\hat{V}^\beta I_\beta = V^\alpha \hat{I}_\alpha.$$

Мы использовали одну и ту же батарею, так что  $\hat{V}^\beta = V^\alpha = \mathcal{E}$ . Поэтому  $\mathcal{E}I_\beta = \mathcal{E}\hat{I}_\alpha$  и  $I_\beta = \hat{I}_\alpha$ . Этот результат интересен тем, что хотя сама цепь не обязана обладать свойствами симметрии, соотношение между приложенным напряжением и возникающим током симметрично.

**Случай 2.** Вставим источник тока  $j$  в ветвь  $\alpha$  и измерим получившееся напряжение на ветви  $\beta$  в разомкнутом состоянии. Потом тот же источник вставим в ветвь  $\beta$  и измерим ток в ветви  $\alpha$  (см. рис. 13.12). Поскольку  $I_\beta = 0$  и  $\hat{I}_\alpha = 0$  (разомкнутые цепи), то теорема взаимности говорит, что в этом случае

$$\hat{V}^\alpha I_\alpha = V^\beta \hat{I}_\beta.$$

Но

$$I_\alpha = \hat{I}_\beta = j,$$

следовательно,

$$\hat{V}^\alpha j = V^\beta j,$$

и мы получаем

$$\hat{V}^\alpha = V^\beta.$$

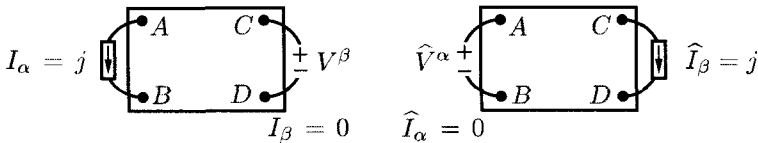


Рис. 13.12

В общем случае мы можем рассмотреть *резисторный  $n$ -порт*, из которого выходят  $n$  пар проводов. Если к различным портам присоединены источники тока  $I_\alpha, I_\beta, \dots$ , то возникающие напряжения  $V^\alpha, V^\beta, \dots$  будут зависеть от токов согласно формуле

$$\mathbf{V} = \mathbf{R}\mathbf{I},$$

где  $R$  — матрица  $n \times n$ . Приведенные соображения показывают, что матрица  $R$  должна быть симметричной, потому что для любых портов  $\lambda$  и  $\mu$  зависимость  $V^\lambda$  от  $I_\mu$  точно такая же, как зависимость  $V^\mu$  от  $I_\lambda$ .

Теорему взаимности Грина можно также вывести как следствие решения, полученного методом круговых токов. В случае, если нет источников тока, т. е.  $\mathbf{K} = \mathbf{0}$ , то мы имеем

$$\mathbf{J} = (sZ\sigma)^{-1}s(-\mathbf{W}).$$

Следовательно,

$$\mathbf{I} = \sigma\mathbf{J} = -\sigma(sZ\sigma)^{-1}s\mathbf{W}.$$

Матрица  $Z$  симметрична, а матрица  $s$  — транспонированная матрица  $\sigma$ . Поэтому матрица

$$G = -\sigma(sZ\sigma)^{-1}s$$

симметрична. Тогда можно написать

$$\mathbf{I} = G\mathbf{W},$$

где матрица  $G$ , выражающая токи в ветвях через приложенное напряжение батареи, симметрична. Аналогично можно начать с решения, полученного методом узловых потенциалов:

$$\Phi = ([\partial]Z^{-1}[d])^{-1}[\partial](\mathbf{K} - \mathbf{Z}^{-1}\mathbf{W}).$$

Если нет источников напряжения, т. е.  $\mathbf{W} = \mathbf{0}$ , то

$$\mathbf{V} = -[d]\Phi = -[d]([\partial]Z^{-1}[d])^{-1}[\partial]\mathbf{K}.$$

На этот раз матрица  $[d]$  — транспонированная матрица  $[\partial]$ . Тогда матрица

$$R = [d]([\partial]Z^{-1}[d])^{-1}[\partial],$$

выражающая напряжения на ветвях через приложенные токи, симметрична.

В связи с этим стоит заметить, что зависимость  $\mathbf{V}$  от  $\mathbf{W}$  или  $\mathbf{I}$  от  $\mathbf{K}$  в общем случае не дается симметричной матрицей.

## 13.4. Цепи с конденсаторами

Немного изменим задачу — вместо цепей с резисторами рассмотрим цепи с конденсаторами. Теперь каждая ветвь может иметь батарею и конденсатор, соединенные последовательно. Когда подсоединена батарея<sup>2</sup>, на пластинах конденсатора накапливается заряд и со временем устанавливается стационарное состояние. Поскольку заряды не могут пересечь расстояние между пластинами конденсатора, то в цепи не будет никакого тока. Нам нужно узнать величину заряда  $Q_\alpha$  на конденсаторе в ветви  $\alpha$  и напряжение  $V^\alpha$  на этой ветви.

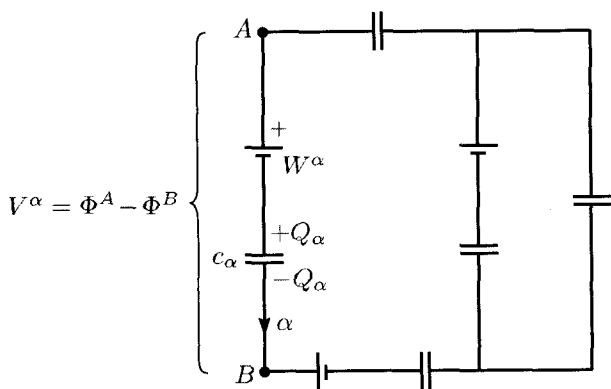


Рис. 13.13

Типичный пример такой цепи изображен на рис. 13.13. На ветви  $\alpha$  обозначены знаки зарядов на пластинах конденсаторов. Заметим, что так же, как и раньше, положительные значения  $V^\alpha$  и  $W^\alpha$  относятся к падениям напряжения, когда ветвь проходит в направлении, указанном стрелкой. Положительный знак  $Q_\alpha$  означает, что сначала проходит положительно заряженная пластина конденсатора. Договорившись об этом, обозначим  $C_\alpha$  емкость ветви и напишем уравнение

$$V^\alpha - W^\alpha = \frac{Q_\alpha}{C_\alpha}.$$

<sup>2</sup> постоянного тока. — Прим. ред.

Мы можем рассматривать вектор  $\mathbf{Q} = (Q_\alpha, Q_\beta, \dots)$  как 1-цепь, а вектор  $\mathbf{V} = (V^\alpha, V^\beta, \dots)^T$  — как 1-коцепь. Однако, теперь связь между напряжением и зарядом дает энергию, а не мощность<sup>3</sup>. В векторных обозначениях мы можем написать

$$\mathbf{V} - \mathbf{W} = C^{-1}\mathbf{Q},$$

где  $C$  — диагональная матрица с матричными элементами  $C_\alpha$ . Если заменить  $Z$  на  $C^{-1}$  и мощность на энергию, то ситуация формально совпадает с той, что возникает в цепях с резисторами. Напряжение  $\mathbf{V}$  по-прежнему получается из потенциала  $\mathbf{V} = -d\Phi$ . Поскольку ток не может течь через конденсаторы, то полный заряд на конденсаторах, соединенных с любым узлом, не может изменяться. Если вначале конденсаторы не были заряжены, то полный заряд на каждом узле будет равен нулю и  $\partial\mathbf{Q} = 0$ , точно так же, как мы имели  $\partial\mathbf{I} = 0$  для цепей с резисторами.

Изолированные цепи конденсаторов, которые вначале не были заряжены и для которых справедливо условие  $\partial\mathbf{Q} = 0$ , можно рассматривать так же, как мы рассматривали цепи с резисторами. В предыдущих результатах надо только заменить  $\mathbf{I}$  на  $\mathbf{Q}$  и  $Z$  на  $C^{-1}$ . Батареи, как и раньше, представляются вектором  $\mathbf{W}$ , и нет ничего, что играло бы роль источника тока. Тогда по аналогии с методом круговых токов можно ввести *заряды ячеек*, описываемые вектором  $\mathbf{P} \in Z_1$ . В этом случае любое распределение зарядов, удовлетворяющее условию  $\partial\mathbf{Q} = 0$ , можно записать в форме  $\mathbf{Q} = \sigma\mathbf{P}$ . Тогда

$$\mathbf{P} = (sC^{-1}\sigma)^{-1}(-s\mathbf{W}).$$

Аналогично, мы можем заменить  $Z^{-1}$  на  $C$  в методе узловых потенциалов и получить

$$\Phi = -([\partial]C[d])^{-1}[\partial]C\mathbf{W}.$$

---

<sup>3</sup>Следует напомнить, что в отличие от выделяемой в резисторном участке цепи с сопротивлением  $R$  при напряжении  $U$  и токе  $I$  мощности  $P = UI = IR^2 = U^2R^{-1}$ , потенциальная энергия конденсатора емкости  $C$  при заряде пластин  $\pm Q$  и напряжении  $U$  составляет  $E = \frac{1}{2}UQ = \frac{1}{2}CU^2 = \frac{1}{2}Q^2C^{-1}$ ; так что аналогия «мощность–энергия» здесь имеет место лишь с точностью до множителя  $\frac{1}{2}$ . — Прим. ред.

В общем случае, если конденсаторы вначале были заряжены, на пластинах, присоединенных к узлу  $A$ , может быть фиксированный полный заряд  $\rho_A$ . Вектор  $\rho = (\rho_A, \rho_B, \dots)$  может рассматриваться как 0-цепь. Поскольку мы договорились, что с конечным узлом ветви соединена отрицательная пластина конденсатора, то в общем случае можно написать  $\partial Q = -\rho$ . (Посмотрите на рис. 13.14 и убедитесь в правильности знака  $-$ .) Сумма зарядов на двух пластинах каждого конденсатора равна нулю, следовательно,

$$\sum_{\text{по всем узлам}} \rho_A = 0.$$

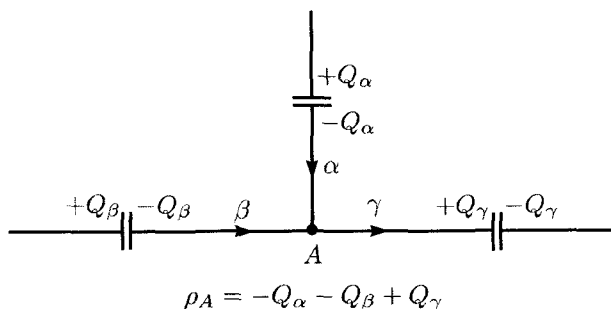


Рис. 13.14

Уравнение  $\partial Q = -\rho$  называется *законом Гаусса*. Закон показывает связь между зарядами узлов и зарядами на пластинах конденсатора в ветвях. В главе 16 мы будем изучать электростатику и там увидим, что аналогом узловых зарядов будет плотность заряда  $\rho$ , а аналогом  $Q$  — диэлектрическое смещение  $D$ . Мы объясним, что  $D$  можно рассматривать как 2-форму в трехмерном пространстве и опишем, как надо ставить физический эксперимент, чтобы измерить значение  $D$ . Окажется, что закон Гаусса — это просто трехмерный вариант теоремы Стокса. В большинстве книг электростатика излагается в рамках векторного анализа, где диэлектрическое смещение  $D$  считается векторным полем. Закон Гаусса становится теоремой о дивергенции. Однако, в рамках нашего исследования цепей с конденсаторами уравнение  $\partial Q = -\rho$  означает, что каждому узлу мы приписываем сумму

зарядов на пластинах конденсатора, присоединенного к нему. До конца этой главы мы будем изучать цепи без источников напряжения, т. е. с  $\mathbf{W} = 0$ , где конденсаторы *имеют* первоначальный заряд. Тогда наши уравнения принимают вид

$$\mathbf{Q} = C\mathbf{V}, \quad \mathbf{V} = -d\Phi \quad \text{и} \quad \partial\mathbf{Q} = -\rho.$$

Эти три уравнения можно собрать вместе и получить

$$-\partial C d\Phi = -\rho.$$

Это уравнение называется *уравнением Пуассона*. Оно связывает функцию потенциала с зарядами узлов. Оператор  $-\partial C d$ , стоящий в левой части уравнения Пуассона, называется *лапласианом* и обозначается  $\Delta$ . Тогда уравнение Пуассона можно переписать в форме

$$\Delta\Phi = -\rho.$$

Теперь оператор  $d$  отображает  $C^0 \rightarrow C^1$ ,  $C$  отображает  $C^1 \rightarrow C_1$  и  $\partial$  отображает  $C_1 \rightarrow C_0$ . Поэтому лапласиан  $\Delta$  отображает  $C^0 \rightarrow C_0$ . Как и следовало ожидать, потенциал  $\Phi$ , стоящий в левой части уравнения Пуассона, принадлежит  $C^0$ , а  $\rho$  из правой части этого уравнения принадлежит  $C_0$ .

Полезно написать явную форму лапласиана. Пусть  $u$  — элемент пространства  $C^0$ , а  $A$  — узел. Если ветвь  $\alpha$  имеет  $A$  в качестве одного из концов, то

$$du(\alpha) = \pm(u(A) - u(B)), \quad \text{если} \quad \partial\alpha = \pm(A - B).$$

Поэтому  $C du$  имеет коэффициент

$$\pm C_\alpha(u(A) - u(B)) \quad \text{для ветви} \quad \alpha.$$

Подеиствуем теперь оператором  $\partial$  и найдем коэффициент в узле  $A$  (суммируя по всем  $\alpha$  с условием  $\partial\alpha = \pm(A - B)$  для узла  $B$ ):

$$(\Delta u)(A) = \sum_{\alpha: \partial\alpha = \pm(B-A)} C_\alpha(u(B) - u(A)).$$

В частности,  $u$  удовлетворяет уравнению Лапласа

$$\Delta u = 0,$$

тогда и только тогда, когда

$$u(A) = \frac{1}{\sum C_\alpha} \sum C_\alpha u(B),$$

где сумма берется по всем узлам  $B$ , являющимся концами ветвей, выходящих из узла  $A$ , как это показано на рис. 13.15. Другими словами, уравнение Лапласа говорит, что значение  $u$  на узле  $A$  есть среднее от всех значений на *ближайших соседних* узлах, причем весами являются емкости соответствующих ветвей.

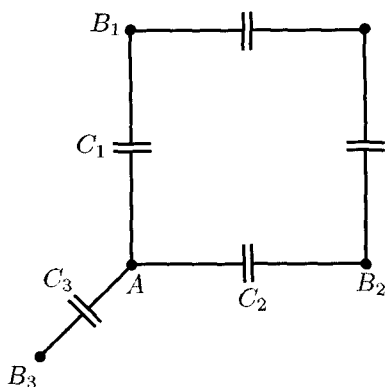


Рис. 13.15

Эту интерпретацию следует сравнить с уравнением в частных производных

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

введенным в главе 6. Это уравнение говорит, что значение  $u$  в точке  $P$  есть среднее от  $u$  по малым окружностям с центром в точке  $A$ . Там, чтобы определить понятие окружности, мы должны были иметь евклидову геометрию на плоскости. Здесь, чтобы определить усреднение по ближайшим соседям, нам требуется матрица емкостей  $C$ . Существует аналогия между выбором матрицы  $C$  и выбором евклидовой геометрии на плоскости. Мы рассмотрим это подробнее в главе 16.

А пока вернемся к изучению уравнения Пуассона. Заметим,

что потенциал  $\Phi$  в левой части уравнения определен с точностью до функции, которая постоянна для всех связанных компонент цепи. Действительно, любая такая функция принадлежит ядру  $d$  и, следовательно, ядру  $\Delta = -\partial C d$ . Величина  $\rho$  в правой части уравнения лежит в подпространстве  $B_0$  пространства  $C_0$ , поскольку, по определению, узловые заряды получаются в результате действия  $\partial$  на заряды ветвей. Вспомним, что мы ввели факторпространство  $P^0 = C^0/Z^0$ , ограниченный кограничный оператор  $[d] : P^0 \rightarrow C^1$  и ограниченный граничный оператор  $[\partial] : C_1 \rightarrow B_0$ . Ситуацию в целом можно изобразить на схеме отображений

$$C^0 \xrightarrow{\quad} P^0 \xrightarrow{[d]} C^1 \xrightarrow{C} C_1 \xrightarrow{[\partial]} B_0 \longrightarrow C_0.$$

$\Delta$

Это значит, что можно ввести *ограниченный лапласиан*  $[\Delta] = -[\partial]C[d]$  и рассмотреть ограниченное уравнение Пуассона

$$[\Delta]\Psi = -\rho$$

для  $\Psi \in P^0$ . (Другими словами,  $\Psi$  является потенциалом, определенным с точностью до аддитивной постоянной на каждой связанной компоненте цепи.) Далее, поскольку оператор  $C$  дается положительной диагональной матрицей в стандартном базисе  $C^1$  и  $C_1$ , из теоремы Кирхгофа (с формулировкой для узловых потенциалов), мы знаем, что матрица  $[\partial]C[d]$  обратима и, следовательно, решение ограниченного уравнения Пуассона имеет вид

$$\Psi = ([\partial]C[d])^{-1}\rho.$$

**Пример.** Задача, где применяется простое уравнение Пуассона, изображена на рис. 13.16. На узле  $B$  заряд равен 4, на узле  $C$  он равен 1. Мы хотим найти потенциалы на всех узлах. (Для конкретности заряд будем определять в микрокулонах, емкость в микрофарадах, а потенциал в вольтах.)

Пусть узел  $A$  заземлен. Тогда ограниченный граничный оператор

$$[\partial] = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

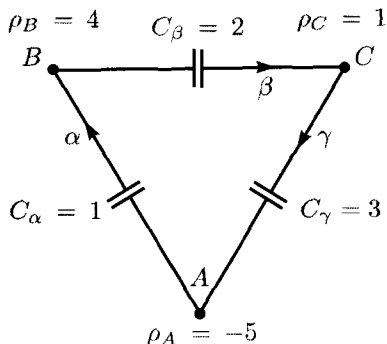


Рис. 13.16

(напоминаем, что убрана строка, соответствующая узлу  $A$ ). Матрица оператора  $[d]$  является транспонированной к матрице  $[\partial]$ , т. е.

$$[d] = \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix}.$$

Матрица  $C$  равна  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ , поэтому

$$[\partial]C[d] = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 3 & -2 \\ -2 & 5 \end{pmatrix}.$$

Решение уравнения Пуассона имеет вид

$$([\partial]C[d])^{-1} = \frac{1}{11} \begin{pmatrix} 5 & 2 \\ 2 & 3 \end{pmatrix}, \quad \Psi = \frac{1}{11} \begin{pmatrix} 5 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

## 13.5. Краевые задачи

Задача для уравнения Пуассона, которую мы только что решили, аналогична электростатической задаче, когда вычисляется потенциал  $\Phi(\mathbf{r})$ , создаваемый распределением заряда  $\rho(\mathbf{r})$  в

вакууме, когда не заданы граничные условия. В электростатике возникают более интересные задачи, когда конкретизируются граничные условия для нулевого потенциала на определенных проводниках, или когда фиксируется потенциал на границе области. Здесь мы рассмотрим несколько таких примеров.

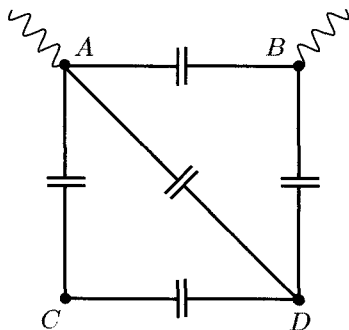


Рис. 13.17

На рис. 13.17 изображена система конденсаторов, причем ни в одной ветви нет батарей. Узлы этой цепи разделим на два класса: *граничные узлы* и *внутренние узлы*. Граничные узлы  $A$  и  $B$  соединяются с внешними источниками, которые поддерживают на них определенные значения потенциалов. Поскольку заряды могут свободно перетекать по проводам, соединяющим эти узлы с внешними источниками, то нет смысла задавать на граничных узлах значение полного заряда. Внутренние узлы  $C$  и  $D$  соединены только с другими узлами цепи, и не подходят к внешним источникам. Полный заряд на таких узлах постоянен и может быть задан. При этом нет смысла задавать потенциал на внутренних узлах.

Для расчета такой цепи в общем случае мы будем задавать потенциал на граничных узлах и полный заряд на внутренних узлах. Задача в том, чтобы найти потенциалы на внутренних узлах и полные заряды на граничных. По существу эта задача может быть представлена как суперпозиция двух более ограниченных проблем.

1. *Задача Дирихле*. Полный заряд на каждом внутреннем узле полагается равным нулю, а потенциал на каждом гранич-

ном узле имеет определенное значение (в общем случае не равное нулю).

2. *Задача для уравнения Пуассона.* Потенциал на каждом граничном узле полагается равным нулю, а полный заряд  $\rho$  на каждом внутреннем узле имеет определенное значение. Задача именно такого типа изображена на рис. 13.15, где узел  $A$  является граничным.

Решение общей задачи всегда может быть представлено как суперпозиция решений задачи Дирихле с соответствующими граничными условиями и задачи Пуассона с соответствующими значениями зарядов на внутренних узлах.

Разделив узлы на два класса, мы получили два векторных пространства:  $C_0$  (узловые заряды) и  $C^0$  (узловые потенциалы). Пространство  $C_0$  есть прямая сумма

$$C_0 = C_0^{\text{гран}} \oplus C_0^{\text{внутр}},$$

где  $C_0^{\text{гран}}$  состоит из 0-цепей, имеющих ненулевые коэффициенты только на граничных узлах, и  $C_0^{\text{внутр}}$  состоит из 0-цепей, имеющих ненулевые коэффициенты только на внутренних узлах. Аналогично мы имеем

$$C^0 = C_{\text{гран}}^0 \oplus C_{\text{внутр}}^0,$$

где  $C_{\text{гран}}^0$  состоит из тех линейных функций, которые обращаются в нуль в  $C_{\text{внутр}}^0$ , а  $C_{\text{внутр}}^0$  состоит из функций, обращающихся в нуль в  $C_{\text{гран}}^0$ .

Потенциал  $\Phi$  может быть разложен единственным способом в сумму элементов из пространства  $C_{\text{гран}}^0$  и  $C_{\text{внутр}}^0$ , т. е.

$$\Phi = \Phi_{\text{гран}} + \Phi_{\text{внутр}}.$$

Для цепи, изображенной на рис. 13.17 с граничными узлами  $A$ ,  $B$  и внутренними узлами  $C$ ,  $D$ , такое разложение имеет вид

$$\begin{pmatrix} \Phi^A \\ \Phi^B \\ \Phi^C \\ \Phi^D \end{pmatrix} = \begin{pmatrix} \Phi^A \\ \Phi^B \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \Phi^C \\ \Phi^D \end{pmatrix}.$$

Аналогично разлагается вектор, представляющий заряды узлов

$$\rho = \rho^{\text{гран}} + \rho^{\text{внутр}}.$$

Тогда в цепи, изображенной на рис. 13.17, мы имеем

$$\begin{pmatrix} \rho_A \\ \rho_B \\ \rho_C \\ \rho_D \end{pmatrix} = \begin{pmatrix} \rho_A \\ \rho_B \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \rho_C \\ \rho_D \end{pmatrix}.$$

Теперь задачи могут быть сформулированы в терминах разложений этих векторных пространств.

**1. Задача Дирихле.** Поскольку потенциал задается на каждом граничном узле, вектор  $\Phi_{\text{гран}}$  известен. Более того, поскольку на внутренних узлах нет зарядов, вектор  $\rho^{\text{внутр}} = 0$ . Мы должны найти такой вектор  $\Phi_{\text{внутр}}$ , чтобы на всех внутренних узлах выполнялось условие

$$\Delta(\Phi_{\text{внутр}} + \Phi_{\text{гран}}) = 0.$$

На граничных узлах мы имеем уравнение

$$\Delta(\Phi_{\text{внутр}} + \Phi_{\text{гран}}) = -\rho^{\text{гран}}.$$

Решив эти уравнения, мы узнаем потенциалы и заряды на всех узлах.

**2. Задача Пуассона.** На каждом внутреннем узле задается заряд. Следовательно, вектор  $\rho^{\text{внутр}}$  известен. Поскольку потенциал равен нулю на каждом граничном узле,  $\Phi_{\text{гран}} = 0$ . Мы должны найти такой вектор  $\Phi_{\text{внутр}}$ , чтобы на всех внутренних узлах выполнялось условие

$$\Delta(\Phi_{\text{гран}} + \Phi_{\text{внутр}}) = -\rho^{\text{внутр}}.$$

Раз потенциал известен, распределение зарядов на граничных узлах определяется уравнением

$$\rho^{\text{гран}} = -\Delta(\Phi_{\text{гран}} + \Phi_{\text{внутр}}).$$

### 13.6. Решение краевых задач методом ортогональных проекций Вейля

Сначала мы рассмотрим связную цепь, имеющую  $n_i$  внутренних узлов и  $n_b$  граничных узлов. Обе задачи могут быть решены методом ортогональных проекций Вейля. Для определения скалярного произведения напряжений на ветвях в пространстве  $C^1$  воспользуемся емкостной матрицей  $C : C^1 \rightarrow C_1$  и напомним формулу

$$(\mathbf{V}, \hat{\mathbf{V}})_C = \int_{C\mathbf{V}} \hat{\mathbf{V}} = \sum_{\text{по всем ветвям}} C_\alpha V^\alpha \hat{V}_\alpha.$$

Это скалярное произведение положительно определено, потому что  $C$  — диагональная матрица, у которой все матричные элементы положительны. Не удивительно, что это выражение может быть интерпретировано через энергию. Действительно, для любого значения  $V$  формула

$$\frac{1}{2}(\mathbf{V}, \hat{\mathbf{V}})_C = \sum \frac{1}{2} C_\alpha (V^\alpha)^2$$

представляет полную энергию, накопленную в конденсаторах.

Относительно этого скалярного произведения пространство  $C^1$  разлагается в ортогональную прямую сумму

$$C^1 = d(C^0) \oplus C^{-1}Z_1.$$

Действительно, мы знаем, что  $d(C^0)$  является аннулирующим подпространством подпространства  $Z_1$  пространства  $C_1$ . Вычисление скалярного произведения элемента  $V$  из  $C_1$  на элемент типа  $C^{-1}I$  равносильно вычислению скалярного произведения  $V$  на  $I$ :

$$(V, C^{-1}I)_C = \int_I V.$$

Разложим далее пространство  $d(C^0)$ . Поскольку  $dC^0_{\text{внутр}}$  является подпространством  $C^0$ , то его ортогональное дополнение внутри  $dC^0$  есть подпространство  $dC^0$ , которое обозначим  $D^1$ . Тогда мы имеем разложение в виде прямой суммы

$$dC^0 = D^1 \oplus dC^0_{\text{внутр}}.$$

Следовательно,

$$C^1 = D^1 \oplus dC_{\text{внутр}}^0 \oplus C^{-1}Z_1 \quad (13.1)$$

является разложением пространства  $C^1$  на три попарно ортогональных подпространства. Заметим, что всякий элемент пространства  $D^1$  имеет вид  $d\Phi$ , потому что  $D^1 \subset dC^0$ . Кроме того, если  $V \in D^1$ , то  $V \perp dC_{\text{внутр}}^0$ , и тогда

$$0 = (V, d\Phi_{\text{внутр}})_C = \int_{CV} d\Phi_{\text{внутр}} = \int_{\partial CV} \Phi_{\text{внутр}}.$$

Для любого внутреннего узла мы можем найти такую функцию  $\Phi_{\text{внутр}}$ , которая не равна нулю на этом узле и обращается в нуль на всех других узлах. Тогда  $\partial CV$  должна иметь нулевые компоненты для всех внутренних узлов, т. е.  $\partial CV \in C_0^{\text{гран}}$ . Но  $V = -d\Phi$ , следовательно,

$$\Delta\Phi = 0 \quad \text{для всех внутренних узлов.}$$

Другими словами,  $D^1$  состоит из  $d\Phi$ , где  $\Phi$  есть решение задачи Дирихле. Тем самым  $\dim C^{-1}Z_1 = \dim Z_1 = m$ . Если есть по крайней мере один граничный узел, то  $d: C_{\text{внутр}}^0 \rightarrow C^1$  инъективно, поэтому  $\dim dC_{\text{внутр}}^0 = n_i$ . Кроме того,  $\dim C^1 = n + m - 1 = n_b - 1 + n_i + m$ . Тогда из уравнения (13.1) следует, что

$$\dim D^1 = n_b - 1.$$

Если исключить заземленный узел, то останется  $n_b - 1$  граничных узлов, на которых мы можем задать произвольные значения  $\Phi$ . Таким образом, задача Дирихле имеет решение всегда.

В качестве примера разложения (13.1) рассмотрим цепь, изображенную на рис. 13.18. Здесь  $A$ ,  $B$  и  $C$  — граничные узлы, а  $D$  — внутренний узел. Поскольку в цепи есть четыре ветви, то  $\dim C^1 = 4$ . Тогда получаются следующие ортогональные пространства.

1. Одномерное подпространство  $C^{-1}Z_1$  с базисным вектором

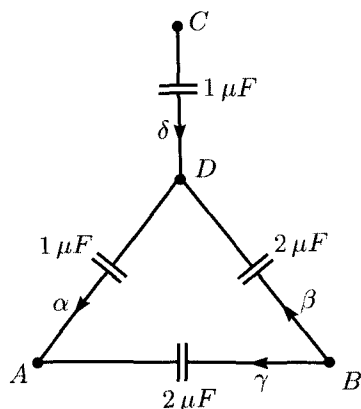


Рис. 13.18

$\begin{pmatrix} 1 \\ 1/2 \\ -1/2 \\ 0 \end{pmatrix}$ . Этот вектор соответствует распределению зарядов  $C$ 
 $\begin{pmatrix} 1 \\ 1/2 \\ -1/2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 0 \end{pmatrix}$ , для которого  $\rho = 0$  во всех узлах.

Конечно, при этом нарушается второй закон Кирхгофа, поэтому эти заряды нельзя получить из потенциала.

2. Одномерное подпространство  $dC_{\text{внутр}}^0$  с базисным вектором  $\begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix}$ . Этот вектор описывает напряжения на ветвях, возникающие от потенциала

$$\Phi^A = \Phi^B = \Phi^C = 0, \quad \Phi^D = 1.$$

3. Двумерное подпространство  $D^1$  с базисными векторами

$$\begin{pmatrix} 2/3 \\ 1/3 \\ 1 \\ 0 \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} 1/3 \\ -1/6 \\ 1/6 \\ 2/3 \end{pmatrix}.$$

Эти вектора описывают напряжения на ветвях, возникшие из

$$\Phi^A = \Phi^C = 0, \quad \Phi^B = 1 \quad \text{и} \quad \Phi^A = \Phi^B = 0, \quad \Phi^C = 1$$

соответственно. Причем в обоих случаях  $\rho_D = 0$ . (В общем случае совсем не просто сразу написать базис для  $D^1$ .)

## Решение уравнения Пуассона

Сейчас мы уже можем решить уравнение Пуассона методом ортогональных проекций Вейля. Предположим, что на всех внутренних узлах заданы значения зарядов  $\rho$ . Мы построим *любое* распределение напряжений  $\hat{V}$ , для которого  $-\partial C \hat{V} = \rho$  имеет данные значения на всех внутренних узлах. (При этом  $\hat{V}$  может даже нарушить второй закон Кирхгофа.) Тогда существует единственное разложение

$$\hat{V} = V + W,$$

где  $V$  является элементом  $dC_{\text{внутр}}^0$  и  $W$  ортогонален  $dC_{\text{внутр}}^0$ . Будучи элементом  $dC_{\text{внутр}}^0$ , распределение напряжений  $V$  может быть получено из потенциала при условии, что на граничных узлах  $\Phi = 0$ . Поскольку вектор  $W$  ортогонален  $V$ , то он лежит в  $D^1 \oplus C^{-1}Z_1$  и поэтому удовлетворяет условию  $\partial C W = 0$  для всех внутренних узлов. Тогда для внутренних узлов  $-\partial C V = -\partial C \hat{V} = \rho$ , и  $V$  является решением задачи. Если  $\pi$  обозначает ортогональную проекцию  $C^1$  на  $dC_{\text{внутр}}^0$ , то решение можно записать в форме

$$V = \pi \hat{V}.$$

## Решение задачи Дирихле

Задача Дирихле тоже может быть решена методом ортогональных проекций. Пусть  $\hat{\Phi}$  обозначает потенциал, который принимает определенные значения на граничных узлах и который равен нулю на всех внутренних узлах. Пусть  $\hat{V} = -d\hat{\Phi}$ . Тогда

$$\hat{V} = \pi \hat{V} + (1 - \pi) \hat{V}.$$

Поскольку  $\pi \hat{\mathbf{V}} \in dC_{\text{внутр}}^0$ , мы можем написать  $\pi \hat{\mathbf{V}} = -d\psi$ , где  $\psi = 0$  на граничных узлах. Если записать

$$(1 - \pi)\hat{\mathbf{V}} = -d\Phi,$$

то поскольку  $(1 - \pi)\hat{\mathbf{V}} \in D^1$ , на всех внутренних узлах удовлетворяется уравнение  $\Delta\Phi = 0$ . Более того, поскольку

$$-d\hat{\Phi} = -d\psi - d\Phi$$

и  $\psi = 0$  для всех граничных узлов, то на границе  $\Phi = \hat{\Phi}$ . Тогда  $\Phi$  и есть желаемое решение задачи Дирихле.

## Примеры

Оба подхода к решению задачи могут быть продемонстрированы на примере цепи, изображенной на рис. 13.19.

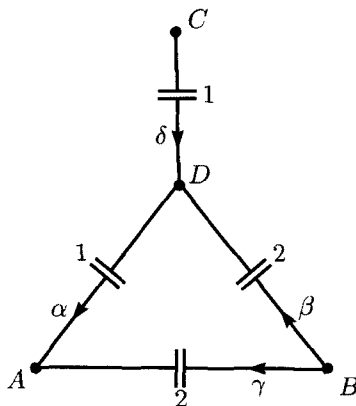


Рис. 13.19

Для этой цепи базисом в пространстве  $dC_{\text{внутр}}^0$  является вектор

$$\mathbf{U} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix},$$

квадрат длины которого (в соответствии с определением скалярного произведения через матрицу  $C$ ) равен  $1 + 2 + 1 = 4$ . Тогда для любого вектора  $\mathbf{V} \in C^1$  мы имеем

$$\pi \mathbf{V} = \frac{1}{4}(\mathbf{V}, \mathbf{U})_C \mathbf{U}.$$

Действуя оператором  $\pi$  на базисные вектора  $\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \dots$ , мы получим матрицы

$$\pi = \frac{1}{4} \begin{pmatrix} 1 & -2 & 0 & -1 \\ -1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix} \quad \text{и} \quad (1 - \pi) = \frac{1}{4} \begin{pmatrix} 3 & 2 & 0 & 1 \\ 1 & 2 & 0 & -1 \\ 0 & 0 & 4 & 0 \\ 1 & -2 & 0 & 3 \end{pmatrix}.$$

Предположим, что мы хотели бы решить уравнение Пуассона для  $\rho_D = 1$ . Заряд, равный единице на узле  $D$ , может быть в случае  $\hat{Q}_\alpha = \hat{Q}_\beta = \hat{Q}_\gamma = 0$  и  $\hat{Q}_\delta = -1$ , так что  $\hat{\mathbf{V}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}$ . Тогда

$$\pi \hat{\mathbf{V}} = \frac{1}{4} \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix} \quad \text{дает нам решение задачи.}$$

Заметим, что  $\hat{\mathbf{V}}$  не обязан удовлетворять второму закону Кирхгофа. Кроме того,  $\rho_D = 1$  и при другом распределении зарядов:  $\hat{Q}_\alpha = \frac{1}{3}$ ,  $\hat{Q}_\beta = -\frac{1}{3}$ ,  $\hat{Q}_\gamma = 1$ ,  $\hat{Q}_\delta = -\frac{1}{3}$ . Тогда

$$\hat{\mathbf{V}} = C^{-1} \hat{\mathbf{Q}} = \begin{pmatrix} 1/3 \\ -1/6 \\ 2 \\ -1/3 \end{pmatrix}.$$

Это распределение напряжений не удовлетворяет второму закону Кирхгофа, потому что сумма падений напряжения при обходе

ячейки  $\beta + \alpha - \gamma$  не равна нулю. И тем не менее

$$\pi \hat{\mathbf{V}} = \frac{1}{4} \begin{pmatrix} 1 & -2 & 0 & -1 \\ -1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ -1 & 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1/3 \\ -1/6 \\ 2 \\ -1/3 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix},$$

что является правильным решением задачи.

Чтобы построить  $\hat{\mathbf{V}}$ , проще всего (когда это возможно) приписать заряд на каждом внутреннем узле одной ветви, которая связывает эту ветвь с граничным узлом. В сложных цепях, где некоторые внутренние узлы связаны только с внутренними узлами, лучше всего начинать именно с таких узлов и двигаться изнутри наружу. Например, на рис. 13.20 граничные узлы — это  $A$  и  $E$ . Лучше всего начинать с узла  $C$ . Если  $\rho_B = 1$ ,  $\rho_C = 2$ ,  $\rho_D = 3$ , то выберем  $\hat{Q}_\gamma = 2$ . Тогда  $\hat{Q}_\delta = 5$ ,  $\hat{Q}_\alpha = -1$ , так что

$$\hat{\mathbf{Q}} = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 5 \end{pmatrix} \quad \text{и} \quad \hat{\mathbf{V}} = \begin{pmatrix} -1 \\ 0 \\ 6 \\ 20 \end{pmatrix}.$$

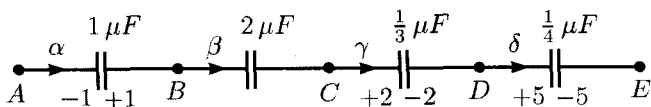


Рис. 13.20

Возвращаясь к рис. 13.19, давайте решим задачу Дирихле для  $\Phi^A = 0$ ,  $\Phi^B = 5$  и  $\Phi^C = 6$ . Положим  $\hat{\Phi} = \Phi$  на границе и  $\hat{\Phi}^D = 0$ ,

так что  $\hat{\mathbf{V}} = -d\hat{\Phi} = \begin{pmatrix} 0 \\ 5 \\ 5 \\ 6 \end{pmatrix}$ . Тогда

$$\mathbf{V} = (1 - \pi)\hat{\mathbf{V}} = \frac{1}{4} \begin{pmatrix} 3 & 2 & 0 & 1 \\ 1 & 2 & 0 & -1 \\ 0 & 0 & 4 & 0 \\ 1 & -2 & 0 & 3 \end{pmatrix} \begin{pmatrix} 0 \\ 5 \\ 5 \\ 6 \end{pmatrix} = \frac{1}{4} \begin{pmatrix} 16 \\ 4 \\ 20 \\ 8 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \\ 5 \\ 2 \end{pmatrix}$$

дает правильное решение при  $\Phi^D = 4$ . То же самое решение можно было бы получить для любого произвольно выбранного значения  $\hat{\Phi}^D$ .

### 13.7. Функции Грина

Отображение, приписывающее распределению зарядов  $\rho$  соответствующий потенциал  $u$ , с которым решается уравнение Пуассона, называется *оператором Грина*. Матричные элементы этого оператора  $\mathbf{G}$  в базисе  $C_0$ , состоящем из узлов, будем обозначать  $G(A, B)$ , где  $A$  и  $B$  — соответствующие узлы. Мы используем обозначение  $G(A, B)$  вместо привычного  $G_{AB}$ , потому что позднее, когда будут рассматриваться непрерывные распределения зарядов, они удобнее для соответствующих формул. Таким образом,

$$\begin{aligned} G(A, B) = \text{потенциалу в узле } B \\ \text{благодаря решению уравнения Пуассона} \\ \text{с единичным зарядом в узле } A. \end{aligned}$$

Функция  $G$ , приписывающая узлам  $A$  и  $B$  матричный элемент  $G(A, B)$ , называется *функцией Грина*. Она зависит от двух переменных. Если фиксировать первую переменную и рассматривать зависимость от второй переменной, то получается элемент пространства  $C_{\text{внутр}}^0$ . Пусть  $\Delta_2$  обозначает оператор  $\Delta$ , примененный к элементу  $G(A, \cdot)$  пространства  $C_{\text{внутр}}^0$ . Тогда

$$\begin{aligned} -\Delta G(A, \cdot) &= \delta_A, \\ G(A, B) &= 0 \quad \text{для узла } B \text{ на границе.} \end{aligned}$$

Повторим эту операцию:

$$\Delta : C_{\text{внутр}}^0 \rightarrow C_0^{\text{внутр}}.$$

Оператор, обратный к  $-\Delta$ , есть  $G$ , поэтому

$$-\Delta u = \rho, \quad \text{если} \quad \mathbf{G}\rho = u.$$

Оператор  $\Delta$  симметричен, следовательно, симметричен и оператор  $G$ . Если взять базис, заданный узлами, то

$$u(B) = \sum \rho(A)G(A, B) \quad \text{для всех } B,$$

где суммирование ведется по всем внутренним узлам  $A$ .

Функцию Грина можно использовать для решения задачи Дирихле. Для этого удобна формула для суммирования, известная как *формула Грина*. Заметим, что если  $u$  и  $v$  — произвольные элементы  $C^0$ , то

$$\begin{aligned} \sum_{\text{по всем } A} u(A)\Delta v(A) &= \sum_{\text{по всем } A} u(A)\partial C dv(A) \\ &= \int_{\partial C} u = \int_C du = (du, dv), \end{aligned}$$

где символ  $(\cdot, \cdot)$  обозначает симметричное скалярное произведение в  $C^1$ . Тогда

$$\sum_{\text{по всем } A} u(A)\Delta v(A) = \sum_{\text{по всем } A} \Delta u(A)v(A).$$

Если мы разделим всю сумму на слагаемое по граничным узлам и слагаемое по внутренним узлам, то получим формулу Грина:

$$\begin{aligned} \sum_{\text{по внутр. } A} (u(A)\Delta v(A) - \Delta u(A)v(A)) \\ = - \sum_{\text{по гран. } B} (u(B)\Delta v(B) - \Delta u(B)v(B)). \quad (13.2) \end{aligned}$$

Давайте возьмем две внутренних точки  $A_1$  и  $A_2$ , и в формуле Грина положим  $u = G(A_1, \cdot)$  и  $v = G(A_2, \cdot)$ . Правая часть (13.2) обращается в нуль, потому что на границе  $G(A_1, \cdot)$  и  $G(A_2, \cdot)$  равны нулю. Во внутренних точках  $u(A) = 0$ , если  $A \neq A_1$  и  $u(A_1) = 1$ . Аналогично для  $v$  и  $A_2$ . Тогда обращение в нуль левой части (13.2) значит, что

$$G(A_1, A_2) = G(A_2, A_1).$$

Если мы напишем  $u = \Phi$ ,  $\Delta u = -\rho$  и  $v = \hat{\Phi}$ ,  $\Delta v = -\hat{\rho}$ , то уравнение (13.2) преобразуется в

$$\sum_{\text{внутр. } C} (\rho_C \hat{\Phi}^C - \hat{\rho}_C \Phi^C) = \sum_{\text{гран. } B} (\hat{\rho}_B \Phi^B - \rho_B \hat{\Phi}^B). \quad (13.3)$$

Итак, мы видим, что формула Грина есть теорема взаимности Грина для цепей, содержащих емкости.

Теперь предположим, что  $\hat{\Phi}$  является решением задачи Дирихле для заданных значений на границе, так что  $\hat{\rho} = 0$  внутри. В формуле (13.3) возьмем  $\Phi = G(A, \cdot)$ . Тогда мы получим

$$\begin{aligned} \sum_{\text{внутр. } C} (-\Delta G(A, C) \hat{\Phi}^C - \hat{\rho}_C G(A, C)) \\ = \sum_{\text{гран. } B} (\hat{\rho}_B G(A, B) + \Delta G(A, B) \hat{\Phi}^B). \end{aligned}$$

Вспомним, что  $-\Delta G(A, C) = 1$ , если  $C = A$  и нулю во всех остальных случаях, потому что  $G(A, C)$  есть решение уравнения Пуассона с единичным зарядом в узле  $A$  и нулевыми зарядами на других внутренних узлах. Кроме того, мы предположили, что  $\hat{\rho}_C = 0$  на внутренних узлах. Поэтому в левой части уравнения выживает только один член  $\hat{\Phi}^A$ . В правой части мы имеем  $G(A, B) = 0$ . Следовательно,

$$\hat{\Phi}^A = \sum_{\text{гран. } B} \Delta G(A, B) \hat{\Phi}^B.$$

Таким образом, значение потенциала  $\hat{\Phi}^A$  на внутреннем узле  $A$  равно среднему значению потенциалов  $\hat{\Phi}^B$  на граничных узлах с весовым множителем  $\Delta G(A, B)$ , который является значением отрицательного заряда, индуцированного на граничном узле  $B$ , если в узле  $A$  помещен единичный положительный заряд, а потенциал на всех границах равен нулю. Матрица  $(\Delta G(A, B))$ , представляющая линейное преобразование от  $C_{\text{гран}}^0$  к  $C_{\text{внутр}}^0$ , называется *ядром Пуассона*.

Давайте применим функцию Грина к расчету цепи, изображенной на рис. 13.18. Поскольку в цепи есть только один внутренний узел  $D$ , нам необходимо решить уравнение Пуассона для

случая  $\rho_D = 1$ . Мы уже это сделали и получили вектор напряжений

$$\mathbf{V} = \frac{1}{4} \begin{pmatrix} 1 \\ -1 \\ 0 \\ -1 \end{pmatrix}.$$

Поэтому, если узлы  $A$ ,  $B$  и  $C$  заземлены, то потенциал на узле  $D$  равен  $G(D, D) = \frac{1}{4}$ . Для этой цепи заряды на граничных узлах равны:  $\rho_A = -\frac{1}{4}$ ,  $\rho_B = -\frac{1}{2}$ ,  $\rho_C = -\frac{1}{4}$ . Тогда мы имеем  $\Delta G(D, A) = \frac{1}{4}$ ,  $\Delta G(D, B) = \frac{1}{2}$  и  $\Delta G(D, C) = \frac{1}{4}$ . С помощью функции Грина мы легко решаем задачу Дирихле для этой цепи. Если  $\rho_D = 0$ , то  $\Phi^D$  является усредненным значением граничных потенциалов с весовыми множителями  $\Delta G$ :

$$\Phi^D = \Delta G(D, A)\Phi^A + \Delta G(D, B)\Phi^B + \Delta G(D, C)\Phi^C$$

или

$$\Phi^D = \frac{1}{4}\Phi^A + \frac{1}{2}\Phi^B + \frac{1}{4}\Phi^C.$$

Существует другой вариант формулы Грина (называемый второй формулой Грина), который иногда бывает полезен. Возьмем граничный узел  $A$ . В формуле

$$\Delta u(A) = \sum_{\pm \partial \alpha = (B-A)}^{\alpha} C_{\alpha}(u(A) - u(B))$$

сумму справа разделим на две части: ветвь  $\alpha$  соединяет две граничные точки или ветвь соединяет граничную точку с внутренней. Тогда получим

$$\begin{aligned} \Delta u(A) = & \sum_{\substack{\pm \partial \alpha = (B-A) \\ B - \text{внутр. узел}}}^{\alpha} C_{\alpha}(u(A) - u(B)) \\ & + \sum_{\substack{\pm \partial \alpha = (B-A) \\ B - \text{гран. узел}}}^{\alpha} C_{\alpha}(u(A) - u(B)). \end{aligned}$$

Вторую сумму обозначим  $\Delta^{\text{гран}} u(A)$ . Ветвь, соединяющую две граничные точки, будем считать отдельной цепью. Тогда  $\Delta^{\text{гран}}$  будет оператором Лапласа для этой подцепи, определяемым емкостями ветвей, соединяющих граничные узлы. Следовательно,

$$\sum_{\text{гран. } A} (u(A)\Delta^{\text{гран}} v(A) - \Delta^{\text{гран}} u(A)v(A)) = 0$$

и поэтому

$$\begin{aligned} & \sum_{\text{гран. } A} (u(A)\Delta v(A) - \Delta u(A)v(A)) \\ &= \sum_{\substack{\text{гран. } A \\ \partial\alpha=\pm(B-A) \\ \text{внутр. } B}} (u(A)C_{\alpha}(v(A) - v(B)) - C_{\alpha}(u(A) - u(B))v(A)). \end{aligned}$$

Члены, содержащие  $u(A)v(A)$  справа, сокращаются, правая часть равенства упрощается и принимает вид

$$\sum_{\substack{\text{гран. } A \\ \partial\alpha=\pm(B-A) \\ \text{внутр. } B}} C_{\alpha}(u(B)v(A) - u(A)v(B)).$$

Подставляя это в формулу Грина, мы получаем *вторую формулу Грина*:

$$\boxed{\begin{aligned} & \sum_{\text{внутр. } A} (u(A)\Delta v(A) - \Delta u(A)v(A)) \\ &= \sum_{\substack{\text{гран. } A \\ \partial\alpha=\pm(B-A) \\ \text{внутр. } B}} C_{\alpha}(u(A)v(B) - u(B)v(A)). \end{aligned}}$$

## 13.8. Ядро Пуассона и случайное блуждание

Продолжим изучение задачи Дирихле. Пусть  $Q = \Delta G$  — ядро Пуассона. Тогда если  $\phi \in C^0_{\text{гран}}$  определяет конкретный граничный потенциал, то  $u$ , определяемое формулой

$$u(A) = \sum_{\text{гран. } B} Q(A, B)\phi(B),$$

дает решение задачи Дирихле с краевым условием  $\phi$ . В предыдущем параграфе мы видели, как найти  $G$  и, следовательно,  $Q$ , используя метод ортогональных проекций. В этом параграфе мы дадим вероятностное построение  $Q$ . И хотя этот метод обычно не применяется в конкретных расчетах, он дает нам как новый взгляд на проблему, так и итерационную процедуру вычисления матрицы  $Q$ .

Сначала вспомним еще раз, в чем состоит задача Дирихле. Пусть  $A$  — любой внутренний узел. Пусть  $B$  — ближайший сосед узла  $A$ , т. е. узел  $B$  является началом или концом некоторой ветви  $\beta$ , соединяющей его с узлом  $A$ . (Узел  $B$  может быть как внутренним, так и граничным.) Пусть

$$P_{B,A} = C_\beta \left( \sum_{\partial\alpha=\pm(B-A)} C_\alpha \right)^{-1}.$$

Тогда условие

$$(\Delta u)(A) = 0$$

совпадает с условием

$$u(A) = \sum u(B)P_{B,A},$$

где сумма берется по всем ближайшим соседям. Заметим, что по построению

$$P_{A,B} > 0, \quad \sum_B P_{B,A} = 1.$$

Положим

$$P_{A,A} = 1, \quad P_{A',A} = 0, \quad A' \neq A$$

для всех граничных узлов. Тогда пусть  $P$  обозначает матрицу  $n \times n$  с матричными элементами  $P_{A,B}$ . Чтобы определиться с обозначениями, будем сначала писать граничные узлы, так чтобы верхний левый угол матрицы  $P$  был единичной матрицей  $k \times k$ , где  $k$  есть число граничных узлов. Столбец, соответствующий внутреннему узлу, будет иметь элемент  $P_{B,A}$  на узле  $B$ , если  $B$  — ближайший сосед узла  $A$ , и нулю во всех остальных случаях. Условие

$$\Delta u = 0 \quad \text{на внутреннем узле } A$$

превращается в

$$u = uP,$$

если рассматривать  $u = (u(B_1), u(B_2), \dots)$  как вектор-строку. Если первые  $k$  компонент вектора  $u$  задаются  $\phi$ , то таковы же будут и для первые  $k$  компонент  $uP$ , поскольку первые  $k$  столбцов  $P$  имеют 1 на  $k$ -м месте и 0 на остальных. Итак, задача Дирихле может быть сформулирована следующим образом:

$$uP = u, \quad \text{и первые } k \text{ элементов } u \text{ определяются } \phi.$$

С другой стороны, матрица  $P$  имеет структуру матрицы *вероятностей перехода*. Действительно, предположим, что у нас есть частица, находящаяся в каком-то узле цепи в каждый дискретный момент времени  $t = 1, 2, \dots$ . Допустим, что она двигается согласно следующему правилу. Если в какой-то момент времени  $t$  она находится в граничном узле, то она там и остается. А если она находится на внутреннем узле, то должна перескочить на какой-либо соседний узел, при этом вероятность ее перехода на узел  $B$  равна  $P_{B,A}$ .

Матрица  $P^N$  представляет вероятность перехода после  $N$  шагов. Очевидно, что предельное поведение матрицы  $P^N$  имеет вероятностную интерпретацию. Начав с любого внутреннего узла, после достаточно большого числа шагов получится ненулевая вероятность частице оказаться на граничном узле, и с вероятностью меньше единицы она останется во внутренней части цепи. Последняя вероятность при некотором  $N$  допускает единую оценку сверху для всех начальных внутренних узлов, поэтому с дальнейшим ростом числа шагов вероятность частице остаться

на внутренних узлах стремится к нулю. Таким образом, предельное выражение матрицы  $P^N$  принимает вид

$$H = \begin{pmatrix} \mathbb{I} & Q \\ 0 & 0 \end{pmatrix},$$

где  $\mathbb{I}$  — единичная матрица  $k \times k$ , а  $Q$  — матрица, у которой  $k$  строк и  $n - k$  столбцов, причем  $n - k$  есть число внутренних узлов. Матричный элемент  $Q(A, B)$  представляет вероятность того, что частица, начав двигаться из точки  $A$ , в конце концов придет в граничную точку  $B$ . (Путь частицы рано или поздно закончится в некой граничной точке.) Тогда  $P^N \rightarrow H$  и, следовательно,  $P^{N+1} = P^N P \rightarrow H$ , откуда  $HP = H$ . Поэтому если  $\mathbf{v}$  — какой-то потенциал, то для  $\mathbf{v}H$  выполняется условие

$$(\mathbf{v}H)P = \mathbf{v}H$$

и, следовательно,

$$\mathbf{u} = \mathbf{v}H$$

является решением задачи Дирихле, когда краевыми условиями будут первые  $k$  компонент вектора  $\mathbf{v}$ . Если первые  $k$  компонент задаются  $\phi$ , т. е. если  $\mathbf{v}$  выбирается в виде

$$\mathbf{v} = (\phi, \mathbf{w}), \quad \mathbf{w} \text{ — любой из } n - k \text{ векторов-строк,}$$

то  $\mathbf{v}H$  является решением задачи Дирихле. Из вида матрицы  $H$  очевидно, что для любого внутреннего узла  $A$  выполняется условие

$$\mathbf{u}(A) = (\mathbf{v}H)(A) = \sum_{\text{гран. } B} \phi(A)Q(A, B).$$

Следовательно,  $Q$  действительно является ядром Пуассона.

Матрице  $H$ , т. е. предельному виду матрицы  $P^N$ , можно дать интересную геометрическую интерпретацию. Матричный элемент  $P^N(A_2, A_1)$  определяет вероятность перехода из  $A_1$  в  $A_2$  после  $N$  шагов. Эта вероятность равна сумме вероятностей всех  $N$ -шаговых траекторий, ведущих из  $A_1$  в  $A_2$ . (Если  $A_2$  — граничный узел, то некоторые из этих траекторий фактически имеют меньшую длину, т. е. в течение последних  $r$  шагов частица стоит

в точке  $A_2$ .) Вероятность любой траектории равна произведению  $P_{B,A}$  по всем ветвям  $\alpha$ , проходимым вдоль траектории. Если  $A$  — внутренний узел, а  $B$  — внешний, то очевидно, что с ростом  $N$  при вычислении  $P^N(B, A)$  мы добавляем все более длинные пути. Тогда

$$Q(B, A) = \sum_{\substack{\text{по всем путям,} \\ \text{соединяющим} \\ A \text{ с } B}} (\text{вероятность пути}).$$

### 13.9. Теорема взаимности Грина в электростатике

В качестве небольшого обобщения цепей с емкостями можно рассмотреть систему заряженных проводников, каждый из которых имеет определенный заряд  $\rho$  и определенный потенциал  $\phi$  (рис. 13.21). Чтобы описать аналог ветви для системы такого типа, требуется введение поля диэлектрического смещения и поэтому должно быть отложено до главы 16, в которой мы будем изучать электростатику. Однако, описание цепи через заряды и потенциалы имеет много общего с описанием емкостных цепей.

Полный заряд на каждом проводнике может описываться вектором  $\rho = \begin{pmatrix} \rho_A \\ \rho_B \\ \vdots \end{pmatrix}$  в пространстве, обозначаемом  $C_0$ ; потенциал

же образует вектор  $\phi = (\phi^A, \phi^B, \dots)$  в его дуальном пространстве  $C^0$ . Для цепей с емкостями скалярное произведение элемента пространства  $C^0$  на элемент пространства  $C_0$  дает электростатическую энергию системы проводников, т.е. энергию, равную  $\frac{1}{2} \int \rho \phi = \frac{1}{2} \sum \rho_A \phi^A$ . (В главе 16 мы увидим, что это выражение равно интегралу

$$\int (d\phi, d\phi) dx dy dz$$

по пространству между проводниками, где  $\phi$  обозначает электростатический потенциал. А сейчас мы просто ограничиваемся описанием того, что происходит в проводниках.)

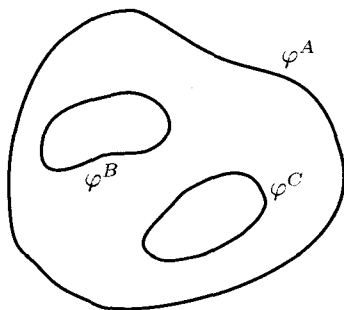


Рис. 13.21

Полный заряд проводника  $\rho$  может быть выражен через потенциал  $\phi$  с помощью оператора Лапласа  $\Delta$ , т. е.

$$\rho = -\Delta\phi.$$

Оператор  $\Delta$  зависит от формы различных проводников, от их распределения в пространстве и физических констант электростатики. Кроме ряда частных случаев, когда в системе есть параллельные плоскости, концентрические сферы или цилиндры, вычисление  $\Delta$  очень трудно. Однако, как мы узнаем в главе 16, для системы проводников выполняется теорема взаимности Грина, точно так же, как она выполняется для цепей с емкостями. Говоря точнее, если  $\rho$  и  $\phi$  обозначают возможное распределение полного заряда и напряжений в системе проводников, а  $\rho'$  и  $\phi'$  обозначают другое распределение, то теорема взаимности Грина утверждает, что

$$\int_{\rho'} \phi = \int_{\rho} \phi'.$$

Отсюда немедленно следует, что оператор  $\Delta$  самосопряжен и представляется симметричной матрицей. Чтобы это увидеть, положим

$$\rho = -\Delta\phi, \quad \rho' = -\Delta\phi'.$$

Тогда

$$\int_{-\Delta\phi'} \phi = \int_{-\Delta\phi} \phi',$$

т. е.

$$(\Delta\phi', \phi) = (\phi', \Delta\phi),$$

откуда следует, что  $\Delta$  самосопряжен. В физической литературе оператор  $\Delta$  обычно называют *матрицей емкостных коэффициентов*.

Матрица, обратная к  $-\Delta$ , называется матрицей потенциальных коэффициентов. Мы напишем эту матрицу для ряда простых частных случаев через хорошо известные формулы элементарной электростатики. Например, рассмотрим систему двух концентрических сфер (рис. 13.22) с радиусами  $r_A$  и  $r_B$ . Для сферы с радиусом  $r_0$ , обладающей единичным положительным зарядом, электрический потенциал, выраженный в гауссовых единицах, имеет вид

$$\phi(r) = \begin{cases} 1/r_0 & (r \leq r_0), \\ 1/r & (r > r_0). \end{cases}$$

(См., например книгу Purcell «Электричество и магнетизм», параграф 1.11.) Позднее в параграфах 16.1 и 16.2 мы выведем этот результат альтернативным путем из первых принципов.

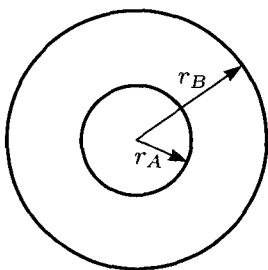


Рис. 13.22

Если  $\rho_A = 1$ ,  $\rho_B = 0$ , то потенциалы  $\phi^A = 1/r_A$ ,  $\phi^B = 1/r_B$ . Однако если  $\rho_A = 0$ ,  $\rho_B = 1$ , то оба потенциала равны  $\phi^A = \phi^B = 1/r_B$ . Отсюда следует, что

$$-\Delta^{-1} = \begin{pmatrix} 1/r_A & 1/r_B \\ 1/r_B & 1/r_B \end{pmatrix}.$$

Эта матрица позволяет нам вычислить потенциал двух сфер для произвольного распределения зарядов. Обратная матрица дает

оператор Лапласа

$$-\Delta = \left( \frac{1}{r_B} \frac{1}{r_A} - \frac{1}{r_B^2} \right)^{-1} \begin{pmatrix} 1/r_B & -1/r_B \\ -1/r_B & 1/r_A \end{pmatrix}$$

или

$$-\Delta = \frac{r_A r_B}{r_B - r_A} \begin{pmatrix} 1 & -1 \\ -1 & r_B/r_A \end{pmatrix} = \frac{r_B}{r_B - r_A} \begin{pmatrix} r_A & -r_A \\ -r_A & r_B \end{pmatrix}.$$

Эта матрица определяет заряды на двух сферах при заданных потенциалах. В случае  $\phi_A = 1$ ,  $\phi_B = 0$  мы имеем

$$\begin{pmatrix} \rho_A \\ \rho_B \end{pmatrix} = \frac{r_A r_B}{r_B - r_A} \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

т. е. на двух сферах есть два заряда с противоположными знаками, равные  $r_A r_B / (r_B - r_A)$ . Эта величина  $r_A r_B / (r_B - r_A)$  называется емкостью двух сфер.

Матрицу  $-\Delta^{-1}$  одинаково легко написать для любого числа концентрических сфер. Например, для трех сфер, изображенных на рис. 13.23, она равна

$$-\Delta^{-1} = \begin{pmatrix} 1/r_A & 1/r_B & 1/r_C \\ 1/r_B & 1/r_B & 1/r_C \\ 1/r_C & 1/r_C & 1/r_C \end{pmatrix}.$$

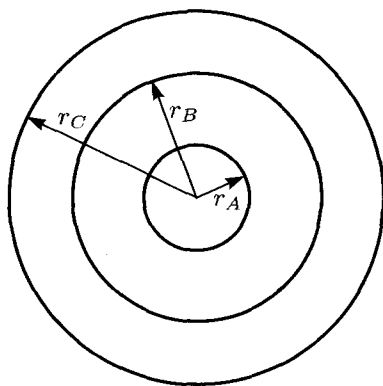


Рис. 13.23

А сейчас представим себе цепь с емкостями и пусть некоторые проводники будут *граничными проводниками*, потенциалы которых определяются подсоединенными батареями, а на других проводниках, называемых *внутренними*, задается заряд. В задаче Пуассона мы имеем потенциалы  $\phi' = 0$  на всех граничных проводниках и заряды  $\rho'$  на всех внутренних. Требуется найти  $\phi'$  для всех внутренних проводников и  $\rho'$  для всех граничных. А в задаче Дирихле нам дано, что  $\rho = 0$  на внутренних проводниках и  $\phi$  на всех граничных. И мы должны найти  $\phi$  для внутренних проводников и  $\rho$  для граничных.

Эти две задачи связаны между собой теоремой взаимности Грина. В общем случае

$$\sum_{\text{гран. } A} \rho'_A \phi^A + \sum_{\text{внутр. } B} \rho'_B \phi^B = \sum_{\text{гран. } A} \rho_A \phi'^A + \sum_{\text{внутр. } B} \rho_B \phi'^B.$$

Нам дано, что  $\phi' = 0$  на граничных проводниках (задача Пуассона) и  $\rho = 0$  на внутренних (задача Дирихле). Тогда получаем

$$\sum_{\text{внутр. } B} \rho'_B \phi^B = \sum_{\text{гран. } A} \rho'_A \phi^A.$$

Допустим, что в задаче Пуассона  $Q(C, A)$  обозначает заряд  $\rho'_A$  на граничном проводнике  $A$ , имеется единичный заряд на внутреннем проводнике  $C$ , а заряды всех остальных внутренних проводников равны нулю. Тогда слева выживает только один член, и мы имеем

$$\phi^C = - \sum Q(C, A) \phi^A, \quad \text{где } Q(C, A) = -\Delta G(C, A).$$

Эта функция Грина является решением задачи Дирихле.

На практике довольно часто для системы проводников можно легко решить задачу Дирихле, а затем с помощью теоремы взаимности решить уравнение Пуассона. Рассмотрим, например, систему из трех больших параллельных проводящих плоскостей, показанную на рис. 13.24. На этой схеме  $A$  и  $C$  — граничные проводники, а  $B$  — внутренний. В этой геометрии потенциал является линейной функцией положения между плоскостями. Следовательно,

$$\phi^B = \phi^A + \frac{1}{3}(\phi^C - \phi^A)$$

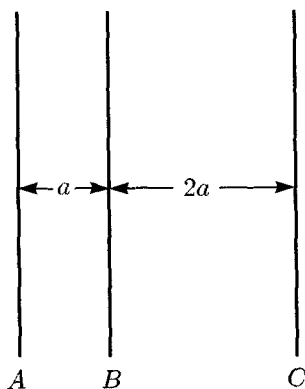


Рис. 13.24

или

$$\phi^B = \frac{2}{3}\phi^A + \frac{1}{3}\phi^C.$$

Это и есть решение задачи Дирихле.

А теперь рассмотрим уравнение Пуассона с зарядом  $\rho'_B$  на внутренней плоскости. При этом  $\phi'^A = \phi'^C = 0$ . Из теоремы взаимности следует, что

$$\rho'_B \phi^B = -\rho'_A \phi^A - \rho'_C \phi^C.$$

Однако для произвольной задачи Дирихле  $\phi^B = \frac{2}{3}\phi^A + \frac{1}{3}\phi^C$ , следовательно,

$$\rho'_B \cdot \frac{2}{3}\phi^A + \rho'_B \cdot \frac{1}{3}\phi^C = -\rho'_A \phi^A - \rho'_C \phi^C.$$

Отсюда следует, что

$$\rho'_A = -\frac{2}{3}\rho'_B, \quad \rho'_C = -\frac{1}{3}\rho'_B,$$

т.е. мы выяснили, как индуцированный заряд распределяется между двумя граничными плоскостями.

В общем случае теорема взаимности Грина может применяться в задачах, где имеются как проводники, так и распределения зарядов в области, ограниченной этими проводниками. Например, если заряд  $+q$  помещен вблизи бесконечной заряженной

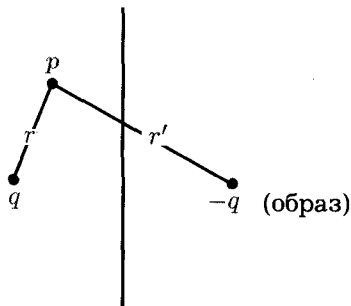


Рис. 13.25

плоскости (рис. 13.25), то электрический потенциал везде на той же стороне плоскости, где находится заряд, может быть определен *методом изображений*, т. е.

$$\phi(P) = \frac{q}{r} - \frac{q}{r'}$$

(см. параграф 16.1). Для этой потенциальной функции можно получить распределение отрицательного заряда вдоль проводящей плоскости. А это, в свою очередь, дает возможность решить задачу Дирихле, когда везде на плоскости задан потенциал и требуется найти распределение потенциала во всем пространстве. Аналогичным способом строится функция Грина для решения уравнения Пуассона, когда точечный заряд находится внутри заземленной проводящей сферы. Потом эта функция используется для решения задачи Дирихле, когда на поверхности сферы задан потенциал и требуется найти потенциал в области, ограниченной сферой и свободной от зарядов.

## Резюме

### А. Ортогональные проекции

Вы должны знать свойства самосопряженного проекционного оператора  $\pi$  для цепей с резисторами и использовать его для расчета таких цепей.

Вы должны уметь строить оператор  $\pi$ , пользуясь процедурой Грама–Шмидта или методом Кирхгофа.

### В. Цепи с емкостями

Вы должны уметь рассчитывать цепи с емкостями с помощью методов Максвелла или обращения оператора Лапласа.

Для расчета цепей с емкостями Вы должны уметь решать уравнение Пуассона или задачу Дирихле с помощью ортогональных проекций; должны описывать связанное с этим разложение  $C^1$  на ортогональные подпространства.

### С. Функции Грина

Вы должны уметь строить функцию Грина для цепи с емкостями и использовать ее для решения уравнения Пуассона и задачи Дирихле.

## Задачи

13.1. Рассмотрите цепь, изображенную на рис. 13.26.

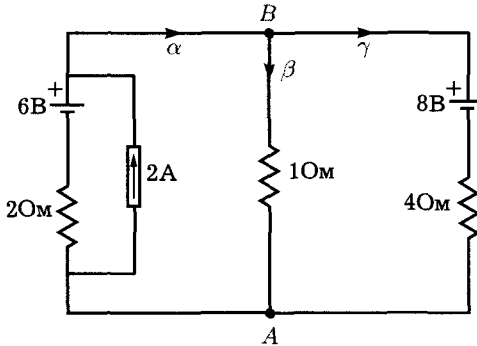


Рис. 13.26

- (а) Найдите базис пространства  $Z_1$ . Объясните смысл утверждения «оператор  $\pi$  представляет проекцию  $C_1$  на подпространство  $Z_1$ ». Выполнив соответствующие вычисления, покажите, что матрица

$$\pi = \frac{1}{7} \begin{pmatrix} 5 & 2 & 2 \\ 4 & 3 & -4 \\ 1 & -1 & 6 \end{pmatrix}$$

представляет проекцию  $C_1$  на  $Z_1$ .

- (b) Для этой цепи найдите токи в ветвях, используя матрицу  $\pi$ . Сделайте рисунок, показывающий образ и ядро  $\pi$ , покажите, как с его помощью решить эту задачу.
- (c) Какое линейное преобразование представляется матрицей  $\pi^T$ , транспонированной к  $\pi$ ? Назовите и дайте характеристику с физической точки зрения пространству, являющемуся ядром этого преобразования. Обоснуйте свой ответ либо с помощью общей теоремы, либо ссылаясь на соответствующие свойства матрицы  $\pi^T$ .
- 13.2. (a) Постройте матрицу  $2 \times 2$ , представляющую  $\pi : C_1 \rightarrow Z$ , для цепи, изображенной на рис. 13.27.

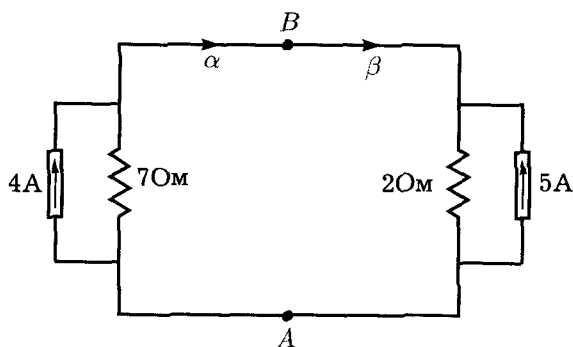


Рис. 13.27

- (b) Сделайте рисунок, показывающий образ и ядро  $\pi$ . На нем покажите вектора  $\mathbf{K}$  и  $\mathbf{I} = \pi \mathbf{K}$ .
- 13.3. Рассмотрите цепь, изображенную на рис. 13.28.
- (a) Используя процедуру Грама–Шмидта, преобразуйте базис

$$\mathbf{M}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{M}_2 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

в ортонормальный базис  $\{\mathbf{E}_1, \mathbf{E}_2\}$  пространства  $Z_1$ . Применяя операцию  $\pi \mathbf{I} = (\mathbf{E}_1, \mathbf{I})_Z \mathbf{E}_1 + (\mathbf{E}_2, \mathbf{I})_Z \mathbf{E}_2$  по очереди к каждому из четырех базисных векторов  $C_1$ , постройте матрицу, представляющую  $\pi$ , и рассчитайте данную цепь.

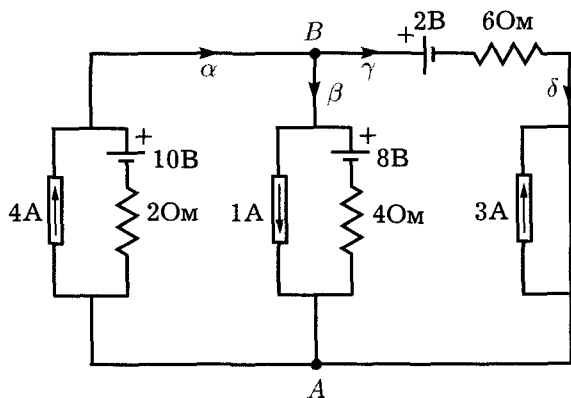


Рис. 13.28

- (b) Для этой цепи докажите формулу  $\pi = \sigma(sZ\sigma)^{-1}sZ$ .  
 (c) Постройте матрицу  $Z(1 - \pi)$  и покажите, что ее образ есть пространство  $B^1$ .

Ответ:  $\pi = \frac{1}{50} \begin{pmatrix} 22 & 28 & 24 & 4 \\ 14 & 36 & -12 & -2 \\ 8 & -8 & 36 & 6 \\ 8 & -8 & 36 & 6 \end{pmatrix}.$

- 13.4. (a) Для цепи, изображенной на рис. 13.28, постройте матрицу, представляющую

$$\tau : C^1 \rightarrow C^1,$$

которая проектирует пространство  $C^1$  на  $B^1$ , причем проектирует ортогонально по отношению к скалярному произведению, определенному преобразованием  $Z^{-1}$ , т. е.

$$(\mathbf{V}, \mathbf{V}')_{Z^{-1}} = \int_{Z^{-1}\mathbf{V}'} \mathbf{V}.$$

Чтобы построить базис для пространства  $B^1$ , положите потенциал равным нулю в узле  $A$  (заземлите его) и  $-1$  в узлах  $B$  и  $C$ .

- (b) Покажите, что  $\mathbf{V} = \tau(\mathbf{W} - Z\mathbf{K})$  и используйте эту формулу для расчета цепи, изображенной на рис. 13.28.
- 13.5. (a) Найдите явное выражение для матрицы  $\tau$  из предыдущего упражнения через  $[d]$ ,  $Z^{-1}$  и  $[\partial]$ . С помощью этой матрицы

покажите, что  $\tau^2 = \tau$ , что образ  $\tau$  есть  $B^1$ , что  $\tau$  — самосопряженная матрица по отношению к скалярному произведению, определенному  $Z^{-1}$ .

- (b) Покажите, что для любой цепи матрица

$$Z^{-1}[d]([\partial]Z^{-1}[d])^{-1}[\partial] + \sigma(sZ\sigma)^{-1}sZ,$$

представляющая линейное преобразование из  $C_1$  в  $C_1$ , является единичной матрицей.

13.6. Для цепи, изображенной на рис. 13.28, с помощью метода Кирхгофа сделайте следующее.

- (a) В цепи есть пять различных максимальных деревьев. Для каждого дерева получите матрицу  $\rho_T$  ( $4 \times 4$ ). Заметим, что каждая пара ветвей, кроме  $\{\alpha, \beta\}$ , образует дерево. (Следите за знаками — все диагональные элементы матрицы  $\rho_T$  должны быть  $\geq 0$ .)
- (b) Для каждого дерева вычислите  $Q_T$  и найдите  $R = \sum Q_T$ . (Получится  $R = 50$ , что совпадает с  $\text{Det}(sZ\sigma)$ . Будет ли это так в общем случае?)
- (c) Воспользуйтесь формулой Кирхгофа

$$\pi = \frac{1}{R} \sum_T Q_T \rho_T,$$

чтобы получить проекционный оператор  $\pi$  методом Вейля.

13.7. Придумайте изменение в методе Кирхгофа, чтобы построить самосопряженный проекционный оператор  $\tau : C^1 \rightarrow C^1$ , обладающий таким же свойством  $\mathbf{V} = \tau(\mathbf{W} - Z\mathbf{K})$ , что и усредненное значение проекций  $C^1$  на  $Z^1$ , взятых с определенными весами, и не являющееся самосопряженным оператором. Один из способов сделать это — начать с метода Кирхгофа в пространстве  $C_1$  и потом обе части умножить на  $Z$ .

13.8. Докажите формулу для проекционного оператора Вейля  $\pi = \sigma(sZ\sigma)^{-1}sZ$ , проверив, что  $\pi$  обладает требуемыми свойствами:

- (a)  $\pi$  является проекцией:  $\pi^2 = \pi$ ;
- (b) образом  $\pi$  является  $Z_1$ ;
- (c)  $\pi$  — самосопряженный оператор относительно  $(\cdot, \cdot)_Z$ , т. е.

$$(\pi \mathbf{I}, \mathbf{I}')_Z = (\mathbf{I}, \pi \mathbf{I}')_Z.$$

- 13.9. (а) Для цепи, изображенной на рис. 13.29, постройте матрицу  $\pi$ , представляющую ортогональную проекцию пространства  $C_1$  на подпространство  $Z_1$ . Воспользуйтесь этим проекционным оператором для вычисления токов в ветвях  $I$ . (Все токи задаются целыми числами.)

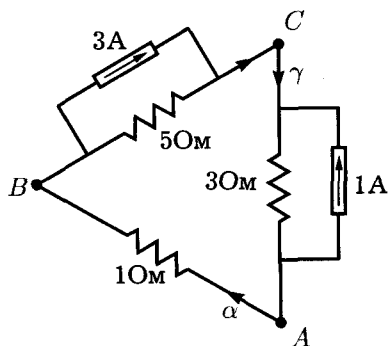


Рис. 13.29

- (b) Для этой цепи постройте базис ядра оператора  $\pi$ . Воспользуйтесь им, чтобы показать в явном виде, что ядро  $\pi$  ортогонально его образу.
- 13.10. В качестве иллюстрации теоремы взаимности Грина рассмотрите цепь, изображенную на рис. 13.30.

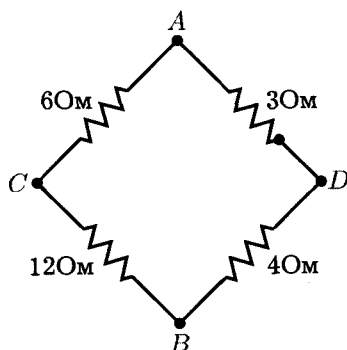


Рис. 13.30

- (а) Если между узлами  $A$  и  $B$  вставить батарею на 60 вольт, то какой ток потечет между узлами  $C$  и  $D$  при замыкании их

накоротко? Если, наоборот, батарею вставить между узлами  $C$  и  $D$  и устроить короткое замыкание  $A$  с  $B$ , то какой ток потечет через  $AB$ ?

- (b) Если источник тока на 25 ампер вставить между узлами  $A$  и  $B$ , то какое напряжение возникнет в  $CD$  (узлы разомкнуты)? Если, в свою очередь, тот же источник тока вставить между узлами  $C$  и  $D$ , то какое напряжение возникнет в  $AB$ ?
- 13.11. (a) Для цепи, изображенной на рис. 13.31, постройте симметричную матрицу  $Z\pi$ . Используйте ее для определения напряжения на ветвях  $\mathbf{V}$ , если в ветвь  $\alpha$  вставлен источник тока, дающий 8 ампер. Найдите  $\mathbf{V}$  также в случае, когда тот же источник вставлен в ветвь  $\gamma$ .
- (b) Для той же цепи постройте симметричную матрицу  $\pi Z^{-1}$ . Воспользуйтесь ей для определения токов в ветвях  $\mathbf{I}$ , если в ветвь  $\alpha$  вставлена батарея, дающая напряжение 16 вольт? Какой будет ток, если ту же батарею вставить в ветвь  $\gamma$ ?

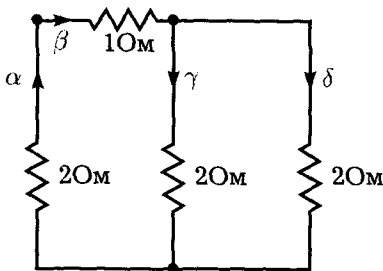


Рис. 13.31

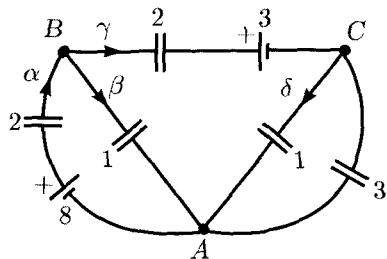


Рис. 13.32

- 13.12. Определите заряд на каждом конденсаторе цепи, изображенной на рис. 13.32, используя метод круговых токов и метод узловых потенциалов. Предположим, что  $[\partial]Q = 0$ , т.е. что до подключения батарей все конденсаторы не были заряжены. (Если нужно, определим единицы: напряжение  $V$  измеряется в вольтах, емкость  $C$  — в микрофарадах, а заряд  $Q$  — в микрокулонах.)
- 13.13. В цепи, изображенной на рис. 13.33,  $\rho_B = 13$  и  $\rho_C = -1$ , поэтому  $\rho_A = -12$ . Решите уравнение Пуассона  $\Delta\Phi = -\rho$  для этой цепи и определите  $\Phi^B$  и  $\Phi^C$ . Вычислите заряды на ветвях  $\mathbf{Q}$  и проверьте,

$$\text{что } \partial\mathbf{Q} = - \begin{pmatrix} -12 \\ 13 \\ -1 \end{pmatrix}.$$

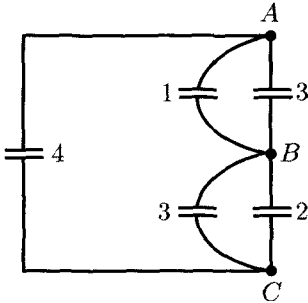


Рис. 13.33

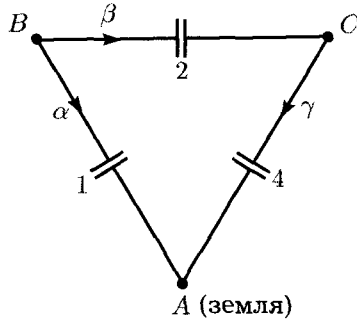


Рис. 13.34

- 13.14. Для цепи, изображенной на рис. 13.34, напишите матрицы  $d$  и  $\partial$ . Для  $\rho_B = 10$  и  $\rho_C = -2$  решите уравнение Пуассона и найдите потенциалы узлов  $\Phi^B$  и  $\Phi^C$ .
- 13.15. Рассмотрите цепь конденсаторов, изображенную на рис. 13.35. Узлы  $A$  и  $B$  — граничные, а узел  $C$  — внутренний.

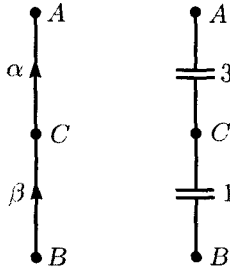


Рис. 13.35

- (а) Постройте матрицу  $2 \times 2$  оператора  $\pi$ , который является ортогональной проекцией (относительно скалярного произведения, определенного  $C$ )  $C'$  на  $dC_{\text{внутр}}^0$ .
- (б) Вектор  $\mathbf{W} = \begin{pmatrix} 0 \\ -4 \end{pmatrix}$  соответствует ситуации, когда на узле  $C$  находится заряд 4. Воспользуйтесь оператором  $\pi$  для нахождения распределения зарядов в случае, если узлы  $A$  и  $B$  заземлены.
- (с) Предположим, что потенциал узла  $A$  равен  $\Phi^A = 0$ , потенциал узла  $B$  равен  $\Phi^B = 4$  и заряд узла  $C$  равен нулю.

Воспользуйтесь оператором  $\pi$  и определите напряжения в ветвях  $V^\alpha$  и  $V^\beta$ .

- 13.16. Рассмотрите систему конденсаторов, показанную на рис. 13.36. В этой цепи узлы  $A$  и  $D$  граничные, а узлы  $B$  и  $C$  внутренние.

(a) Для этой цепи напишите матрицы  $\partial$  и  $d$ .

(b) Пусть  $\phi_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$  и  $\phi_2 = \begin{pmatrix} 0 \\ 3 \\ 1 \\ 0 \end{pmatrix}$  — векторы в пространстве

$C^0$ . Вычислите  $d\phi_1$  и  $d\phi_2$ . Покажите, что они ортогональны по отношению к скалярному произведению, определенному матрицей емкостей  $C$ .

(c) Пусть  $\pi$  представляет ортогональную проекцию из  $C^1$  на подпространство, заданное векторами  $d\phi_1$  и  $d\phi_2$ . Вычислите  $\pi \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ .

(d) Пусть на узле  $B$  находится единичный заряд, на узле  $C$  заряд равен нулю, а узлы  $A$  и  $D$  заземлены. Воспользуйтесь ортогональной проекцией и вычислите потенциалы узлов  $B$  и  $C$ . Найдите заряды на узлах  $A$  и  $D$ .

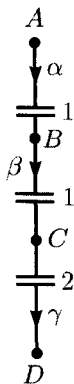


Рис. 13.36

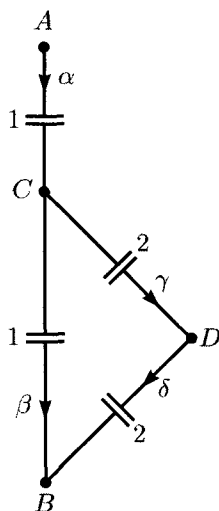


Рис. 13.37

- (е) Пусть  $\Phi^A = 5$ ,  $\Phi^D = 0$ , на узлах  $B$  и  $C$  нет зарядов. Воспользуйтесь методом ортогональных проекций и определите напряжения на ветвях  $\mathbf{V}$ .
- (f) Для задачи, поставленной в пункте (е), покажите, как найти  $\Phi^B$ , используя функцию Грина из пункта (d).

13.17. Рассмотрите цепь конденсаторов, показанную на рис. 13.37. Узлы  $A$  и  $B$  — граничные, а узлы  $C$  и  $D$  — внутренние.

- (a) Вектора

$$\psi_1 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{и} \quad \psi_2 = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

являются элементами пространства  $C^0$ . Назовите, какое пространство они задают, и дайте ему физическую характеристику. Определите  $d\psi_1$  и  $d\psi_2$ . Покажите, что они образуют ортогональный базис подпространства  $C^1$  относительно скалярного произведения, определенного  $C$ .

- (b) Пусть  $\pi$  обозначает ортогональную проекцию пространства  $C^1$  на подпространство, для которого только что получен базис. С физической точки зрения дайте характеристику векторам, которые принадлежат ядру оператора  $\pi$ . Вычис-

лите  $\pi \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$ .

- (с) Пусть на узле  $D$  находится единичный положительный заряд, на узле  $C$  заряда нет, а узлы  $A$  и  $B$  заземлены. Используя результат пункта (b), найдите заряды и напряжения всех ветвей.
- (d) Пусть на узле  $D$  находится заряд 2, на узле  $C$  заряда нет,  $\Phi^A = 4$  и  $\Phi^B = 1$ . Используя только что полученную функцию Грина, определите потенциал узла  $D$ . (Указание: эта задача является суперпозицией задачи Дирихле и задачи Пуассона.)

13.18. Рассмотрите цепь, изображенную на рис. 13.38. Узлы  $A$ ,  $B$  и  $C$  — граничные, а узлы  $D$  и  $E$  — внутренние.

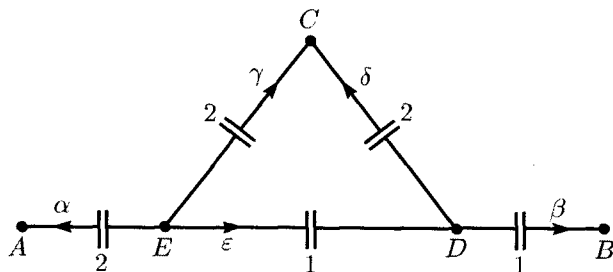


Рис. 13.38

- (a) Получите базис подпространства  $C^0_{\text{внутр}} \subset C^0$  и  $dC^0_{\text{внутр}} \subset C^1$ . Проверьте, что матрица

$$\pi = \frac{1}{19} \begin{pmatrix} 8 & 1 & 8 & 2 & 3 \\ 2 & 5 & 2 & 10 & -4 \\ 8 & 1 & 8 & 2 & 3 \\ 2 & 5 & 2 & 10 & -4 \\ 6 & -4 & 6 & -8 & 7 \end{pmatrix}$$

проектирует  $C^1$  на это подпространство.

- (b) Постройте матрицу  $1 - \pi$  и проверьте, что ее образ ортогонален образу  $\pi$ .
- (c) Проверьте, что  $H^1$  (образ  $Z_1$  при действии  $C^{-1}$ ) является ядром  $\pi$ .
- (d) Постройте матрицу  $(1 - \pi)d$ , представляющую преобразование из  $C^0$  в  $C^1$ . Проверьте в явном виде, что если  $\hat{\Phi}$  — произвольный базисный вектор в  $C^0_{\text{гран}}$ , то  $V = -(1 - \pi)d\hat{\Phi}$  описывает множество напряжений ветвей, для которых  $\partial CV = 0$  на внутренних узлах  $D$  и  $E$ .
- 13.19. Для цепи, изображенной на рис. 13.38, используйте матрицу  $(1 - \pi)d$  и решите задачу Дирихле для потенциалов  $\Phi^A = 3$ ,  $\Phi^B = -5$ ,  $\Phi^C = 0$  при  $\rho_D = \rho_E = 0$ . Получите  $\Phi^D$ ,  $\Phi^E$ ,  $\rho_A$ ,  $\rho_B$  и  $\rho_C$ .
- 13.20. В предыдущем упражнении решите уравнение Пуассона для случая, когда

$$\rho_E = 1, \quad \rho_D = 0, \quad \Phi^A = \Phi^B = \Phi^C = 0.$$

- (a) Выберите любое значение  $\hat{Q}$ , удовлетворяющее условиям  $\rho_D = 0$ ,  $\rho_E = 1$ . Постройте вектор напряжений в ветвях  $V = C^{-1}\hat{Q}$ .

- (b) Вычислите  $\mathbf{V} = \pi \hat{\mathbf{V}}$  и определите такой потенциал, чтобы на границе  $\Phi = 0$  и  $-d\Phi = \mathbf{V}$ .
- (c) Повторите пункты (a) и (b) с другими значениями  $\hat{Q}$ . Вы должны получить тот же самый ответ.
- (d) Решите уравнение Пуассона для случая  $\rho_D = 1$ ,  $\rho_E = 0$ . Для этой цепи уже построена функция Грина. Воспользуйтесь ей и решите уравнение Пуассона для  $\rho_D = 10$  и  $\rho_E = 7$ .
- (e) Используя функцию Грина, постройте ядро Пуассона  $\Delta G(A, B)$ . Это будет матрица с тремя столбцами и двумя строками, которая представляет отображение

$$\begin{pmatrix} \Phi^A \\ \Phi^B \\ \Phi^C \end{pmatrix} \quad \text{в} \quad \begin{pmatrix} \Phi^D \\ \Phi^E \end{pmatrix}.$$

Транспонированная к ней матрица отображает

$$\begin{pmatrix} \rho_D \\ \rho_E \end{pmatrix} \quad \text{в} \quad \begin{pmatrix} \rho_A \\ \rho_B \\ \rho_C \end{pmatrix}.$$

- (f) Проверьте, что ядро Пуассона дает правильное решение задачи Дирихле, поставленной в упражнении 13.18(c).
- 13.21. Конденсатор состоит из трех длинных коаксиальных цилиндров с длиной  $l$  и радиусами  $r_A = a$ ,  $r_B = 2a$ ,  $r_C = 8a$ .
- (a) Решите задачу Дирихле, когда цилиндр  $B$  не заряжен, а разница потенциалов между цилиндрами  $A$  и  $C$  равна  $\Phi^0$ .
- (b) Используя теорему взаимности Грина, решите уравнение Пуассона, если цилиндры  $A$  и  $C$  заземлены, а заряд  $Q$  находится на цилиндре  $B$ .
- 13.22. Три концентрических проводящих сферы имеют радиусы  $r_A = 2$ ,  $r_B = 3$ ,  $r_C = 4$  соответственно.
- (a) Для этой системы напишите матрицу  $-\Delta^{-1}$ , обратите ее и постройте оператор Лапласа  $-\Delta$ .
- (b) Пусть потенциалы равны  $\phi^A = 0$ ,  $\phi^B = 1$ ,  $\phi^C = 0$ . Определите заряд на каждой сфере.
- 13.23. Четыре больших проводящих плоскости (площади  $A_0$  каждая) находятся на расстоянии  $l$  друг от друга (рис. 13.39).

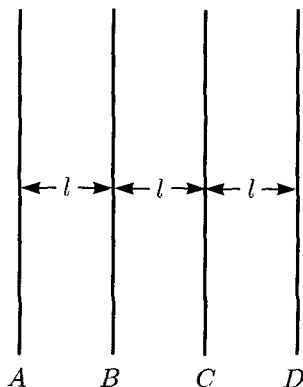


Рис. 13.39

- (a) Решите задачу Дирихле, когда заданы потенциалы  $\phi^A, \phi^D$ , а плоскости  $B$  и  $C$  не заряжены.
- (b) Используя теорему взаимности Грина, решите уравнение Пуассона для задачи, когда плоскости  $A$  и  $D$  заземлены, а на внутренних плоскостях находятся заряды  $Q_B$  и  $Q_C$ . Определите потенциалы  $\phi^B, \phi^C$  и заряды  $Q_A, Q_D$ .

## Глава 14

---

# Комплексы высших размерностей

---

В этой главе мы закончим введение в алгебраическую топологию. Вкратце опишем, как одномерные результаты, полученные в главах 12 и 13, обобщаются на случай большего числа измерений.

### 14.1. Комплексы и гомологии

В этой главе мы прервем изучение электрических цепей и рассмотрим ряд обобщений тех математических методов, которыми мы занимались в предыдущих главах. Мы определили одномерный комплекс как систему, состоящую из объектов нулевой размерности, названных узлами, и одномерных объектов, названных ветвями. Ветви соединяются между собой в узлах. Такие комплексы представляют особый интерес с точки зрения связности, что нашло свое отражение в пространстве  $H_0$ . Важным также является число независимых ячеек, которое равно размерности пространства  $H_1$ . В этой главе мы будем рассматривать комплексы с большим числом измерений, т. е. содержащие двумерные, трехмерные и т. д. элементы. В двумерном комплексе или *2-комплексе*, помимо узлов и ветвей, имеются двумерные элементы (куски поверхности), которые мы будем называть *2-ячейками* или *гранями*. На рис. 14.1 показано, что эти грани соприкасаются друг с другом вдоль ветвей. Граница грани состоит из нескольких ветвей, причем каждая входит со знаком плюс или минус.

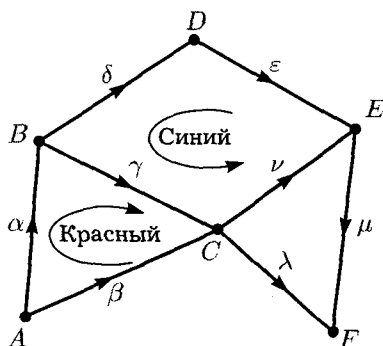


Рис. 14.1

Трехмерный комплекс содержит трехмерные элементы (объемные куски), называемые 3-ячейками, которые касаются друг друга вдоль граней. Граница 3-ячейки состоит из нескольких граней, причем каждая входит со знаком плюс или минус.

В общем случае  $n$ -комплекс состоит из конечного множества  $n + 1$  элементов  $S_0, S_1, S_2, \dots, S_n$ , связанных между собой рядом условий, которые мы опишем позже, после введения обозначений. Множество  $S_0$  состоит из элементов нулевой размерности, т.е. из узлов. Множество  $S_1$  состоит из одномерных элементов, т.е. из ветвей. Множество  $S_2$  состоит из двумерных элементов, т.е. из граней и т.д. Поскольку у нас нет достаточного числа подходящих алфавитов для обозначения объектов различной размерности, то элементы  $S_0$  обозначаем  $0_1, 0_2, 0_3, \dots$ , элементы  $S_1$  обозначаем  $1_1, 1_2, 1_3, \dots$ , элементы  $S_2$  обозначаем  $2_1, 2_2, 2_3 \dots$  и так далее. В последующих примерах для обозначения элементов  $S_0$  будем пользоваться латинскими буквами, элементы  $S_1$  будем обозначать греческими буквами, элементы  $S_2$  будем обозначать словами (например, цветами Красный, Синий и т.д.).

Для каждого множества  $S_k$  введем векторное пространство  $k$ -цепей  $C_k$ , состоящее из векторов, компоненты которых индексированы элементами  $S_k$ . Из изучения электрических цепей нам известны пространства  $C_0$  и  $C_1$ . В общем случае размерность пространства  $C_k$  равна числу элементов  $S_k$ . Как и раньше, элемент множества  $S_k$  будем идентифицировать с вектором, у которого одна компонента, соответствующая этому элементу, равна еди-

нице, а остальные компоненты равны нулю. Например, для комплекса, изображенного на рис. 14.1, мы имеем вектора в пространстве  $C_0$ :  $\mathbf{0}_1 = A = (1, 0, 0, \dots)^T$ ,  $\mathbf{0}_2 = B = (0, 1, 0, \dots)^T$ , и т. д.; в пространстве  $C_1$ :  $\mathbf{1}_1 = \alpha = (1, 0, 0, \dots)^T$ ,  $\mathbf{2}_1 = \beta = (0, 1, 0, \dots)^T$  и т. д.; в пространстве  $C_2$ :  $\mathbf{2}_1 = \text{Красный} = (1, 0)^T$ ,  $\mathbf{2}_2 = \text{Синий} = (0, 1)^T$ .

Чтобы закончить описание комплекса, мы должны определить граничный оператор, в действительности последовательность граничных операторов, которую будем обозначать тем же символом  $\partial : C_k \rightarrow C_{k-1}$  для любого  $k$ . Оператор  $\partial$  приписывает каждому элементу  $S_k$  определенную комбинацию элементов множества  $S_{k-1}$  (с соответствующим знаком), которая и образует границу этого элемента. Оператор  $\partial$  определен для каждого базисного элемента  $C_k$  и, следовательно, определяет линейное отображение  $C_k$  на  $C_{k-1}$ . (Для  $k = 0$  мы выбираем  $\partial = 0$ , поскольку не существует элементов отрицательной размерности.)

Для 2-комплекса, показанного на рис. 14.1, мы имеем

$$\partial\alpha = B - A, \quad \partial\beta = C - A, \quad \text{и т. д.}$$

Эти соотношения определяют граничное отображение из  $C_1$  в  $C_0$ , которое представляется матрицей  $8 \times 6$ . Чтобы написать этот оператор в явном виде, необходимо на рисунке изобразить стрелками ориентацию каждой ветви. Аналогично, чтобы написать граничный оператор из  $C_2$  в  $C_1$ , мы должны задать ориентацию каждой грани (по часовой стрелке или против). Границей грани является замкнутый путь вдоль ее периметра, проходимый в направлении, совпадающим с ориентацией грани. Например, на рис. 14.1  $\partial(\text{Красный}) = \alpha + \gamma - \beta$ , а  $\partial(\text{Синий}) = \gamma + \nu - \varepsilon - \delta$ . В качестве отображения из  $C_2$  в  $C_1$  граничный оператор задается матрицей

$$\partial = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 1 & 1 \\ 0 & -1 \\ 0 & -1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

в базисах {Красный, Синий} пространства  $C_2$  и  $\{\alpha, \beta, \gamma, \varepsilon, \lambda, \nu\}$  пространства  $C_1$ .

При изучении одномерных комплексов мы обнаружили, что полезно рассматривать подпространства ядра и образа оператора  $\partial$ . Обобщим эту идею, определив пространство  $k$ -циклов  $Z_k$ , которое будет подпространством  $C_k$ , состоящим из элементов с нулевой границей, т. е. *ядро*  $\partial : C_k \rightarrow C_{k-1}$ . Например, на рис. 14.1

- $Z_2$  является нулевым подпространством, потому что нет такой нетривиальной линейной комбинации Красный и Синий, которая имеет нулевую границу;
- $Z_1$  является трехмерным подпространством, порожденным ячейками

$$\mathbf{M}_1 = \alpha + \gamma - \beta, \quad \mathbf{M}_2 = \gamma + \nu - \varepsilon - \delta, \quad \mathbf{M}_3 = \lambda - \mu - \nu;$$

- $Z_0$  есть все пространство  $C_0$ , поскольку по определению  $\partial c = 0$  для любой 0-цепи  $c$ .

Аналогично определим  $B_k$  — пространство  $k$ -границ, которое будет подпространством  $C_k$ , состоящим из элементов, являющимися границами элементов пространства  $C_{k+1}$ , т. е. *образ*  $\partial : C_{k+1} \rightarrow C_k$ . Для комплекса, изображенного на рис. 14.1,

- $B_2$  является нулевым подпространством, потому что не существует трехмерных элементов в двумерном комплексе (точно так же, как в одномерных комплексах, которые мы рассматривали,  $B_1$  всегда было нулевым подпространством);
- $B_1$  является двумерным пространством, базис которого образуют  $\partial(\text{Красный}) = \alpha + \gamma - \beta$  и  $\partial(\text{Синий}) = \gamma + \nu - \varepsilon - \delta$ ;
- $B_0$  является пятимерным подпространством; так же, как если бы комплекс был одномерным, а не двумерным.

В этом примере каждое пространство  $B_k$  является подпространством соответствующего  $Z_k$ , т. е. *каждая граница является циклом*. Немного подумав, можно убедиться, что это свойство должно выполняться для любого двумерного комплекса,

схема которого подобна рис. 14.1. Поскольку  $B_2$  нулевое, очевидно, что оно является подпространством  $Z_2$ . Поскольку граница любого многоугольника — замкнутая траектория, то  $B_1$  есть подпространство  $Z_1$ . И наконец, поскольку  $Z_0$  — это все пространство  $C_0$ , очевидно, что  $B_0$  есть подпространство  $Z_0$ . Условие  $B_k \subset Z_k$  можно вывести и по-другому. Поскольку  $B_k$  есть образ  $\partial : C_{k+1} \rightarrow C_k$  и  $Z_k$  есть ядро  $\partial : C_k \rightarrow C_{k-1}$ , то условие  $B_k \subset Z_k$  означает, что композиция  $\partial \circ \partial : C_{k+1} \rightarrow C_{k-1}$  равна нулю. Другими словами: граница границы есть нуль. Из последующих примеров станет ясно, почему это справедливо для любого комплекса.

## Гомологии

Поскольку  $B_k$  является подпространством  $Z_k$ , то можно образовать факторпространство  $H_k = Z_k/B_k$ , которое называется *пространством  $k$ -мерных гомологий* комплекса. Оказывается, что оно имеет более фундаментальное значение, чем  $Z_k$  и  $B_k$  по отдельности. Пространства  $Z_k$  и  $B_k$  связаны со свойствами данного комплекса, а  $H_k$ , как мы скоро увидим, связано со свойствами пространства, построенного из элементов комплекса.

При изучении электрических цепей мы уже встречались с пространствами  $H_0$  и  $H_1$ . Вспомним, что размерность  $H_0$  равна числу связных компонент, а его дуальное пространство  $H^0$  интерпретируется как пространство потенциалов, которые постоянны на каждой связной компоненте. Поскольку пространство границ  $B_1$  нулевое для 1-комплекса (в нем нет двумерных элементов), то  $H_1 = Z_1/B_1 = Z_1$ . Таким образом, для любого один-комплекса  $H_1$  есть пространство циклов  $Z_1$ , которое играет столь важную роль в методе круговых токов Максвелла.

Давайте рассмотрим пространство  $H_k$  для 2-комплекса, изображенного на рис. 14.1. Поскольку у этого комплекса нет 2-циклов, пространство  $Z_2$  есть  $\{0\}$ , аналогично пространство  $H_2$  тоже нулевое. Поскольку  $Z_1$  — трехмерное пространство, а  $B_1$  — двумерное, то факторпространство  $H_1 = Z_1/B_1$  одномерно. Если посмотреть на рисунок, то видно, почему: путь  $\lambda - \mu - \nu$  не является границей никакого элемента пространства  $C_2$ . Если треугольник  $CEF$  добавить к комплексу в качестве третьего элемента  $C_2$ , то

путь  $\lambda - \mu - \nu$  стал бы границей и размерность пространства  $H_1$  была бы равна нулю. И наконец, поскольку комплекс связан, пространство  $H_0$  одномерно.

В качестве примера трехмерного комплекса рассмотрим тетраэдр со всеми его гранями, ребрами (ветвями) и вершинами (узлами), показанный на рис. 14.2(а). Развернутая поверхность тетраэдра изображена на рис. 14.2(б). В комплексе имеется четыре вершины:  $A, B, C, D$ ; шесть ребер:  $\alpha, \beta, \gamma, \delta, \varepsilon, \eta$ ; четыре грани: Красный, Жёлтый, Зелёный, Синий. Ориентация каждой грани выбирается так, чтобы стрелка, показанная на рисунке, вращалась против часовой стрелки, если смотреть на тетраэдр снаружи (см. Зелёный и Жёлтый на рис. 14.2(а)), и по часовой стрелке, если смотреть на него изнутри (см. рис. 14.2(б)).

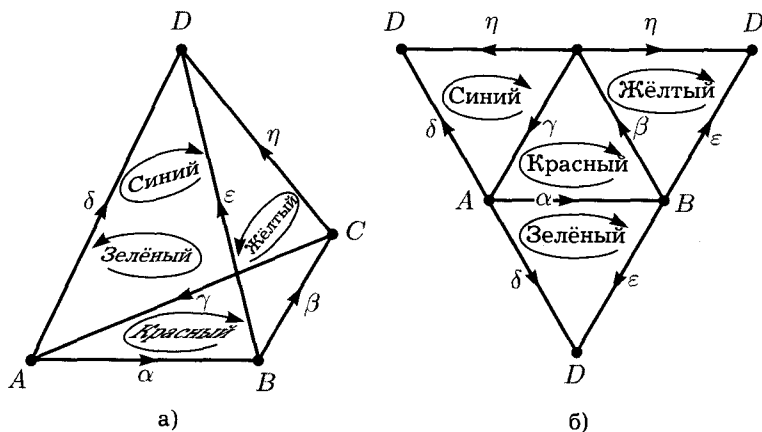


Рис. 14.2

Из рисунка легко увидеть, что граничный оператор из  $C_2$  в  $C_1$  определяется формулами

$$\partial(\text{Красный}) = -\alpha - \beta - \gamma,$$

$$\partial(\text{Зелёный}) = \alpha + \varepsilon - \delta.$$

Поскольку граница каждой грани является циклом, то каждая такая граница имеет нулевую границу, т. е.

$$\partial[\partial(\text{Зелёный})] = \partial(\alpha + \varepsilon - \delta) = (B - A) + (D - B) - (D - A) = 0.$$

В этом комплексе имеется один трехмерный элемент — сам тетраэдр. Чтобы написать граничный оператор из  $C_3$  в  $C_2$ , надо приписать этому элементу ориентацию. (Хотя совсем не просто изобразить эту ориентацию на рисунке.) Мы должны сделать это в абстрактной форме. Выберем так называемую *правостороннюю ориентацию*. Это значит, что когда мы смотрим на элемент снаружки, то все грани, составляющие его границу, должны быть ориентированы против часовой стрелки. Ориентация называется правосторонней, потому что если положить *правую* руку на поверхность тетраэдра так, чтобы большой палец смотрел наружу, то пальцы правой руки будут поворачиваться так, что грань войдет в границу со знаком *плюс*. Конечно, это утверждение выглядит короче, если записать его алгебраически через граничный оператор из  $C_3$  в  $C_2$ :

$$\partial(\text{Тетраэдр}) = \text{Красный} + \text{Жёлтый} + \text{Зелёный} + \text{Синий}.$$

Если для тетраэдра выбирается левосторонняя ориентация, то все четыре грани войдут в выражение для границы со знаком *минус*. Таким образом, изменение ориентации любой грани ведет к изменению ее знака в выражении для границы.

Очевидно, что граница тетраэдра — замкнутая поверхность, имеющая нулевую границу. Это легко показать алгебраически, используя выражения для граничных операторов их  $C_3$  в  $C_2$  и из  $C_2$  в  $C_1$ :

$$\begin{aligned} \partial[\partial(\text{Тетраэдр})] &= \partial[\text{Красный} + \text{Жёлтый} + \text{Зелёный} + \text{Синий}] \\ &= (-\alpha - \beta - \gamma) + (\beta + \eta - \varepsilon) + (\alpha + \varepsilon - \delta) + (\gamma + \delta - \eta) = 0. \end{aligned}$$

Этот результат никак не зависит от выбора ориентации самого тетраэдра, его граней или ребер. (Попробуйте изменить ориентацию какой-нибудь грани и пары ребер и убедитесь в этом.)

Итак, наши комплексы первоначально были изображены рисунками, потом мы преобразовали рисунки в алгебраические выражения для граничного оператора  $\partial$ . Очевидно, что граничный оператор, как и рисунок, определяет комплекс полностью. Поэтому очень хочется уметь задавать комплекс, не рисуя его, а просто выписывая его элементы и граничный оператор. Когда мы так сделаем, то покажем, что для любого элемента  $c$  комплекса выполняется условие  $\partial(\partial c) = 0$ .

путь  $\lambda - \mu - \nu$  стал бы границей и размерность пространства  $H_1$  была бы равна нулю. И наконец, поскольку комплекс связан, пространство  $H_0$  одномерно.

В качестве примера трехмерного комплекса рассмотрим тетраэдр со всеми его гранями, ребрами (ветвями) и вершинами (узлами), показанный на рис. 14.2(а). Развернутая поверхность тетраэдра изображена на рис. 14.2(б). В комплексе имеется четыре вершины:  $A, B, C, D$ ; шесть ребер:  $\alpha, \beta, \gamma, \delta, \varepsilon, \eta$ ; четыре грани: Красный, Жёлтый, Зелёный, Синий. Ориентация каждой грани выбирается так, чтобы стрелка, показанная на рисунке, вращалась против часовой стрелки, если смотреть на тетраэдр снаружи (см. Зелёный и Жёлтый на рис. 14.2(а)), и по часовой стрелке, если смотреть на него изнутри (см. рис. 14.2(б)).

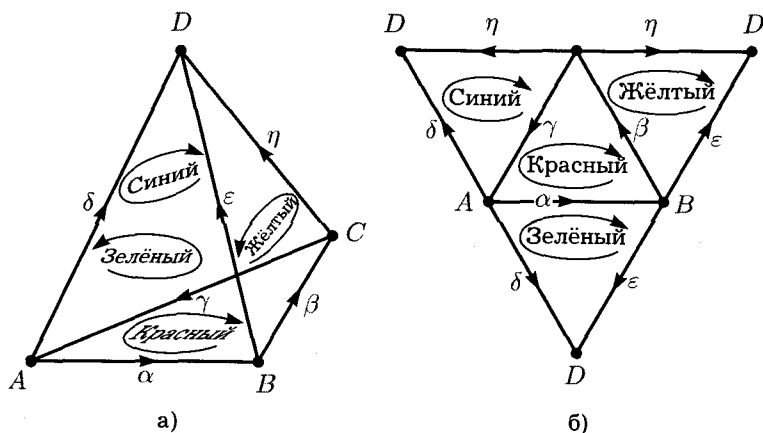


Рис. 14.2

Из рисунка легко увидеть, что граничный оператор из  $C_2$  в  $C_1$  определяется формулами

$$\partial(\text{Красный}) = -\alpha - \beta - \gamma,$$

$$\partial(\text{Зелёный}) = \alpha + \varepsilon - \delta.$$

Поскольку граница каждой грани является циклом, то каждая такая граница имеет нулевую границу, т. е.

$$\partial[\partial(\text{Зелёный})] = \partial(\alpha + \varepsilon - \delta) = (B - A) + (D - B) - (D - A) = 0.$$

В этом комплексе имеется один трехмерный элемент — сам тетраэдр. Чтобы написать граничный оператор из  $C_3$  в  $C_2$ , надо приписать этому элементу ориентацию. (Хотя совсем не просто изобразить эту ориентацию на рисунке.) Мы должны сделать это в абстрактной форме. Выберем так называемую *правостороннюю ориентацию*. Это значит, что когда мы смотрим на элемент снаружи, то все грани, составляющие его границу, должны быть ориентированы против часовой стрелки. Ориентация называется правосторонней, потому что если положить *правую* руку на поверхность тетраэдра так, чтобы большой палец смотрел наружу, то пальцы правой руки будут поворачиваться так, что грань войдет в границу со знаком *плюс*. Конечно, это утверждение выглядит короче, если записать его алгебраически через граничный оператор из  $C_3$  в  $C_2$ :

$$\partial(\text{Тетраэдр}) = \text{Красный} + \text{Жёлтый} + \text{Зелёный} + \text{Синий}.$$

Если для тетраэдра выбирается левосторонняя ориентация, то все четыре грани войдут в выражение для границы со знаком *минус*. Таким образом, изменение ориентации любой грани ведет к изменению ее знака в выражении для границы.

Очевидно, что граница тетраэдра — замкнутая поверхность, имеющая нулевую границу. Это легко показать алгебраически, используя выражения для граничных операторов их  $C_3$  в  $C_2$  и из  $C_2$  в  $C_1$ :

$$\begin{aligned} \partial[\partial(\text{Тетраэдр})] &= \partial[\text{Красный} + \text{Жёлтый} + \text{Зелёный} + \text{Синий}] \\ &= (-\alpha - \beta - \gamma) + (\beta + \eta - \epsilon) + (\alpha + \epsilon - \delta) + (\gamma + \delta - \eta) = 0. \end{aligned}$$

Этот результат никак не зависит от выбора ориентации самого тетраэдра, его граней или ребер. (Попробуйте изменить ориентацию какой-нибудь грани и пары ребер и убедитесь в этом.)

Итак, наши комплексы первоначально были изображены рисунками, потом мы преобразовали рисунки в алгебраические выражения для граничного оператора  $\partial$ . Очевидно, что граничный оператор, как и рисунок, определяет комплекс полностью. Поэтому очень хочется уметь задавать комплекс, не рисуя его, а просто выписывая его элементы и граничный оператор. Когда мы так сделаем, то покажем, что для любого элемента  $c$  комплекса выполняется условие  $\partial(\partial c) = 0$ .

Давайте разберемся, почему условие  $\partial(\partial c) = 0$  разумно для любого комплекса, даже, например, для поверхности, которую нельзя построить в трехмерном пространстве, или для комплексов, элементы которых включают четырехмерное пространство-время. Чтобы доказать, что условие  $\partial(\partial c) = 0$  выполняется для всех элементов  $c$ , достаточно проверить, что оно выполняется для генераторов, т. е. для элементов пространства  $S_k$ . Для любого такого элемента мы можем построить *подкомплекс*, состоящий из этого элемента, всех элементов пространства  $S_{k-1}$ , которые входят в выражение для его границ нетривиальным образом, всех границ и так далее. Граничные соотношения между этими элементами определяются свойствами того элемента, с которого мы начали. Несмотря на то, что взаимосвязи между различными строительными блоками могут быть достаточно сложными, мы ожидаем, что отдельные элементы, из которых построен наш комплекс, будут вполне знакомыми геометрическими объектами. Например, предположим, что каждый элемент пространства  $S_3$  является выпуклым многогранником в трехмерном пространстве (по крайней мере, если нам интересны его соотношения с границами, мы допускаем любые непрерывные искажения реальной геометрической формы объекта — так же, как раньше мы не интересовались формой проводов при изучении электрических цепей). Граница такого многогранника будет состоять из конечного числа многоугольных граней, каждая из которых ориентирована против часовой стрелки, если смотреть на многогранник снаружи (рис. 14.3). Когда мы вычислим сумму границ этих многоугольных граней, каждое ребро будет входить дважды от каждой грани с противоположными знаками. Поэтому сумма всех границ будет равна нулю. В этом случае граница каждого элемента  $S_3$  является циклом. Аналогично, мы могли бы предположить, что каждый элемент  $S_2$  является выпуклым плоским многоугольником. Очевидно, что граница каждого элемента  $S_2$  будет циклом (по существу это соответствует ячейке).

Давайте исследуем пространства гомологий комплекса, являющегося тетраэдром. Если нас интересует  $H_0$ , то мы можем проигнорировать  $C_2$  и  $C_3$  и рассмотреть 1-комплекс, заданный вершинами и ребрами. Это связный 1-комплекс с  $\dim H_0 = 1$ .

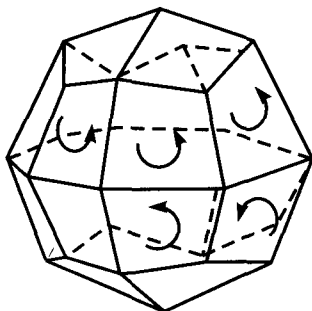


Рис. 14.3

Чтобы определить  $H_1$ , рассмотрим 2-комплекс, заданный вершинами, ребрами и гранями. Пространство  $Z_1$  проанализируем с помощью метода, которым мы пользовались при изучении электрических цепей. Рассмотрим рис. 14.2. На нем ветви  $\delta$ ,  $\varepsilon$  и  $\eta$  образуют максимальное дерево. Каждая из остальных ветвей  $\alpha$ ,  $\beta$ ,  $\gamma$  определяет независимый цикл. Следовательно,  $\dim Z_1 = 3$ . Каждый из этих циклов является границей грани, следовательно,  $\dim B_3 = 3$ . Отсюда следует, что  $\dim H_1 = \dim Z_1 - \dim B_1 = 0$ , т. е.  $H_1 = \{0\}$ .

Теперь рассмотрим  $H_2$ . Сначала возьмем пространство 2-циклов  $Z_2$  — ядро оператора  $\partial$ , действующего на  $C_2$ . Мы уже видели, что образ  $\partial$ , т. е. пространство  $B_1 \subset C_2$ , трехмерен. Поскольку  $C_2$  — четырехмерное пространство, ядро  $\partial$  одномерно. Базисом пространства  $Z_2$  является элемент Красный + Жёлтый + Зелёный + Синий, т. е. поверхность тетраэдра, которая имеет нулевую границу. Конечно, этот элемент принадлежит и пространству  $B_2$ , поскольку он является границей тетраэдра. Следовательно,  $\dim B_2 = 1$ ,  $\dim Z_2 = 1$  и  $\dim H_2 = 0$ .

Наконец, рассмотрим  $H_3$ . Заметим, что в нем нет 3-циклов, так что  $\dim Z_3 = 0$  и  $\dim H_3 = 0$ .

## Теорема Эйлера

Изучая электрические цепи, мы доказали, что

$$\dim H_0 - \dim H_1 = \dim C_0 - \dim C_1.$$

Теорема Эйлера обобщает этот результат на  $n$ -комплексы:

$$\begin{aligned} \dim H_0 - \dim H_1 + \dim H_2 - \cdots \pm \dim H_n \\ = \dim C_0 - \dim C_1 + \dim C_2 - \cdots \pm \dim C_n. \end{aligned}$$

Стоящее в обеих частях этого равенства число называется *эйлеровой характеристикой* комплекса. Мы скоро докажем этот результат в общем случае, но сначала проверим его для случая тетраэдра. Поскольку  $\dim H_0 = 1$ , а пространства  $H_1$ ,  $H_2$  и  $H_3$  тривиальны (равны  $\{0\}$ ), имеем

$$\dim H_0 - \dim H_1 + \dim H_2 - \dim H_3 = 1.$$

С другой стороны,

$$\dim C_0 - \dim C_1 + \dim C_2 - \dim C_3 = 4 - 6 + 4 - 1 = 1.$$

Предположим, что вместо тетраэдра (с внутренностью) мы рассматриваем 2-комплекс, состоящий из поверхности тетраэдра. Единственное отличие в том, что теперь в комплексе нет трехмерного элемента. Как и раньше,  $\dim H_0 = 1$  и  $\dim H_1 = 0$ . По-прежнему  $\dim Z_2 = 1$ , поскольку поверхность тетраэдра имеет нулевую границу. Однако теперь  $\dim B_2 = 0$ , поскольку нет трехмерного элемента, границей которого был бы какой-нибудь двумерный элемент. Поэтому имеем  $\dim H_2 = \dim Z_2 - \dim B_2 = 1$ . Тогда для этого комплекса  $\dim H_0 - \dim H_1 + \dim H_2 = 1 - 0 + 1 = 2$ , а  $\dim C_0 - \dim C_1 + \dim C_2 = 4 - 6 + 4 = 2$

Давайте сделаем то же самое для сплошного (трехмерного) куба. Он изображен в перспективе на рис. 14.4(а), а его развернутая поверхность изображена на рис. 14.4(б). Как и в предыдущем случае,  $\dim H_0 = 1$ , потому что комплекс связан, и  $\dim H_3 = 0$ , потому что нет 3-циклов.

Чтобы найти размерность  $Z_1$ , рассмотрим максимальное дерево, состоящее из семи ветвей, показанное на рис. 14.4 жирными линиями. Каждая из оставшихся пяти ветвей определяет цикл. Следовательно,  $\dim Z_1 = 5$ , но  $\dim B_1 = 5$ , поэтому  $\dim H_1 = 0$ .

Теперь займемся пространством  $H_2$ . Поскольку  $\dim C_2 = 6$  и  $\dim B_2 = 5$ , мы делаем вывод, что  $\dim Z_2 = 1$ . Базисом  $Z_2$  является вся поверхность куба. Но эта поверхность является границей куба, поэтому  $\dim B_2 = 1$  и  $\dim H_2 = \dim Z_2 - \dim B_2 = 0$ .

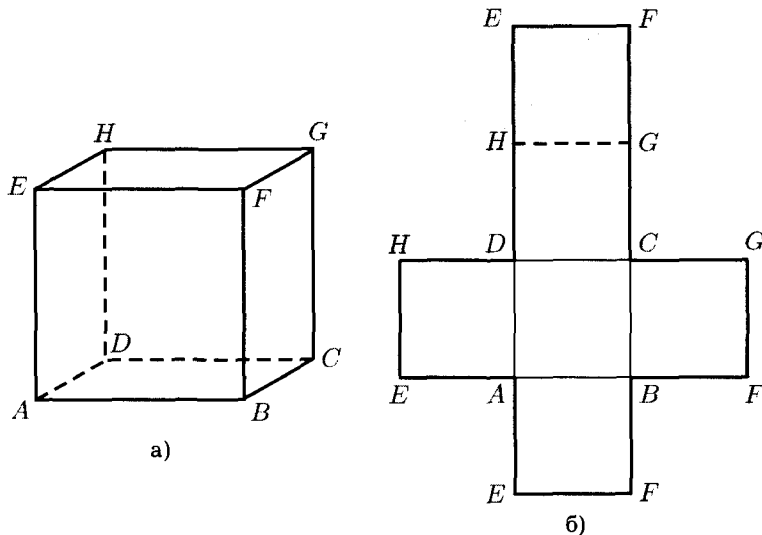


Рис. 14.4

Оказывается, что пространства гомологий куба и тетраэдра одни и те же. Для куба мы получаем  $\dim H_0 - \dim H_1 + \dim H_2 - \dim H_3 = 1$ . Нетрудно также проверить, что

$$\dim C_0 - \dim C_1 + \dim C_2 - \dim C_3 = 8 - 12 + 6 - 1 = 1.$$

### Инвариантность гомологий

С точки зрения топологии есть фундаментальная причина, почему комплекс куба и комплекс тетраэдра имеют одни и те же пространства гомологий. *Пространство*, заполняющее куб, — это то же самое пространство, которое заполняет тетраэдр, и, в свою очередь, совпадает с пространством, заполняющим шар. Все эти пространства могут быть непрерывным образом преобразованы друг в друга<sup>1</sup>. Аналогично, поверхности куба, тетраэдра и шара представляют одно и то же пространство. Рис. 14.5(а) показывает, как поверхность сферы может перейти при разбиении на

<sup>1</sup>Точнее, говорят, что такие пространства *гомеоморфны*, или топологически эквивалентны — каждое из них допускает взаимно-однозначное непрерывное в обе стороны отображение (*гомеоморфизм*) на другое. — Прим. ред.

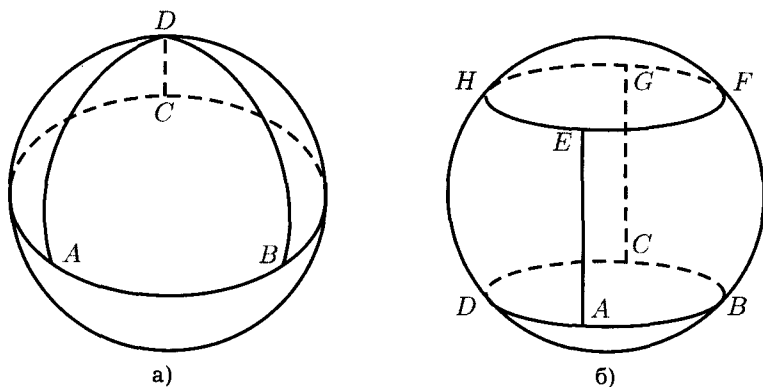


Рис. 14.5

грани в поверхность тетраэдра, а рис. 14.5(б) показывает, как эта же поверхность преобразуется в поверхность куба. Конечно, *комплекс* тетраэдра и *комплекс* куба отличаются, но они представляют собой различные способы разбиения *одного и того же пространства*.

В топологии имеется теорема, которая, грубо говоря, утверждает следующее. Предположим, что мы строим пространство  $X$ , соединяя различные ячейки друг с другом вдоль граничных ячеек, размерность которых на единицу меньше. Например, построим поверхность сферы, соединяя вместе четыре 2-ячейки вдоль их ребер, как это показано на рис. 14.5(а), или соединяя вместе шесть 2-ячеек, как это показано на рис. 14.5(б). Другими словами, мы можем построить трехмерный ансамбль многогранников, соединяя ячейки вдоль их граней, подобно скоплению мыльных пузырей. Тогда теорема говорит, что *пространства гомотопий результирующего комплекса зависят только от пространства  $X$  и не зависят от способа соединения*. Если пространство  $X$  разложить на ячейки каким-то другим способом, то мы получим другой комплекс. Пространства  $H_k$  для разных комплексов будут изоморфны при каждом  $k$ .

На рис. 14.6 показано, как сфера может быть разложена в комплекс еще проще, чем преобразование в куб или тетраэдр. Поверхность сферы разбивается на два треугольника, соответ-

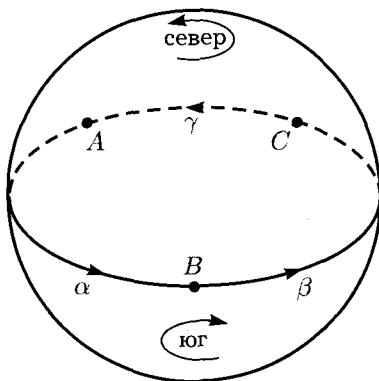


Рис. 14.6

ствующие северной и южной полусферам, соединяющимся вдоль экватора.

У этого 2-комплекса существует только одна ячейка  $\alpha + \beta + \gamma$ , которая является границей обеих полусферических поверхностей. Следовательно,  $\dim Z_1 = \dim B_1 = 1$  и опять  $\dim H_1 = 0$ . Сумма этих двух граней имеет нулевую границу, поскольку

$$\partial(\text{Север}) = \alpha + \beta + \gamma, \quad \text{а} \quad \partial(\text{Юг}) = -\alpha - \beta - \gamma.$$

Поэтому так же, как и для поверхности куба или тетраэдра, имеем  $\dim H_2 = \dim Z_2 = 1$ . И опять замечаем, что для 2-комплекса

$$\begin{aligned} \dim C_0 - \dim C_1 + \dim C_2 &= 3 - 3 + 2 = 2 \\ &= \dim H_0 - \dim H_1 + \dim H_2. \end{aligned}$$

К сожалению, мы не готовы доказать упомянутую выше теорему и даже дать ее точную формулировку. Чтобы это сделать, надо иметь точные определения понятий *пространство*, *деформация*, *соединение* и т. д. Когда эти понятия строго определены, то совсем не сложно сформулировать и доказать теорему, но развитие необходимой для этого базы уведет нас слишком далеко. Тем не менее, мы будем пользоваться этой теоремой, не заботясь о том, насколько точно она сформулирована. С ее помощью мы проделаем ряд вычислений, чтобы лучше познакомиться с пространствами  $H_k$ .

## Top

Рассмотрим пример 2-комплекса, имеющего пространства гомологий, отличные от тех, которые мы уже рассматривали. Для этого получим разложение двумерного тора, т. е. поверхности бублика. (Вместо того, чтобы разрезать мячик, мы разрежем камеру колеса.) Рис. 14.7(а) показывает, как тор разбивается на две области; рис. 14.7(б) показывает, как эти области делаются плоскими; рис. 14.7(в) показывает, как эти области разделяются. Можно считать, что рис. 14.7(в) показывает элементы тора, и попытаться понять, в чем принципиальное отличие этого набора элементов от набора элементов сферы, который получился бы, если бы мы разделили на элементы рис. 14.5(а).

Посчитав вершины, ребра и грани, мы находим  $\dim C_0 = 4$ ,  $\dim C_1 = 8$  и  $\dim C_2 = 4$ . Откуда получаем, что  $\dim C_0 - \dim C_1 + \dim C_2 = 4 - 8 + 4 = 0$ . (Вспомним, что для 2-комплекса, образованного из поверхности сферы, эта величина всегда равна 2.)

Рассмотрим пространства гомологий для комплекса тора. Так как этот комплекс связный, то пространство  $H_0$  одномерно. Поскольку вся поверхность тора имеет нулевую границу, то  $\dim Z_2 = 1$ . (Проверьте самостоятельно, что для указанных ориентаций  $\partial(\text{Красный} + \text{Жёлтый} + \text{Зелёный} + \text{Синий}) = 0$ .) Мы рассматриваем 2-комплекс, поэтому в пространстве  $C_2$  не может быть элементов, являющихся границами. Значит,  $\dim B_2 = 0$  и  $\dim H_2 = \dim Z_2 - \dim B_2 = 1$ . (В общем случае очевидно, что по тем же причинам для любого 2-комплекса, полученного делением гладкой связной поверхности, являющейся границей области в пространстве  $\mathbb{R}^3$ , имеет место  $\dim H_2 = 1$ .)

Наконец, рассмотрим пространство  $H_1$ . Рис. 14.8 показывает максимальное дерево, состоящее только из трех ребер. Самостоятельно проверьте, что если к этому максимальному дереву добавить любое из оставшихся пяти ребер, то получается цикл! Например,  $\alpha + \beta$  является циклом, так же, как  $\varepsilon + \lambda$  и  $\mu + \eta$ . Поэтому для этого комплекса  $\dim Z_1 = 5$ . С другой стороны,  $\dim C_2 = 4$ . Ядро оператора  $\partial$ , действующего на  $C_2$ , — пространство  $Z_2$  — является одномерным, а образ  $\partial$  в пространстве  $B_1$  имеет размерность 3. Поэтому

$$\dim H_1 = \dim Z_1 - \dim B_1 = 5 - 3 = 2.$$

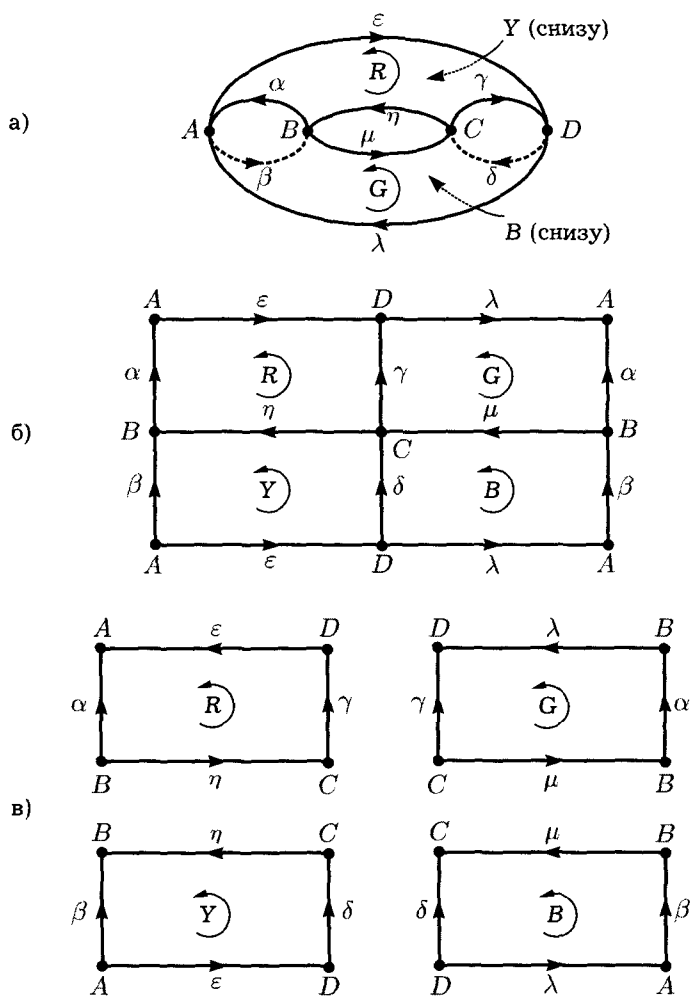


Рис. 14.7

Полезно далее построить базисы для пространств  $B_1$  и  $H_1$ . Проще всего в качестве базисных элементов в  $B_1$  взять ячейки, которые являются границами четырех граней, например,

$$\begin{aligned}\partial(\text{Жёлтый}) &= \eta - \beta + \varepsilon + \delta, \\ \partial(\text{Зелёный}) &= \alpha - \lambda - \gamma - \mu, \\ \partial(\text{Синий}) &= \lambda + \beta + \mu - \delta.\end{aligned}$$

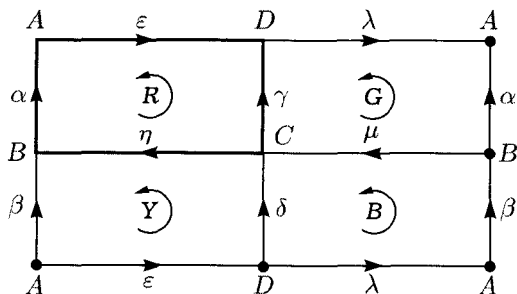


Рис. 14.8

Элементами факторпространства  $H_1 = Z_1/B_1$  являются классы эквивалентности *по модулю*  $B_1$ . Поэтому мы должны найти два цикла, которые не являются границами и, более того, любая их линейная комбинация тоже не граница. Если посмотреть на рис. 14.7(а), то там сразу видно два таких цикла: цикл  $\alpha + \beta$ , идущий вокруг тора в одном направлении, и цикл  $\varepsilon + \lambda$ , идущий вокруг тора в другом независимом направлении. Поэтому базис пространства  $H_1$  состоит из классов эквивалентности  $\overline{\alpha + \beta}$  и  $\overline{\varepsilon + \lambda}$ . Конечно, есть много других циклов, не являющихся границами, но все они попадают в класс эквивалентности некоторой линейной комбинации  $\alpha + \beta$  и  $\varepsilon + \lambda$ . Например, цикл  $\gamma + \delta$  лежит в классе эквивалентности  $\alpha + \beta$ , потому что

$$\gamma + \delta = \alpha + \beta + \partial(\text{Красный}) + \partial(\text{Жёлтый}).$$

Цикл  $\alpha + \varepsilon + \delta - \mu$  может быть представлен в виде

$$(\alpha + \beta) + (\varepsilon + \lambda) - \partial(\text{Синий}).$$

Поэтому он лежит в том же классе эквивалентности, что и

$$\alpha + \beta + \varepsilon + \lambda.$$

### Смысл пространства $H_1$

Сейчас мы уже можем в общем случае понять смысл  $\dim H_1$ . Грубо говоря, это число независимых замкнутых кривых, которые не ограничивают никакой области. На сфере любая замкнутая

кривая обязательно ограничивает какую-то область и поэтому  $\dim H_1 = 0$ . С другой стороны, на торе существует два различных типа замкнутых кривых, которые не являются границами подобластей, как это показано на рис. 14.9. Поэтому для тора  $\dim H_1 = 2$ . Заметим, что для комплекса, который мы построили для тора, выполняется условие  $\dim C_0 - \dim C_1 + \dim C_2 = 4 - 8 + 4 = \dim H_0 - \dim H_1 + \dim H_2$ . Посмотрим, как изменится пространство  $H_1$ , если мы приделаем к тору ручку (рис. 14.10). Это значит, что на его поверхности вырезаются два диска и на их место вставляется (искривленный) цилиндр, открытый с обоих концов: верхнюю часть цилиндра мы прикрепляем к границе одного диска, а нижнюю — к границе второго диска. Выберем такое разбиение исходной поверхности, чтобы каждый диск был ячейкой этого разбиения. Пусть это будет квадрат, а цилиндр мы представим, как будто он склеен из четырех граней. Присоединяя цилиндр таким образом, мы не создаем новых узлов, но добавляем четыре новых ветви (ребра «квадратного» цилиндра). Теперь удалим из  $S_2$  два элемента (внутренние части дисков) и добавим четыре новых (границ цилиндра), см. рис. 14.11. Для получившейся в результате поверхности по-прежнему пространства  $H_0$  и  $H_1$  одномерны. Вычисляя  $Z_1$  для нового комплекса, мы можем выбрать максимальное дерево, которое не содержит ни одну из новых ветвей (поскольку каждая новая ветвь лежит на ячейке — вспомним процедуру построения максимального дерева, которая была сформулирована при изучении электрических цепей). Итак, мы построили максимальное дерево, целиком состоящее из ветвей первоначального комплекса. Каждая ветвь цилиндра теперь дает новую независимую ячейку, что приводит к увеличе-

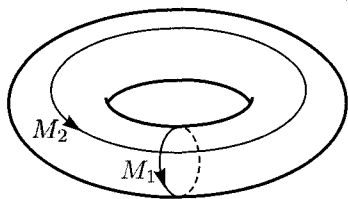


Рис. 14.9

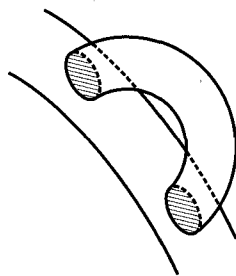


Рис. 14.10

нию размерности  $B_1$  на 4. С другой стороны, размерность  $C_2$  и, следовательно,  $B_1$  увеличилась на 2. Но тогда и размерность  $H_1$  тоже увеличилась на 2. Итак, мы показали, что добавляя ручку, мы увеличиваем размерность  $H_1$  на 2. Если к сфере прикрепить  $k$  ручек, то  $\dim H_0 = 1$ ,  $\dim H_1 = 2k$  и  $\dim H_2 = 1$ .<sup>2</sup> Например, тор можно представить себе как сферу, к которой прикрепили одну ручку (рис. 14.12). Это дает нам еще одно доказательство результатов, полученных для тора.

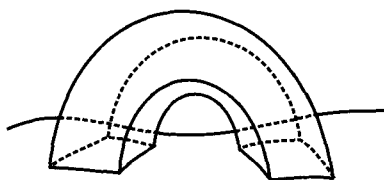


Рис. 14.11

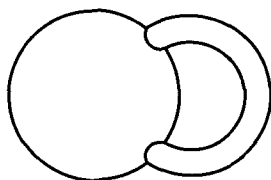


Рис. 14.12

Можно исследовать результат вырезания отверстия в поверхности. В качестве примера рассмотрим комплекс, который (вместе с внутренностью) образует на плоскости квадрат. Очевидно, что для такого комплекса  $\dim H_0 = 1$ ,  $\dim H_1 = 0$  и  $\dim H_2 = 0$ . Предположим, что из внутреннего квадрата мы вырезали отверстие (рис. 14.13). Если бы у нас было разбиение исходного квадрата, для которого вырезанное отверстие было одним из элементов  $S_2$ , то мы бы увидели, что удаление отверстия не изменяет  $Z_1$ , но уменьшает размерность  $C_2$  и, следовательно,  $B_1$  на 1. Но тогда размерность  $H_1$  увеличивается на 1. Итак, если на плоскости имеется область, из которой вырезаны  $k$  отверстий, то  $\dim H_1 = k$ ,  $\dim H_0 = 1$  и  $\dim H_2 = 0$ . Аналогичные рассуждения в трехмерном пространстве показывают, что если из трехмерной области вырезать  $k$  (трехмерных) отверстий, то  $\dim H_0 = 1$ ,  $\dim H_1 = 0$ ,  $\dim H_2 = k$  и  $\dim H_3 = 0$ .

<sup>2</sup>Известно, что всякая замкнутая (не имеющая края) компактная связная ориентируемая двумерная поверхность топологически эквивалентна сфере с добавлением определенного количества ручек, которое называется *родом* поверхности и обозначается  $g$  (genus). Таким образом, род равен половине размерности пространства  $H_1$  одномерных гомологий поверхности. — *Прим. ред.*

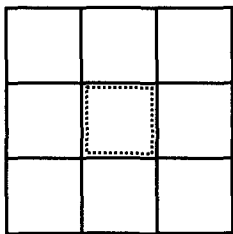


Рис. 14.13

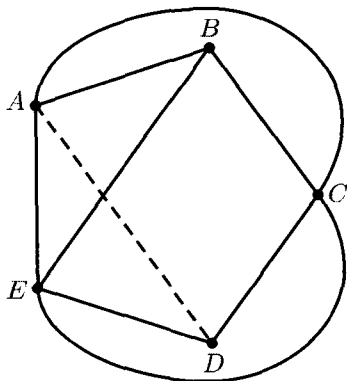


Рис. 14.14

## Бутылка Клейна

При изучении электрических цепей Вы, должно быть догадались, что существуют цепи, подобные изображенной на рис. 14.14, которые не могут быть уложены на плоскости так, чтобы не пересекались отдельные провода. Однако, такие цепи всегда можно собрать в трехмерном пространстве. Точно так же существуют 2-комплексы, которые нельзя собрать в трехмерном пространстве. Хорошо известным примером является *бутылка Клейна*, из поверхности которой получается комплекс, изображенный на рис. 14.15(а). Несмотря на близкое сходство с тором<sup>3</sup> (см. рис. 14.7(б)), эту поверхность в трехмерном пространстве построить нельзя. Этот (развернутый) комплекс можно собрать в цилиндр, изображенный на рис. 14.15(б). Но при этом оказывается невозможно соединить вместе ребра  $\epsilon$  и  $\lambda$  с соблюдением их ориентации. Для комплекса, представляющего бутылку Клейна, гомологические пространства отличаются от гомологических пространств тора. Конечно,  $\dim H_0 = 1$ , потому что комплекс связный. Однако  $\dim H_2 = 0$ , потому что в пространстве  $C_2$  нет

<sup>3</sup>Бутылка Клейна, в отличие от сферы и тора, относится к неориентируемым (односторонним) поверхностям. Неориентируемая поверхность может быть получена из ориентируемой заменой на лист Мёбиуса (одностороннюю поверхность с краем-окружностью) одного или нескольких дисков. — Прим. ред.

элемента с нулевой границей. (Границей всего распрямленного комплекса (см. рис. 14.15(а)) является  $2\varepsilon + 2\lambda$ . Проверьте самостоятельно, что при изменении знака одной или нескольких 2-ячеек в пространстве  $C_2$  не появится элемент с нулевой границей.) Для комплекса, изображенного на рис. 14.7,  $\dim Z_1 = 5$ , а для комплекса бутылки Клейна  $\dim B_1 = 4$  и  $\dim H_1 = 5 - 4 = 1$ . Это значит, что существует только один независимый цикл, который не является границей. Цикл  $\alpha + \beta$  не является границей ни на рис. 14.7(б), ни на рис. 14.15(а). А вот цикл  $\varepsilon + \lambda$ , который не является элементом пространства  $B_1$  в случае тора, в случае бутылки Клейна становится элементом  $B_1$ . Это половина границы всего комплекса, изображенного на рис. 14.15(а).

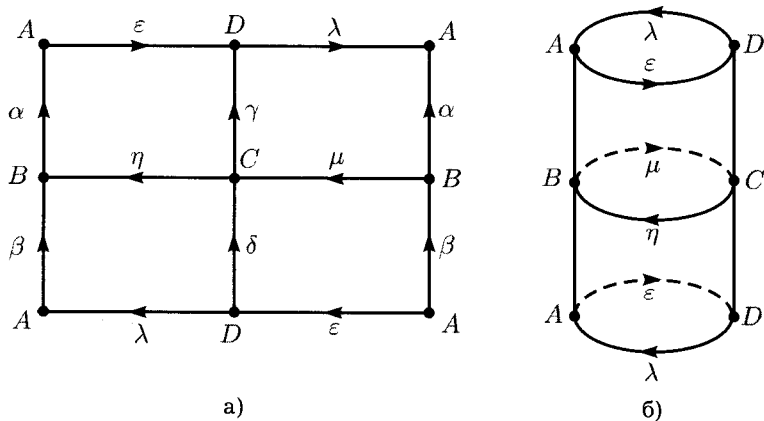


Рис. 14.15

И для тора, и для бутылки Клейна мы имеем  $\dim C_0 = 4$ ,  $\dim C_1 = 8$  и  $\dim C_2 = 4$ , так что

$$\dim C_0 - \dim C_1 + \dim C_2 = 4 - 8 + 4 = 0.$$

Полезно заметить, что для бутылки Клейна мы также имеем

$$\dim H_0 - \dim H_1 + \dim H_2 = 1 - 1 + 0.$$

## Доказательство теоремы Эйлера

В общем случае докажем формулу

$$\begin{aligned} \dim H_0 - \dim H_1 + \dim H_2 - \cdots \pm \dim H_n \\ = \dim C_0 - \dim C_1 + \cdots \pm \dim C_n, \end{aligned}$$

которую мы уже проверили в нескольких частных случаях. Доказательство состоит из очень простых шагов.

- (1) Поскольку  $H_k = Z_k/B_k$ , то  $\dim H_k = \dim Z_k - \dim B_k$ .
- (2) Если теорему о ранге и дефекте применить к  $\partial : C_{k-1} \rightarrow C_k$ , то  $\dim C_{k+1} = \dim Z_{k+1} + \dim B_k$ .
- (3) Для  $n$ -комплекса  $\dim B_n = 0$ , потому что не существует объектов размерности  $n+1$ .
- (4)  $\dim Z_0 = \dim C_0$ , потому что граница любой ячейки нулевой размерности по определению равна нулю, а в комплексе нет объектов отрицательной размерности.

Собирая вместе (1) и (2), получаем

$$\dim H_k = \dim Z_k - \dim C_{k+1} + \dim Z_{k+1}$$

для  $1 \leq k < n$ . Для  $k = 0$  в силу (4) мы имеем

$$\dim H_0 = \dim C_0 - \dim C_1 + \dim Z_1,$$

а в силу (3) для  $k = n$  имеем

$$\dim H_n = \dim Z_n.$$

Таким образом,

$$\begin{aligned} \dim H_0 - \dim H_1 + \dim H_2 + \cdots \pm \dim H_n \\ = (\dim C_0 - \dim C_1 + \dim Z_1) - (\dim Z_1 - \dim C_2 + \dim Z_2) \\ + (\dim Z_2 - \dim C_3 + \dim Z_3) - \cdots \pm \dim Z_n \\ = \dim C_0 - \dim C_1 + \dim C_2 - \cdots \pm \dim C_n. \end{aligned}$$

В качестве применения этого результата рассмотрим выпуклый многогранник, который будем рассматривать как комплекс, образованный в результате разрезания поверхности сферы на отдельные ячейки. Если мы считаем, что согласно теореме, упомянутой раньше, размерности пространств  $H_0$ ,  $H_1$  и  $H_2$  такие же, как у любого комплекса, образованного разбиением данной поверхности, то можно сделать вывод, что  $\dim H_0 - \dim H_1 + \dim H_2 = 1 - 0 + 1 = 2$ . Этот результат мы уже получили раньше для куба и тетраэдра. Таким образом, для любого выпуклого многогранника

$$\dim C_0 - \dim C_1 + \dim C_2 = 2$$

или «(число вершин) – (число ребер) + (число граней) = 2».

### Ориентация и оператор $\partial$

Как уже было сказано, мы не дали строгого определения понятия пространства, лежащего в основе комплекса. Здесь мы попытаемся изложить, как следует понимать данную ситуацию. Каждая  $k$ -ячейка, т. е. каждый элемент пространства  $S_k$ , следует понимать как связный выпуклый многогранник в пространстве  $\mathbb{R}^k$  с определенной ориентацией. (Напоминаем, что в пространстве  $\mathbb{R}^k$  и, следовательно, в любом его открытом подмножестве мы можем выбирать одну из двух ориентаций, фиксируя некоторый упорядоченный базис.) Теперь границей этого выпуклого многогранника будет объединение конечного множества выпуклых многогранников размерности  $k - 1$ , лежащих в некотором аффинном подпространстве размерности  $k - 1$ . Выбрав начало координат и базис, мы можем отождествить каждое такое подпространство с  $\mathbb{R}^{k-1}$ . Предположим, что в итоге таких идентификаций в нашем комплексе каждый  $(k - 1)$ -мерный многогранник соответствует  $(k - 1)$ -ячейке, т. е. элементу пространства  $S_{k-1}$ . Но эти элементы пространства  $S_{k-1}$  имеют определенную ориентацию. Поэтому мы должны увидеть, как ориентация  $k$ -ячейки связана с ориентацией  $(k - 1)$ -ячейки на ее границе. Такая информация закладывается в оператор  $\partial$ .

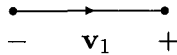
Пусть  $\mathbf{v}_1, \dots, \mathbf{v}_k$  является базисом пространства  $\mathbb{R}^k$ , так что  $\mathbf{v}_2, \dots, \mathbf{v}_k$  касателен к выбранной  $(k - 1)$ -мерной грани нашей

$k$ -ячейки  $C$ . Кроме того, вектор  $\mathbf{v}_1$  направлен из  $k$ -ячейки наружу. Предположим, что ориентация пространства  $\mathbb{R}^k$ , определенная векторами  $\mathbf{v}_1, \dots, \mathbf{v}_k$ , совпадает с ориентацией  $C$ . Тогда  $\mathbf{v}_2, \dots, \mathbf{v}_k$  определяют ориентацию конкретной  $(k-1)$ -мерной грани, соответствующей, скажем,  $F \subset S_{k-1}$ . Эта ориентация может как совпадать, так и не совпадать с ориентацией  $F$ . Если ориентации совпадают, то мы приписываем подпространству  $F$  знак  $+$ , в противном случае — знак  $-$ . Тогда

$$\partial C = \sum \pm F,$$

где сумма берется по тем  $F$ , которые соответствуют граничным  $(k-1)$ -граням  $C$ . Конечно, если базис  $\mathbf{v}_1, \dots, \mathbf{v}_k$  соответствует противоположной ориентации  $C$ , то знак меняется на противоположный.

Если  $k \geq 2$ , то вектор  $\mathbf{v}_k$  можно заменить на  $-\mathbf{v}_k$ , если нам требуется получить «хороший» базис. При  $k = 1$  пространство  $\mathbb{R}^{k-1}$  есть нулевое векторное пространство. Здесь «ориентация» означает выбор знака  $+$  или  $-$ . Согласно нашим правилам, мы выбираем знак  $+$ , если вектор  $\mathbf{v}_1$ , показывающий ориентацию, показывает наружу, и знак  $-$  в противном случае:



Конечно, каждая  $(k-1)$ -ячейка может появиться в границах нескольких  $k$ -ячеек. Но поскольку  $k$ -ячейки образуют базис  $C$ , то, как мы показывали раньше, каждая  $k$ -ячейка определяет линейное отображение  $\partial : C_k \rightarrow C_{k-1}$ .

Чтобы проверить равенство  $\partial\partial = 0$ , достаточно проверить, что для каждой  $k$ -ячейки выполнено  $\partial(\partial C) = 0$ . Теперь каждая  $(k-2)$ -ячейка  $E$ , которая может появиться в выражении  $\partial(\partial C)$ , действительно там появляется, потому что это граница для двух  $(k-1)$ -мерных граней ячейки  $C$ , скажем, для  $F_1$  и  $F_2$ . Читателю рекомендуется самому убедиться в справедливости равенства  $\partial(\partial C) = 0$ , проверив свое понимание понятия ориентации и определения оператора  $\partial$ .

Рассматривая каждый элемент  $S_k$  как выпуклый многогранник, мы можем определить гладкое отображение комплекса в целом в пространство  $\mathbb{R}^N$ . Это сводится к правилу, согласно которому  $f$  приписывает каждой точке  $k$ -ячейки точку в пространстве  $\mathbb{R}^N$ . Если отождествить тем или иным образом  $k$ -ячейку  $C$  с выпуклым многогранником в пространстве  $\mathbb{R}^N$ , то отображение  $f$ , ограниченное на  $C$ , можно рассматривать как отображение из выпуклого многогранника в  $\mathbb{R}^k$  в пространство  $\mathbb{R}^N$ . Мы хотим, чтобы это отображение было гладким. По существу мы хотим работать с отображением  $g$ , определенном в большей области, где оно гладко (так, чтобы включить внутрь области  $C$  и все его границы). Мы подробно займемся этим в главе 15.

## 14.2. Дуальные пространства и когомологии

А теперь займемся семейством дуальных пространств к пространствам  $C_k$ . Эти пространства будем обозначать  $C^k$ . Элемент такого пространства  $C^k$ , называемый  $k$ -коцепью, является линейной функцией элементов пространства  $C_k$ , которые мы называем  $k$ -цепями. Если  $\sigma$  является  $k$ -коцепью и  $c$  —  $k$ -цепь, то значение коцепи  $\sigma$  на цепи  $c$  будем обозначать  $\int_c \sigma$ , точно также, как мы делали в случаях  $k = 0$  и  $k = 1$  при рассмотрении электрических цепей.

Рассмотрим транспонированное преобразование  $d$  к граничному оператору  $\partial : C_{k+1} \rightarrow C_k$ , которое мы будем называть кограничным оператором. Итак,  $d : C^k \rightarrow C^{k+1}$  определяется формулой

$$\int_c d\sigma = \int_{\partial c} \sigma, \quad \text{где } \sigma \in C^k \text{ и } c \in C_{k+1}.$$

Из того факта, что граница границы равняется нулю, следует, что  $d \circ \partial = 0$ . (Конечно, имеется в виду композиция  $d : C^k \rightarrow C^{k+1}$  с  $\partial : C^{k+1} \rightarrow C^k$ .) В только что введенных обозначениях можно написать

$$\int_c d(d\sigma) = \int_{\partial c} d\sigma = \int_{\partial(\partial c)} \sigma = 0, \quad \text{поскольку } \partial(\partial c) = 0.$$

Для того, чтобы почувствовать смысл коцепи и важность кограничного оператора  $d$ , рассмотрим 2-комплекс, изображенный на рис. 14.16. Элемент  $C^0$ , т. е. 0-коцепь  $\Phi$ , полностью определяется заданием значений в каждом узле. Например,

$$\Phi^A = 1, \quad \Phi^B = 3, \quad \Phi^C = 2, \quad \Phi^D = 5.$$

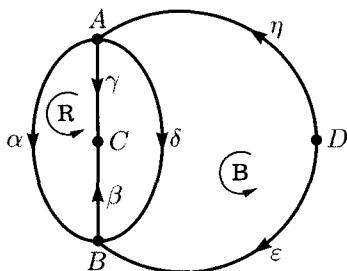


Рис. 14.16

Потенциальные функции, которые играют огромную роль в теории электрических цепей, являются примерами 0-коцепей.

Действие кограничного оператора  $d$  на 0-коцепь дает 1-коцепь  $d\Phi$ . Чтобы вычислить значение  $d\Phi$  на ветви, мы воспользуемся тем, что  $d$  сопряжен к  $\partial$ . Например, поскольку  $\partial\alpha = B - A$ , то  $\int_{\alpha} d\Phi = \int_{\partial\alpha} \Phi = \Phi^B - \Phi^A$ . Как нам уже известно, оператор  $d: C^0 \rightarrow C^1$  преобразует потенциальную функцию, являющуюся 0-коцепью или линейной функцией, заданной в узлах, в функцию скачка напряжения, т. е. в 1-коцепь, определенную для ветви как линейная функция.

Продолжим исследовать смысл оператора  $d: C^1 \rightarrow C^2$ . Примером 1-коцепи может служить линейная функция  $\mathbf{W}$ , принимающая следующие значения:

$$W^{\alpha} = 1, \quad W^{\beta} = 2, \quad W^{\gamma} = 1 \quad W^{\delta} = 2, \quad W^{\epsilon} = -3 \quad W^{\eta} = 2.$$

Напряжения на батареях в одном из примеров электрической цепи определяли такую 1-коцепь.

Действие оператора  $d$  на 1-коцепь  $\mathbf{W}$  должно давать 2-коцепь  $d\mathbf{W}$ , которая является линейной функцией 2-цепей. Мы можем

вычислить значение  $d\mathbf{W}$  на 2-ячейке Красный, используя определение

$$\int_{\text{Красный}} d\mathbf{W} = \int_{\partial(\text{Красный})} \mathbf{W}.$$

Поскольку  $\partial(\text{Красный}) = \alpha + \beta - \gamma$ , мы имеем

$$\int_{\text{Красный}} d\mathbf{W} = W^\alpha + W^\beta - W^\gamma = 1 + 2 - 1 = 2.$$

Аналогично

$$\int_{\text{Синий}} d\mathbf{W} = W^\delta - W^\varepsilon + W^\eta = 2 + 3 + 2 = 7.$$

Теперь известны значения  $d\mathbf{W}$  в базисе пространства  $C_2$ . Но тогда из линейности могут быть получены ее значения на любом элементе пространства  $C_2$ .

В качестве последнего примера, рассмотрим 3-комплекс тетраэдра, показанный на рис. 14.17. Пусть  $R$ ,  $Y$ ,  $G$  и  $B$  обозначают четыре грани тетраэдра, причем все они ориентированы против часовой стрелки, если смотреть на них снаружи.

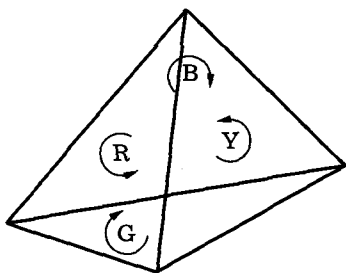


Рис. 14.17

Тогда 2-коцепь  $\mathbf{T}$  определяется ее значениями на гранях, т. е.  $T^R = 1$ ,  $T^Y = 2$ ,  $T^G = 4$ ,  $T^B = -1$ . Действие оператора  $d : C^2 \rightarrow C^3$  на 2-коцепь  $\mathbf{T}$  дает 3-коцепь  $d\mathbf{T}$ , определенную на элементах пространства  $C_3$ . Нам надо только вычислить значение  $d\mathbf{T}$  для сплошного тетраэдра. Если тетраэдр имеет правостороннюю ориентацию, так что

$$\partial(\text{Тетраэдр}) = R + Y + B + G,$$

то

$$\int_{\text{Тетраэдр}} d\mathbf{T} = T^R + T^Y + T^G + T^B = 1 + 2 + 4 - 1 = 6.$$

Читатель, знакомый с электромагнитной теорией, возможно, уже заметил сходство между кограничным оператором и некоторыми операциями физического происхождения. Например, 2-коцепь  $\mathbf{T}$  может рассматриваться как поток электрического поля через грани тетраэдра. В этом случае 3-коцепь  $d\mathbf{T}$  по закону Гаусса будет пропорциональна полному электрическому заряду внутри тетраэдра. В предыдущем примере, если 1-коцепь  $\mathbf{W}$  представляет напряжения, индуцированные изменением магнитного поля, то 2-коцепь  $d\mathbf{W}$  по закону Фарадея будет пропорциональна скорости изменения магнитного потока через соответствующие 2-ячейки. Действительно, кограничный оператор  $d$  преобразуется в дифференциальный оператор, который позволяет дать точные формулировки всех уравнений Максвелла.

Чтобы выделить подпространства пространств  $C^k$ , рассмотрим образ и ядро кограничного оператора. Говоря конкретнее, образ  $d : C^{k-1} \rightarrow C^k$  называется подпространством *кограниц* и обозначается  $B^k$ , а ядро  $d : C^k \rightarrow C^{k+1}$  называется подпространством *коциклов* и обозначается  $Z^k$ . Поскольку  $d \circ d = 0$ , любая кограница обязательно является коциклом, и тем самым  $B^k$  является подпространством  $Z^k$ . Кроме того, мы образуем факторпространство  $H^k = Z^k/B^k$ , называемое *пространством  $k$ -мерных когомологий* комплекса.

Многие аспекты этих пространств уже проявлялись при изучении электрических цепей. Например, в случае  $k = 0$  пространство  $Z^0$  (пространство 0-коциклов) было пространством физически осмысленных потенциалов, которые были постоянны на каждой связной компоненте цепи. Пространство  $B^0$  было нулевым, поскольку 0-коцепь никоим образом не могла принадлежать образу  $d$ . В результате получалось факторпространство  $H^0 = Z^0/B^0 = Z^0$ . (Раньше при обсуждении этого пространства мы использовали только обозначение  $H^0$ .)

Для 1-комплекса пространство  $Z^1$  совпадает со всем пространством  $C^1$ . Причина этого очень проста и состоит в том, что при отсутствии 2-коцепей каждая 1-коцепь должна лежать

в ядре  $d$ . Пространство кограниц  $B^1$  содержит те распределения напряжений, которые получаются из потенциала. В этом случае факторпространство  $H^1 = Z^1/B^1$  совпадает с пространством  $C^1/B^1$ . Вспомните, что пространство  $H^1$  дуально пространству  $H_1$  (в случае 1-комплекса оно совпадает с  $Z_1$ ).

Простейшим примером ситуации, когда  $Z^k$ ,  $B^k$  и  $H^k$  не тривиальны, является случай  $k = 1$  для 2-комплекса. Рассмотрим комплекс, изображенный на рис. 14.18. Пространство  $C^1$  шестимерно, потому что в комплексе шесть ветвей. Чтобы определить размерность пространства  $Z^1$ , мы должны посмотреть на образ и ядро оператора  $d : C^1 \rightarrow C^2$ .

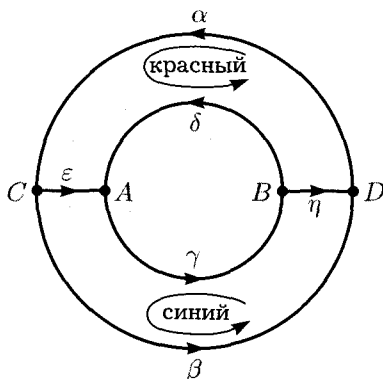


Рис. 14.18

Сначала рассмотрим образ  $d : C^1 \rightarrow C^2$ , т.е. пространство  $B^2$ . Поскольку пространство  $C^2$  двумерно, то  $B^2$  не может иметь размерность больше, чем 2. Чтобы показать, что его размерность равна 2, достаточно найти базис  $C^2$ , который лежит в образе  $d$ . 1-коцепь  $\mathbf{V}$ , которая приписывает значение 1 ветви  $\alpha$  и 0 другим ветвям, обладает свойством:

$$d\mathbf{V}(\text{Красный}) = 1, \quad d\mathbf{V}(\text{Синий}) = 0.$$

Аналогично, 1-коцепь  $\mathbf{W}$ , которая приписывает 1 ветви  $\beta$  и 0 другим ветвям, обладает свойством:

$$d\mathbf{W}(\text{Красный}) = 0, \quad d\mathbf{W}(\text{Синий}) = 1.$$

Поэтому  $B^2$ , т.е. образ  $d$ , заданный векторами  $dV$  и  $dW$ , имеет размерность 2. В силу теоремы о ранге и дефекте, ядро  $d$ , т.е. пространство  $Z^1$ , удовлетворяет условию

$$\dim Z^1 = \dim C^1 - \dim B^2 = 6 - 2 = 4.$$

Обратим теперь внимание на пространство  $B^1$  и рассмотрим  $d : C^0 \rightarrow C^1$ . Ядро этого оператора, т.е. пространство  $Z^0$  (которое совпадает с  $H^0$ ), имеет размерность 1, потому что комплекс связан. Следовательно,

$$\dim B^1 = \dim C^0 - \dim Z^0 = 4 - 1 = 3.$$

Значит, существуют три независимых 1-коцепи, которые выводятся из потенциалов.

Мы видим, что пространство  $Z^1$ , состоящее из 1-коциклов, имеет размерность 4, а пространство  $B^1$ , состоящее из 1-кограниц, имеет размерность 3. Поэтому  $\dim H^1 = \dim Z^1 - \dim B^1 = 4 - 3 = 1$ . Это значит, что существуют «интересные» коциклы, которые лежат в ядре  $d$ , хотя они и не могут быть выведены из потенциала. Класс эквивалентности одного такого коцикла дает базис одномерного факторпространства  $Z^1$ . Давайте, например, выберем коцикл  $V$ , для которого  $V^\alpha = \frac{1}{2}$ ,  $V^\beta = \frac{1}{2}$ ,  $V^\gamma = \frac{1}{2}$ ,  $V^\delta = \frac{1}{2}$ ,  $V^\varepsilon = 0$ ,  $V^\eta = 0$ . Очевидно, что этот коцикл не может быть выведен из потенциала, потому что сумма падений напряжений при обходе цикла  $\alpha + \beta$  или  $\gamma + \delta$  не равна нулю. С другой стороны,  $V$  принадлежит ядру  $d : C^1 \rightarrow C^2$ , потому что  $dV(\text{Красный}) = V^\alpha - V^\delta = \frac{1}{2} - \frac{1}{2} = 0$  и  $dV(\text{Синий}) = V^\beta - V^\gamma = \frac{1}{2} - \frac{1}{2} = 0$ .

Из этого анализа очевидно, что существование «интересных» коциклов тесно связано с наличием дырок в комплексе. Если добавить к комплексу диск с ветвями  $\gamma$  и  $\delta$  (назовем его Жёлтый), то кограницами будут только коциклы. Например, коцепь  $V$ , рассмотренная выше, будет удовлетворять условию  $dV(\text{Жёлтый}) = 1$  и не будет коциклом. Именно таким образом пространства когомологий комплекса показывают наличие дырки, точно так же, как анализ пространств гомологий.

В этом примере сразу же видно, что пространство  $H^1$  дуально  $H_1$  в том смысле, что базисный элемент  $V$  приписывает

1 любому циклу, который обходит дырку против часовой стрелки. Рассмотрим, например, цикл  $\mathbf{I} = \alpha + \varepsilon - \delta + \eta$ . Мы видим, что  $\int_{\mathbf{I}} \mathbf{V} = V^\alpha + V^\varepsilon - V^\delta + V^\eta = \frac{1}{2} - \frac{1}{2} = 0$ . Однако, все циклы, обходящие дырку один раз против часовой стрелки, образуют класс эквивалентности, который может быть базисом  $H_1$ . Кроме того, конкретный выбор элемента  $\mathbf{V}$  в качестве представителя этого класса эквивалентности не имеет значения. Предположим, что мы выбрали  $\mathbf{W} = \mathbf{V} + d\Phi$ . Тогда для любого цикла  $\mathbf{M}$  выполняется условие

$$\int_{\mathbf{M}} \mathbf{W} = \int_{\mathbf{M}} \mathbf{V} + \int_{\mathbf{M}} d\Phi = \int_{\mathbf{M}} \mathbf{V} + \int_{\partial\mathbf{M}} \Phi = \int_{\mathbf{M}} \mathbf{V},$$

поскольку  $\partial\mathbf{M} = 0$ , если  $\mathbf{M}$  — цикл. В этом смысле класс эквивалентности  $\bar{\mathbf{V}}$  (базис для  $H^1$ ) определяет линейную функцию от класса эквивалентности  $\bar{\mathbf{I}}$  (базис для  $H_1$ ). Для вычисления этой функции находим  $\int_{\mathbf{I}} \mathbf{V}$ , взяв *любой* элемент класса эквивалентности.

Попробуем доказать в общем случае, что  $H^k$  — это дуальное пространство  $H_k$ . Сначала установим, что  $\dim H^k = \dim H_k$ , потом покажем, что любой элемент  $H^k$  определяет линейную функцию в пространстве  $H_k$ .

Результат получается из того, что мы уже знаем. Ядро сопряженного оператора аннулирует образ; образ сопряженного оператора аннулирует ядро. Рассмотрим следующую диаграмму:

$$C^{k-1} \xrightarrow{d} C^k \xrightarrow{d} C^{k+1}$$

$$C_{k-1} \xleftarrow{\partial} C_k \xleftarrow{\partial} C_{k+1}$$

Сначала посмотрим на  $\partial : C_{k+1} \rightarrow C_k$  и его сопряженный оператор  $d : C^k \rightarrow C^{k+1}$ . Ядро сопряженного оператора  $d$  является пространством коциклов  $Z^k$ ; оно аннулирует образ  $\partial$ , которое является пространством границ  $B_k$ . А теперь рассмотрим  $\partial : C_k \rightarrow C_{k-1}$  и его сопряженный  $d : C^{k-1} \rightarrow C^k$ . Образ сопряженного оператора есть пространство кограниц  $B^k$ ; он аннулирует ядро  $\partial$ , которое является пространством циклов  $Z_k$ .

Однако, мы знаем, что для любого подпространства пространства  $C_k$  сумма размерности подпространства и размерности его

аннулирующего пространства должна быть равна размерности  $C_k$ . Следовательно,

$$\dim B_k + \dim Z^k = \dim C_k = \dim Z_k + \dim B^k.$$

Отсюда следует, что

$$\dim Z^k - \dim B^k = \dim Z_k - \dim B_k,$$

т. е.

$$\dim H^k = \dim H_k,$$

поскольку

$$H^k = Z^k/B^k \quad \text{и} \quad H_k = Z_k/B_k.$$

Типичным элементом пространства  $H^k$  является класс эквивалентности  $\bar{\mathbf{V}}$ , состоящий из коциклов вида  $\mathbf{V} + d\Phi$ . Если  $\bar{\mathbf{V}}$  не является нулевым элементом, то  $\mathbf{V}$  — не кограница. Аналогично, типичным элементом пространства  $H_k$  является класс эквивалентности  $\bar{\mathbf{I}}$ , состоящий из циклов вида  $\mathbf{I} + \partial\tau$ . Если  $\bar{\mathbf{I}}$  — не нулевой элемент, то  $\mathbf{I}$  — не граница. Мы определяем значение  $\bar{\mathbf{V}}$ , действующего на  $\bar{\mathbf{I}}$ , т. е.  $\int_{\bar{\mathbf{I}}} \bar{\mathbf{V}}$ , давая возможность любому элементу класса  $\bar{\mathbf{V}}$  действовать на любой элемент класса  $\bar{\mathbf{I}}$ . Тогда

$$\begin{aligned} \int_{\bar{\mathbf{I}}} \bar{\mathbf{V}} &= \int_{\mathbf{I} + \partial\tau} \mathbf{V} + d\Phi = \int_{\mathbf{I}} \mathbf{V} + \int_{\mathbf{I}} d\Phi + \int_{\partial\tau} \mathbf{V} + \int_{\partial\tau} d\Phi \\ &= \int_{\mathbf{I}} \mathbf{V} + \int_{\partial\mathbf{I}} \Phi + \int_{\tau} d\mathbf{V} + \int_{\tau} d(d\Phi). \end{aligned}$$

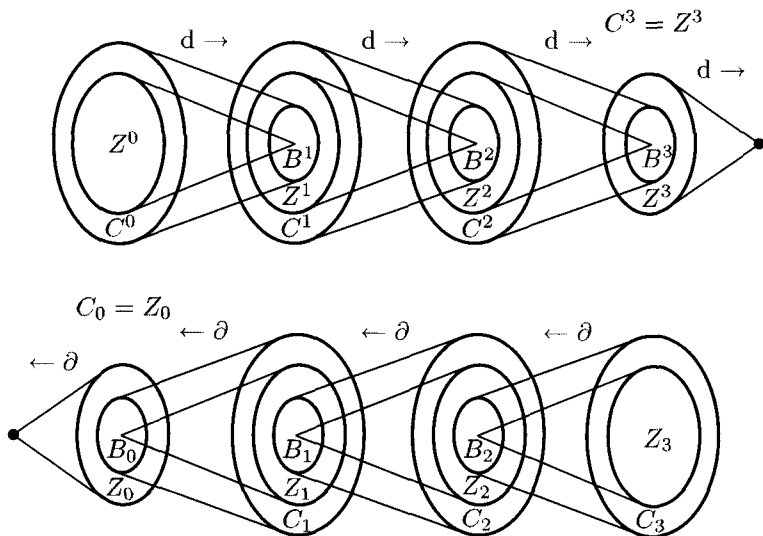
Однако,  $\partial\mathbf{I} = 0$ , потому что  $\mathbf{I}$  — цикл;  $d\mathbf{V} = 0$ , потому что  $\mathbf{V}$  — коцикл и, конечно,  $d(d\Phi) = 0$ . Следовательно,

$$\int_{\bar{\mathbf{I}}} \bar{\mathbf{V}} = \int_{\mathbf{I}} \mathbf{V}.$$

Другими словами, не имеет значения, какой элемент класса эквивалентности мы выбираем. Мы имеем корректное определение элементов  $H^k$  как линейных функций в пространстве  $H_k$ . Таким образом, мы произвели отождествление пространства  $H^k$  с дуальным пространством  $H_k$ .

Ранее говорилось, что пространства гомологий  $H_k$ , относящиеся к комплексу, зависят только от пространства, лежащего в основе комплекса, и не зависят от того, как комплекс был образован. Теперь понятно, что то же самое справедливо для пространств когомологий  $H^k$ . Поскольку коцепи, из которых состоят  $H^k$ , тесно связаны с физическими величинами, то эти пространства открывают связь топологии и теории электромагнетизма.

Прежде чем закончить рассмотрение комплексов, имеет смысл подвести итоги того, что мы изучили, на диаграмме, изображенной на рис. 14.19. Диаграмма нарисована для 3-комплекса, на ней показаны все связи между пространствами, которые мы изучили.



$Z^0$  аннулирует  $B_0$

$B^0$  нулевое

$H^0 = Z^0$  дуально  $H_0 = C_0/B_0$

$Z^2$  аннулирует  $B_2$

$B^2$  аннулирует  $Z_2$

$H^2 = Z^2/B^2$  дуально  $H_2 = Z_2/B_2$

$Z^1$  аннулирует  $B_1$

$B^1$  аннулирует  $Z_1$

$H^1 = Z^1/B^1$  дуально  $H_1 = Z_1/B_1$

$B_3$  нулевое

$B^3$  аннулирует  $Z_3$

$H^3 = C^3/B^3$  дуально  $H_3 = Z_3$

Рис. 14.19

## Резюме

### А. Комплексы высших размерностей

Вы должны знать определения пространств  $Z_k$ ,  $B_k$  и  $H_k$  для  $n$ -комплексов и уметь строить базисы этих пространств.

Для заданной диаграммы, представляющей ячейки 2-комплекса, надо уметь определить, что это за комплекс: сфера, тор или что-то еще.

### В. Когомологии и оператор $d$

Для  $n$ -комплекса Вы должны уметь определить оператор  $d$  и пространства  $Z^k$ ,  $B^k$ ,  $H^k$ ; должны уметь строить базисы этих пространств.

Вы должны понимать, как доказать и применять дуальность пространств  $H^k$  и  $H_k$ .

## Задачи

14.1. Пять 2-комплексов, изображенных на рис. 14.20, представляют в каком-то порядке следующие фигуры:

- (a) искривленная поверхность цилиндра;
- (b) лист Мёбиуса («перекрученный» цилиндр);
- (c) сфера;
- (d) тор;
- (e) бутылка Клейна.

Опознайте эти фигуры и вычислите для каждой из них  $\dim H_1$  и  $\dim H_2$ . В каждом случае найдите базисы пространств  $Z_1$ ,  $B_1$  и  $H_1$ . Проверьте, что

$$\dim C_0 - \dim C_1 + \dim C_2 = \dim H_0 - \dim H_1 + \dim H_2.$$

14.2. Для каждого комплекса из предыдущего упражнения постройте базисы пространств  $Z^1$ ,  $B^1$  и  $H^1$ .

14.3. Резиновая надувная игрушка разделена на три части, окрашенные в разные цвета. На рис. 14.21 показаны две из этих областей.

- (a) Определите  $\partial(\text{Красный})$  и  $\partial(\text{Синий})$ .
- (b) Определите размерности пространств  $Z_2$ ,  $H_2$  и  $H_0$ .

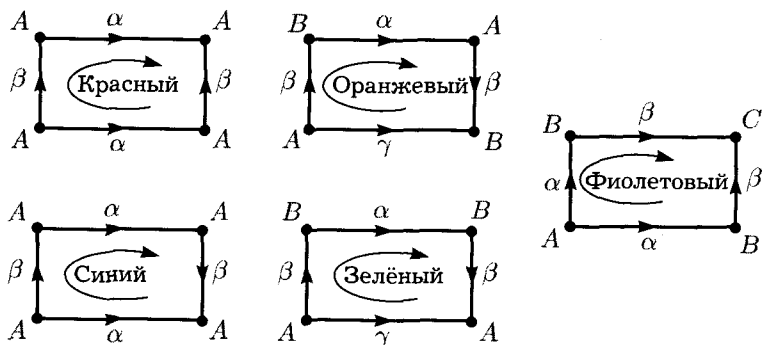


Рис. 14.20

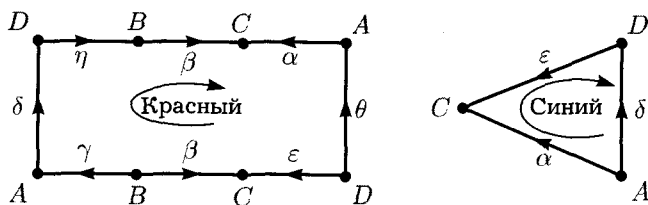


Рис. 14.21

- (с) Найдите базисы пространств  $Z_1$ ,  $B_1$  и  $H_1$ .
- (d) На рис. 14.22 показана третья область. Каковы будут размерности пространств  $C_0$ ,  $C_1$ ,  $C_2$ ,  $H_1$ ,  $H_2$  после того, как вся игрушка собрана? Получится ли мячик ( $\dim H_0 = 1$ ,  $\dim H_1 = 0$ ,  $\dim H_2 = 1$ ), плавательный круг ( $\dim H_0 = 1$ ,  $\dim H_1 = 2$ ,  $\dim H_2 = 1$ ) или что-то другое?

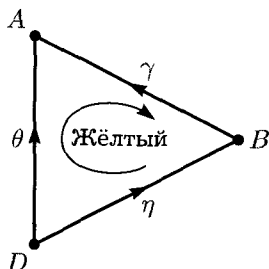


Рис. 14.22

14.4. Трое рабочих завода, выпускающего автомобильные шины, разделили поверхность внутренней камеры на три ориентированных области и подписали каждую своим именем. Получившийся 2-комплекс показан на рис. 14.23.

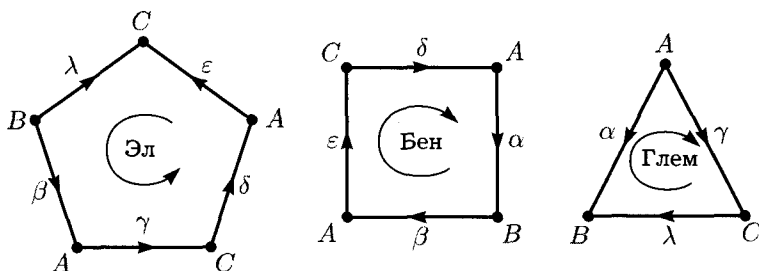


Рис. 14.23

- Какова размерность пространства  $H_0$ ?
- Определите размерность пространства  $H_2$  и выпишите базис этого пространства (2-ячейки имеют имена Эл, Бен и Глем; постройте соответствующие линейные комбинации, следя за ориентацией). Что означает полученный результат для собранной конструкции?
- Определите размерности  $\dim C_0$ ,  $\dim C_1$  и  $\dim C_2$ , потом вычислите  $\dim H_1$ .
- Выпишите базисы пространств  $Z_1$ ,  $B_1$  и базис факторпространства  $H_1 = Z_1/B_1$ .
- Ниже выписаны три 1-коцепи, т. е. элементы  $C^1$ :

$$\begin{array}{lll}
 \sigma(\alpha) = 1 & \tau(\alpha) = 1 & \omega(\alpha) = 1 \\
 \sigma(\beta) = 1 & \tau(\beta) = -1 & \omega(\beta) = -1 \\
 \sigma(\gamma) = 2 & \tau(\gamma) = 2 & \omega(\gamma) = 2 \\
 \sigma(\delta) = -2 & \tau(\delta) = 2 & \omega(\delta) = -2 \\
 \sigma(\varepsilon) = -2 & \tau(\varepsilon) = -2 & \omega(\varepsilon) = 2 \\
 \sigma(\lambda) = 1 & \tau(\lambda) = -1 & \omega(\lambda) = -1
 \end{array}$$

Найдите коцепь, которая не является коциклом, и определите 2-коцепь, которая является ее кограницей ( $d\Omega(\text{Эл}) = 1$ ,  $d\Omega(\text{Бен}) = 3$ ,  $d\Omega(\text{Глем}) = 2$ ).

- Определите коцепь, являющуюся кограницей, и напишите 0-цепь  $f$ , чьей кограницей она будет.

- (g) Покажите, что оставшаяся коцепь является коциклом, но не будет кограницей. Вычислите ее для двух базисных элементов пространства  $H_1$ .
- (h) Постройте второй коцикл, который не является кограницей, но будет вторым базисным элементом пространства  $H^1$ .
- (i) Напишите базисы пространств  $B^1$  и  $Z^1$ .

14.5. Возрожденная Римская империя послала уполномоченного представителя Гая Юлия Цезаря обследовать ситуацию на международной космической станции «Новая Галлия». В его отчете, начинавшемся словами: «Вся Новая Галлия разделена на три части», были три карты, изображенные на рис. 14.24. Карты показывают территории, контролируемые аквитанцами (Aqu), белгами (Bel) и кельтами (Cel) соответственно. Эти враждующие племена не смогли договориться о единой (непротиворечивой) ориентации всей поверхности этой станции. Дайте ответы на следующие вопросы относительно комплекса «Новой Галлии».

- (a) Вычислите границу каждого региона:  $\partial(\text{Aqu})$  и т. д.

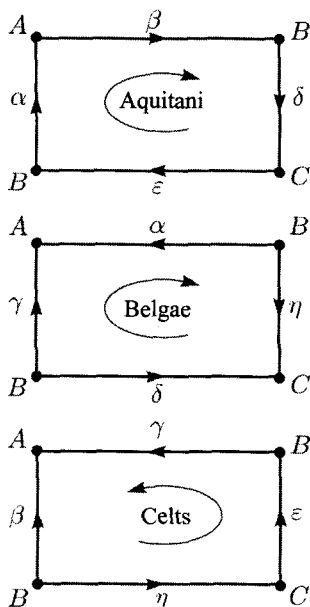


Рис. 14.24

- (b) Найдите максимальное дерево. С помощью этого дерева вычислите  $\dim Z_1$ ,  $\dim B_1$  и  $\dim H_1$ . Постройте базис  $H_1$ . В этом базисе выразите  $\alpha - \gamma + \delta - \eta$  и  $\beta + \gamma$ .
- (c) Если  $\omega$  есть 1-кограница (элемент  $B^1$ ), удовлетворяющая условиям  $\omega^\alpha = 2$  и  $\omega^\delta = 1$ , то каковы значения  $\omega^\gamma$  и  $\omega^\eta$ ?
- (d) Если  $\tau$  есть 1-коцепь, значения которой для каждой ветви равны 1, то чему равны  $d\tau(\text{Aqu})$  и  $d\tau(\text{Bel})$ ?
- (e) Вычислите  $\dim Z^1$ . Если  $\sigma$  есть 1-цикл (элемент  $Z_1$ ), удовлетворяющий условиям  $\sigma^\alpha = 3$ ,  $\sigma^\beta = 1$ ,  $\sigma^\delta = 2$ ,  $\sigma^\eta = 4$ , то каковы значения  $\sigma^\gamma$  и  $\sigma^\epsilon$ ?
- (f) Постройте базис  $H^1$ , дуальный к базису  $H_1$  из пункта (b). В этом базисе выразите класс эквивалентности  $\bar{\sigma}$ .
- (g) Может ли поверхность «Новой Галлии» быть сферой? цилиндром? тором?

14.6. На небольшой планете живет колония, разделенная на две провинции, Красный и Зелёный, как показано на рис. 14.25.

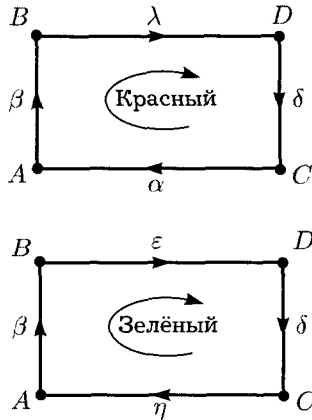


Рис. 14.25

- (a) Вычислите  $\partial(\text{Красный})$  и  $\partial(\text{Зелёный})$ . С помощью результата определите размерность  $H_2$  для этого 2-комплекса. Объясните полученный результат и скажите: занимает ли колония всю поверхность планеты или нет?
- (b) Максимальное дерево для этого комплекса состоит из ветвей  $\alpha$ ,  $\beta$  и  $\delta$ . Найдите базисы пространств  $Z_1$ ,  $B_1$  и  $H_1$ . Дайте

географическое объяснение базисному элементу пространства  $H_1$ .

- (с) Определите пространство  $B^1$ . Какое пространство оно будет аннулировать? Докажите свой ответ. Рассмотрите 1-коцепь  $\mathbf{W}$ , для которой  $W^\alpha = 2$ ,  $W^\beta = -1$ ,  $W^\delta = 1$ .
- (d) Найдите такие значения  $W^\gamma$ ,  $W^\varepsilon$ ,  $W^\eta$ , чтобы  $\mathbf{W}$  было кограницей.
- (e) Пусть  $W^\gamma = 3$ ,  $W^\varepsilon = 2$ ,  $W^\eta = 1$ . Вычислите  $d\mathbf{W}$ .
- (f) Найдите такие значения  $W^\gamma$ ,  $W^\varepsilon$ ,  $W^\eta$ , чтобы  $\mathbf{W}$  было коциклом, но не кограницей. Вычислите эту коцепь в базисе пространства  $H_1$ .
- 14.7. (a) Найдите размерности пространств:  $C_1$ ,  $C_2$ ,  $Z_1$ ,  $B_1$  и  $H_1$  для комплекса, изображенного на рис. 14.26(a).
- (b) Найдите размерности пространств  $H_0$  и  $H_2$ ; докажите, что
- $$\dim C_0 - \dim C_1 + \dim C_2 = \dim H_0 - \dim H_1 + \dim H_2. \quad (14.1)$$
- (с) Пусть к комплексу добавили новую ветвь  $\sigma$  и новый узел  $C$ , как показано на рис. 14.26(б). В явном виде напишите матрицы операторов  $\partial : C_1 \rightarrow C_0$  и  $\partial : C_2 \rightarrow C_1$  и вычислите матрицу  $\partial \circ \partial : C_2 \rightarrow C_0$ .

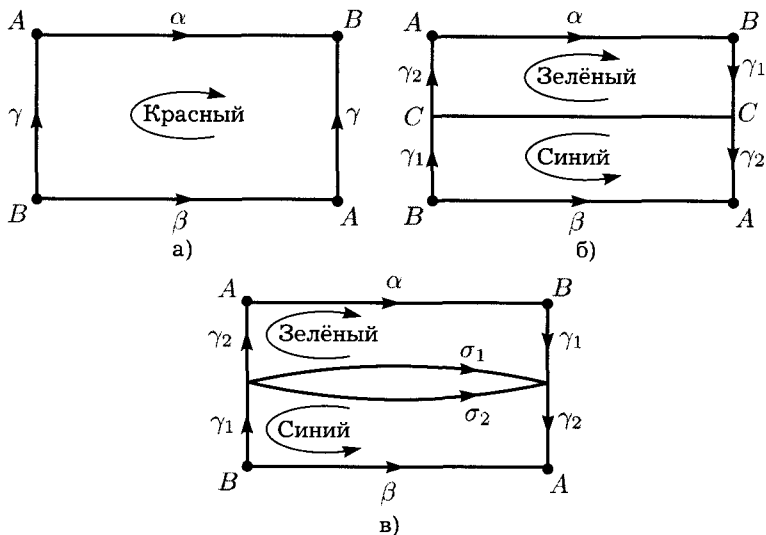


Рис. 14.26

- (d) Вдоль ветви  $\sigma$  сделан разрез, разделивший ее на  $\sigma_1$  (на стороне Зелёный) и  $\sigma_2$  (на стороне Синий), рис. 14.26(в). Напишите  $\partial : C_1 \rightarrow C_0$ . Построив базис, найдите размерность  $H_0$ . Что это говорит о данной цепи? Постройте базис пространства  $H_1$  и проверьте уравнение (14.1).
- (e) Пусть 1-коцепь  $\tau$  имеет значение 1 на всех ветвях между различными точками и значение 0 во всех остальных случаях. Чему равны  $d\tau(\text{Зелёный})$  и  $d\tau(\text{Синий})$ ?

14.8. На рис. 14.27 изображен 2-комплекс, состоящий из двух частей.

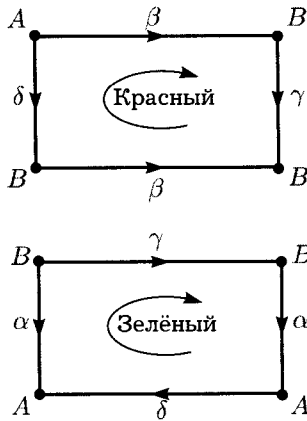


Рис. 14.27

- (a) Вычислите  $\partial(\text{Красный})$  и  $\partial(\text{Зелёный})$ , отсюда определите  $\dim Z_2$ .
- (b) Для этого комплекса найдите максимальное дерево. Используйте его для нахождения  $\dim Z_1$ ,  $\dim B_1$  и  $\dim H_1$ .
- (c) Найдите базис пространства  $H_1$ , и в этом базисе выразите  $\bar{\gamma}$  и  $\alpha + \delta + \beta$ .
- (d) Пусть  $\omega$  — это 1-кограница, удовлетворяющая условию  $\omega^\beta = 2$ . Вычислите  $\omega^\alpha$ .
- (e) Пусть 1-коцепь  $\tau$  принимает на каждой ветви значение 1. Вычислите  $d\tau$ .
- (f) Определите  $\dim B^2$  для этого комплекса.

## Глава 15

---

# Комплексы в пространстве $\mathbb{R}^n$

---

В главах 15–18 внешнее дифференциальное исчисление излагается как непрерывная версия дискретной теории комплексов. В главе 15 сформулированы основные положения внешнего исчисления: внешняя алгебра,  $k$ -формы, операция переноса (pullback), внешняя производная и теорема Стокса.

### Введение

До сих пор, рассматривая различные комплексы, мы не интересовались тем, как комплекс может быть расположен в пространстве  $\mathbb{R}^n$ . Изучая электрические цепи, мы пренебрегали формой проводов и даже не ставили вопрос, может ли данная цепь быть реализована на плоскости. Сейчас мы начнем изучать частный случай комплекса, помещенного в пространство  $\mathbb{R}^n$ , чтобы увидеть, как естественным образом из геометрических или физических соображений возникают коцепи.

Сначала обсудим 1-комплекс в пространстве  $\mathbb{R}^3$  типа изображенного на рис. 15.1. В этом пространстве узлы являются точками, а ветви — дифференцируемыми кривыми. Поскольку коцепь однозначно определена своими значениями в каждом узле комплекса, мы можем рассматривать ее как функцию узлов. Один из очевидных путей получения такой функции — это задать функцию, определенную в  $\mathbb{R}^3$ . Каждая такая функция  $\phi(x, y, z)$

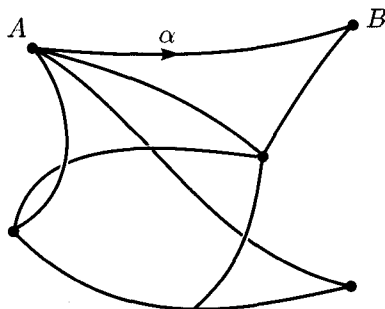


Рис. 15.1

определяет 0-коцепь  $\Phi$  согласно правилу  $\Phi^A = \phi(x^A, y^A, z^A)$ , где  $x^A, y^A, z^A$  являются координатами узла  $A$ .

Функция  $\phi$  может рассматриваться как *зародыш 0-коцепи* в том смысле, что она определяет 0-коцепь для любого комплекса, расположенного в этом пространстве.

Чтобы таким же способом получить 1-коцепь, нам надо иметь правило, по которому любой ветви комплекса приписывается число. Такое правило получается, например, из дифференциальной 1-формы  $\omega$ , определенной в пространстве  $\mathbb{R}^3$ . Тогда если каждая ветвь комплекса является дифференцируемой дугой, то, интегрируя  $\omega$  вдоль каждой ветви, мы можем построить 1-коцепь. Например, если ветвь  $\alpha$  описывается параметризованной кривой  $\alpha(t)$ , то 1-коцепь, определенная  $\omega$ , будет принимать на ветви  $\alpha$  значение  $\int_{\alpha} \omega$ . Таким образом, 1-форма  $\omega$  может рассматриваться как *зародыш 1-коцепи*. Она определяет 1-коцепь для любого комплекса, расположенного в  $\mathbb{R}^3$ .

Предположим, что  $\phi$  — непрерывно дифференцируемая функция, и мы образуем ее дифференциал, т. е. 1-форму  $\omega = d\phi$ . Предположим, что для ветви  $\alpha$  выполняется условие  $\partial\alpha = B - A$ , так

что  $\alpha$  описывается кривой  $\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$  с условиями:

$$\begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix} = \begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} x(1) \\ y(1) \\ z(1) \end{pmatrix} = \begin{pmatrix} x_B \\ y_B \\ z_B \end{pmatrix}.$$

Тогда

$$\begin{aligned}\int_{\alpha} \omega &= \int_{\alpha} d\phi = \int_0^1 \alpha^* d\phi = \int_0^1 d(\alpha^* \phi) \\ &= \alpha^* \phi(1) - \alpha^* \phi(0) = \phi(B) - \phi(A).\end{aligned}$$

Функция  $\phi$  определяет 0-коцепь  $\Phi$  с условиями  $\Phi^A = \phi(A)$  и  $\Phi^B = \phi(B)$ . Из этой 0-коцепи мы можем образовать 1-коцепь  $d\Phi$ , значение которой на ветви  $\alpha$  равно

$$\int_{\alpha} d\Phi = \int_{\partial\alpha} \Phi = \Phi^B - \Phi^A = \phi(B) - \phi(A).$$

Мы используем одно и то же обозначение  $d$  для двух различных ситуаций: во-первых, как линейный оператор, действующий на 0-коцепь  $\Phi$  и образующий 1-коцепь  $d\Phi$ ; во-вторых, как дифференциальный оператор, который действует на функцию  $\phi$  и образует дифференциальную форму  $d\phi$ . Сейчас становится очевидным, что такое использование  $d$  корректно в том смысле, что если из функции  $\phi$  получается 0-коцепь  $\Phi$ , то из 1-формы  $d\phi$  получается 1-коцепь  $d\Phi$ . В пространстве  $\mathbb{R}^3$  рассмотрим пространство гладких функций  $\Omega^0(\mathbb{R}^3)$  и пространство гладких 1-форм  $\Omega^1(\mathbb{R}^3)$ . Тогда можно построить диаграмму, отражающую непротиворечивость определения  $d$ :

$$\begin{array}{ccc} C^0 & \xrightarrow{d} & C^1 \\ \uparrow & & \uparrow \\ \Omega^0(\mathbb{R}^3) & \xrightarrow{d} & \Omega^1(\mathbb{R}^3) \end{array}$$

Эта диаграмма *коммутативна* в том смысле, что от функции  $\phi$  к 1-коцепи можно перейти как через  $C^0$ , так и через  $\Omega^1$ .

Теперь мы должны проделать нечто похожее для 2-коцепей и 3-коцепей. В следующем параграфе мы введем объекты, называемые дифференциальными 2-формами и 3-формами, из которых в результате соответствующего интегрирования получатся 2-коцепи и 3-коцепи соответственно. Надо будет определить такой дифференциальный оператор  $d$ , который преобразует 1-формы

в 2-формы и 2-формы в 3-формы. Причем это преобразование должно быть совместимым с действием оператора  $d$  на соответствующие коцепи.

В частном случае 2-комплекса, расположенного на плоскости, мы можем увидеть, как будет работать эта программа. *Зародыш 2-цепи* должен будет приписывать число каждой ориентированной области на плоскости; значения, приписываемые каждой ячейке  $Q$  этого комплекса, будут определять 2-коцепь. Сейчас мы можем перейти от 1-формы (зародыша 1-коцепи)  $\omega$  к 2-форме  $\mathbf{T}$  двумя различными способами, как это показано на рис. 15.2. Рассмотрим, например, путь  $\alpha + \beta + \gamma$ , являющийся границей ячейки  $Q$ . Один способ состоит в том, чтобы образовать 1-коцепь  $\mathbf{W}$  из  $\omega$ , вычисляя интеграл по пути вдоль каждой ветви комплекса. А уже потом подействовать оператором  $d : C^1 \rightarrow C^2$ . Это дает

$$\int_Q \mathbf{T} = \int_Q d\mathbf{W} = \int_{\partial Q} \mathbf{W} = W^\alpha + W^\beta + W^\gamma = \int_\alpha \omega + \int_\beta \omega + \int_\gamma \omega.$$

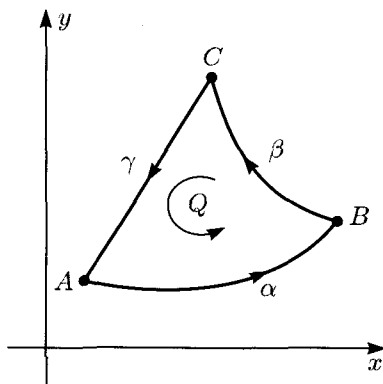


Рис. 15.2

Другой способ состоит в том, что действуя дифференциальным оператором  $d$  на  $\omega$ , есть надежда получить 2-форму  $\tau$ , которую можно проинтегрировать по области  $Q$  и получить значение 2-коцепи  $T$  на ячейке  $Q$ .

В этом случае теорема Грина на плоскости говорит нам, какую форму должен иметь оператор  $d$ . Вспомним, что согласно

этой теореме, если ориентированная область  $Q$  на плоскости имеет границу  $\partial Q$ , то интеграл по траектории<sup>1</sup>

$$\int_{\partial Q} A dx + B dy = \int_Q \left( \frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right) dx \wedge dy.$$

В этом частном случае кажется целесообразным ввести 2-формы вида  $\tau = f(x, y) dx \wedge dy$ . В результате повторного интегрирования из 2-формы получается 2-коцепь, и 2-форма может быть получена из 1-формы согласно правилу

$$d(A dx + B dy) = \left( \frac{\partial B}{\partial x} - \frac{\partial A}{\partial y} \right) dx \wedge dy.$$

## 15.1. Внешняя алгебра

В случае 0-коцепей и 1-коцепей мы уже имеем дифференциальный оператор, прекрасно согласующийся с кограничным оператором. Зародышем 0-коцепи является дифференцируемая функция  $\phi$ , которая каждой точке приписывает вещественное число. Зародышем 1-цепи является 1-форма  $\omega$ , приписывающая каждой точке элемент пространства  $V^*$ , дуального к используемому нами основному пространству  $V$ . Функция  $\phi$  принадлежит пространству  $\Omega^0(V)$ , а 1-форма  $\omega$  принадлежит пространству  $\Omega^1(V)$ . Мы определили оператор  $d : \Omega^0(V) \rightarrow \Omega^1(V)$ , согласованный с кограничным оператором.

Сейчас мы будем строить последующие пространства  $\Omega^2(V)$ ,  $\Omega^3(V)$ , ...,  $\Omega^k(V)$ , которые позволят обобщить дифференциальный оператор  $d$ . Элемент  $\omega$ , принадлежащий пространству  $\Omega^k(V)$ , называется *дифференциальной  $k$ -формой*. Каждой точке области  $\omega$  он будет приписывать элемент пространства  $\Lambda^k(V^*)$ . Обычно это называют « $k$ -я внешняя степень пространства  $V^*$ », для нас это — обобщение дуального пространства  $V^*$ . Чтобы определить  $k$ -коцепь, мы будем интегрировать  $k$ -форму. После этого мы сможем определить дифференциальный оператор

<sup>1</sup>Напомним, что в русской математической литературе он называется криволинейным интегралом второго рода. — Прим. ред.

$d : \Omega^k(V) \rightarrow \Omega^{k+1}(V)$ , который будет согласован с преобразованием комплексов  $d : C^k \rightarrow C^{k+1}$ . В качестве первого шага в этом направлении введем пространство  $\Lambda^k(V^*)$ .

Пусть  $V$  — некое векторное пространство размерности  $n$ , а  $V^*$  — его дуальное пространство. Дуальное пространство  $V^*$  является пространством линейных функций из  $V$  в  $\mathbb{R}$ , т. е.  $\mathcal{L}(V; \mathbb{R})$ . Сначала рассмотрим  $\mathcal{B}(V, V; \mathbb{R})$  — пространство *билинейных* функций из  $V \times V$  в  $\mathbb{R}$ . Мы говорим, что функция  $\tau(\mathbf{v}_1, \mathbf{v}_2)$  билинейна, если для фиксированного значения  $\mathbf{v}_2$  функция  $\tau(\mathbf{v}_1, \mathbf{v}_2)$  линейно зависит от первого аргумента, т. е.  $\tau(\alpha \mathbf{v}_1 + \beta \mathbf{w}_1, \mathbf{v}_2) = \alpha \tau(\mathbf{v}_1, \mathbf{v}_2) + \beta \tau(\mathbf{w}_1, \mathbf{v}_2)$  и, аналогично, для фиксированного значения  $\mathbf{v}_1$  функция  $\tau$  линейно зависит от второго аргумента. Пространство  $\mathcal{B}(V, V; \mathbb{R})$  является векторным пространством, где выполняются обычные правила сложения функций и умножения функции на число.

Внутри пространства функций  $\mathcal{B}(V, V; \mathbb{R})$  есть подпространство, обозначаемое  $\Lambda^2(V^*)$ . Это пространство *кососимметричных* билинейных функций, удовлетворяющих условию  $\tau(\mathbf{v}_1, \mathbf{v}_2) = -\tau(\mathbf{v}_2, \mathbf{v}_1)$ . Из любой билинейной функции  $\sigma(\mathbf{v}_1, \mathbf{v}_2)$  мы можем построить кососимметричную функцию

$$\tau(\mathbf{v}_1, \mathbf{v}_2) = \sigma(\mathbf{v}_1, \mathbf{v}_2) - \sigma(\mathbf{v}_2, \mathbf{v}_1).$$

Из элементов пространства  $V^*$  не сложно построить элементы  $\Lambda^2(V^*)$ . Для заданных элементов  $\omega^i, \omega^j \in V^*$  введем функцию  $\omega^i \wedge \omega^j$  согласно определению

$$\omega^i \wedge \omega^j(\mathbf{v}_1, \mathbf{v}_2) = \omega^i(\mathbf{v}_1)\omega^j(\mathbf{v}_2) - \omega^i(\mathbf{v}_2)\omega^j(\mathbf{v}_1).$$

Поскольку функции  $\omega^i$  и  $\omega^j$  линейны, то  $\omega^i \wedge \omega^j$  билинейна. Более того,

$$\omega^i \wedge \omega^j(\mathbf{v}_2, \mathbf{v}_1) = \omega^i(\mathbf{v}_2)\omega^j(\mathbf{v}_1) - \omega^i(\mathbf{v}_1)\omega^j(\mathbf{v}_2) = -\omega^i \wedge \omega^j(\mathbf{v}_1, \mathbf{v}_2),$$

следовательно, функция  $\omega^i \wedge \omega^j$  — кососимметричная.

Из определения внешнего произведения следует, что оно дистрибутивно по сложению:

$$(\omega^1 + \omega^2) \wedge \omega^3 = \omega^1 \wedge \omega^3 + \omega^2 \wedge \omega^3.$$

Чтобы доказать это утверждение, рассмотрим произвольные вектора  $\mathbf{v}_1$  и  $\mathbf{v}_2$ . Тогда

$$\begin{aligned} & [(\omega^1 + \omega^2) \wedge \omega^3](\mathbf{v}_1, \mathbf{v}_2) \\ &= [(\omega^1 + \omega^2)(\mathbf{v}_1)]\omega^3(\mathbf{v}_2) - [(\omega^1 + \omega^2)(\mathbf{v}_2)]\omega^3(\mathbf{v}_1) \\ &= [\omega^1(\mathbf{v}_1) + \omega^2(\mathbf{v}_1)]\omega^3(\mathbf{v}_2) - [\omega^1(\mathbf{v}_2) + \omega^2(\mathbf{v}_2)]\omega^3(\mathbf{v}_1) \\ &= \omega^1 \wedge \omega^3(\mathbf{v}_1, \mathbf{v}_2) + \omega^2 \wedge \omega^3(\mathbf{v}_1, \mathbf{v}_2). \end{aligned}$$

Предположим, что мы выбрали дуальные базисы: базис  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  пространства  $V$  и базис  $\varepsilon^1, \varepsilon^2, \dots, \varepsilon^n$  пространства  $V^*$ ; таким образом,

$$\varepsilon^i(\mathbf{e}_j) = \begin{cases} 1, & \text{если } i = j, \\ 0, & \text{если } i \neq j. \end{cases}$$

Если выразить  $\omega^1$  и  $\omega^2$  через базис  $V^*$ , то

$$\omega^1 \wedge \omega^2 = (a_1 \varepsilon^1 + a_2 \varepsilon^2 + \dots) \wedge (b_1 \varepsilon^1 + b_2 \varepsilon^2 + \dots)$$

будет суммой членов вида  $\varepsilon^i \wedge \varepsilon^j$ . Однако,  $\varepsilon^j \wedge \varepsilon^i = -\varepsilon^i \wedge \varepsilon^j$ . Поэтому каждый элемент вида  $\omega^1 \wedge \omega^2$  может быть выражен через элементы вида  $\varepsilon^i \wedge \varepsilon^j$ , где  $i < j$ . В общем случае, пусть  $\tau$  является элементом пространства  $\Lambda^2(V^*)$ . Вследствие билинейности и кососимметричности  $\tau$  оказывается, что  $\tau$  определяется значениями  $\tau(\mathbf{e}_i, \mathbf{e}_j)$  для  $i < j$ . Введем новое обозначение  $\tau(\mathbf{e}_i, \mathbf{e}_j) = b_{ij}$ . Величина  $\sum b_{ij} \varepsilon^i \wedge \varepsilon^j$  принимает те же значения, что и  $\tau$ , на всех  $\mathbf{e}_k, \mathbf{e}_l$ . Следовательно,  $\tau = \sum b_{ij} \varepsilon^i \wedge \varepsilon^j$ . Значит, элементы  $\varepsilon^i \wedge \varepsilon^j$  порождают пространство  $\Lambda^2(V^*)$ . С другой стороны, легко проверить, что все они линейно независимы. Тогда из  $n$  базисных элементов пространства  $V^*$  мы можем построить  $\frac{1}{2}n(n-1)$  линейно независимых элементов пространства  $\Lambda^2(V^*)$ , потому что элементы  $\varepsilon^i \wedge \varepsilon^j$  при  $i < j$  образуют базис пространства  $\Lambda^2(V^*)$ .

В случае, когда пространством  $V$  является  $\mathbb{R}^3$ , в качестве базисных элементов  $V^*$  удобно взять  $dx, dy$  и  $dz$ . Тогда базис  $\Lambda^2(\mathbb{R}^{3*})$  состоит из трех элементов  $dx \wedge dy, dx \wedge dz$  и  $dy \wedge dz$ .

Взяв два элемента пространства  $V^*$ , мы можем, воспользовавшись внешним произведением и правилом  $\omega^j \wedge \omega^i = -\omega^i \wedge \omega^j$ , умножить и упростить, например,  $(dx + dy) \wedge (dx + 2dy - 3dz) = 2dx \wedge dy - 3dx \wedge dz + dy \wedge dx - 3dy \wedge dz = dx \wedge dy - 3dx \wedge dz - 3dy \wedge dz$ .

Действие элемента пространства  $\Lambda^2(V^*)$  удобно выразить через определитель, т. е.

$$\omega^1 \wedge \omega^2(\mathbf{v}_1, \mathbf{v}_2) = \text{Det} \begin{pmatrix} \omega^1(\mathbf{v}_1) & \omega^1(\mathbf{v}_2) \\ \omega^2(\mathbf{v}_1) & \omega^2(\mathbf{v}_2) \end{pmatrix}.$$

Мы воспользуемся этой формулой для обобщения внешнего произведения на случай, когда множителей больше двух.

Функция  $f$ , зависящая от трех векторов в пространстве  $V$ , называется *трилинейной*, если  $f(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$  зависит линейно от  $\mathbf{v}_i$  ( $i = 1, 2, 3$ ), когда два других вектора фиксированы. Трилинейная функция  $f$  называется *кососимметричной*, если

$$f(\mathbf{v}_2, \mathbf{v}_1, \mathbf{v}_3) = -f(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) \quad \text{и} \quad f(\mathbf{v}_1, \mathbf{v}_3, \mathbf{v}_2) = -f(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3).$$

Из этих соотношений следует, в частности, равенство

$$f(\mathbf{v}_3, \mathbf{v}_2, \mathbf{v}_1) = -f(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3).$$

Например, определитель матрицы  $3 \times 3$  является трилинейной кососимметричной функцией столбцов.

А теперь давайте построим пространство  $\Lambda^3(V^*)$ , т. е. пространство кососимметричных трилинейных функций из  $V$  в  $\mathbb{R}$ . Элементы этого пространства можно построить, воспользовавшись обозначениями внешнего произведения и обобщив определение на языке определителей следующим образом:

$$\omega^i \wedge \omega^j \wedge \omega^k(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = \text{Det} \begin{pmatrix} \omega^i(\mathbf{v}_1) & \omega^i(\mathbf{v}_2) & \omega^i(\mathbf{v}_3) \\ \omega^j(\mathbf{v}_1) & \omega^j(\mathbf{v}_2) & \omega^j(\mathbf{v}_3) \\ \omega^k(\mathbf{v}_1) & \omega^k(\mathbf{v}_2) & \omega^k(\mathbf{v}_3) \end{pmatrix}.$$

Функции  $\omega^i \wedge \omega^j \wedge \omega^k$  обладают рядом свойств, являющихся следствием линейности  $\omega^i(\mathbf{v})$  и элементарных свойств определителя.

- (а) Для фиксированных векторов  $\mathbf{v}_2$  и  $\mathbf{v}_3$  функция  $\omega^i \wedge \omega^j \wedge \omega^k$  линейно зависит от  $\mathbf{v}_1$ , потому что элементы первого столбца определителя линейны по  $\mathbf{v}_1$ , а определитель является линейной функцией элементов первого столбца.
- (б) Функция  $\omega^i \wedge \omega^j \wedge \omega^k$  линейна также по  $\mathbf{v}_2$  и  $\mathbf{v}_3$ , т. е. она трилинейна.

- (с) Функция  $\omega^i \wedge \omega^j \wedge \omega^k$  изменяет знак при перестановке двух соседних аргументов (т.е.  $\mathbf{v}_1$  и  $\mathbf{v}_2$  или  $\mathbf{v}_2$  и  $\mathbf{v}_3$ ) меняются местами, потому что определитель меняет знак при перестановке соседних столбцов.
- (d) Функция  $\omega^i \wedge \omega^j \wedge \omega^k$  меняется знак на противоположную при перестановке двух соседних «множителей» в ней (т.е., например,  $\varepsilon^j \wedge \varepsilon^i \wedge \varepsilon^k = -\varepsilon^i \wedge \varepsilon^j \wedge \varepsilon^k$ ), так как определитель меняет знак при перестановке соседних строк.
- (е)  $\omega^i \wedge \omega^i \wedge \omega^k = 0$ , потому что определитель с двумя одинаковыми строками равен нулю.
- (f) Тройное внешнее произведение дистрибутивно относительно сложения по каждому множителю, потому что определитель дистрибутивен по каждой строке.
- (g) Если  $\mathbf{v}_1$ ,  $\mathbf{v}_2$  и  $\mathbf{v}_3$  линейно зависимы, то

$$\omega^i \wedge \omega^j \wedge \omega^k(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = 0,$$

потому что определитель равен нулю, если линейно зависимы его столбцы.

- (h) Если  $\omega^i$ ,  $\omega^j$  и  $\omega^k$  — линейно зависимые элементы пространства  $V^*$ , то  $\omega^i \wedge \omega^j \wedge \omega^k = 0$ , потому что определитель равен нулю, если линейно зависимы его строки.

Как и раньше, несложно доказать, что элементы  $\varepsilon^i \wedge \varepsilon^j \wedge \varepsilon^k$  при  $i < j < k$  образуют базис пространства  $\Lambda^3(V^*)$ , если  $\varepsilon^i$ ,  $\varepsilon^j$ ,  $\varepsilon^k$  являются базисными элементами пространства  $V^*$ . Существует  $n$  возможностей выбрать  $i$ , затем  $n - 1$  возможность выбора  $j$  и  $n - 2$  возможности выбора  $k$ , потому что если любые два индекса равны, то  $\varepsilon^i \wedge \varepsilon^j \wedge \varepsilon^k = 0$ . Но комбинация из трех индексов допускает  $3 \cdot 2 \cdot 1 = 6$  перестановок, соответствующих только одному элементу с  $i < j < k$ . Следовательно,

$$\dim \Lambda^3(V^*) = \frac{1}{6}n(n-1)(n-2) = \frac{n!}{3!(n-3)}.$$

Теперь мы можем дать определение пространства  $\Lambda^k(V^*)$  при произвольном  $k$ . Это подпространство пространства<sup>2</sup> полилинейных функций  $\mathcal{L}(V, V, \dots, V; \mathbb{R})$ , которые кососимметричны в том смысле, что функция изменяет знак при перестановке любых двух соседних элементов. Элемент пространства  $\Lambda^k(V^*)$ , называемый  $k$ -формой, может быть построен из  $k$  элементов пространства  $V^*$  согласно правилу:

$$\begin{aligned} & \omega^{i_1} \wedge \omega^{i_2} \wedge \dots \wedge \omega^{i_k}(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k) \\ &= \text{Det} \begin{pmatrix} \omega^{i_1}(\mathbf{v}_1) & \omega^{i_1}(\mathbf{v}_2) & \dots & \omega^{i_1}(\mathbf{v}_k) \\ \omega^{i_2}(\mathbf{v}_1) & \omega^{i_2}(\mathbf{v}_2) & \dots & \omega^{i_2}(\mathbf{v}_k) \\ \vdots & \vdots & \ddots & \vdots \\ \omega^{i_k}(\mathbf{v}_1) & \omega^{i_k}(\mathbf{v}_2) & \dots & \omega^{i_k}(\mathbf{v}_k) \end{pmatrix}. \end{aligned}$$

Размерность пространства  $\Lambda^k(V^*)$  равна  $n!/(k!(n-k)!)$ ,<sup>3</sup> а его базис состоит из всех элементов  $\varepsilon^{i_1} \wedge \varepsilon^{i_2} \wedge \dots \wedge \varepsilon^{i_k}$  при условии  $i_1 < i_2 < \dots < i_k$ .<sup>4</sup> Заметим, что пространство  $\Lambda^n(V^*)$  одномерно, и его единственный базисный элемент равен  $\varepsilon^1 \wedge \varepsilon^2 \wedge \dots \wedge \varepsilon^n$ , а в случае  $k > n$  пространство  $\Lambda^k(V^*) = \{0\}$ .

Для нас будет важна еще одна алгебраическая операция — «внешнее» умножение элемента пространства  $\Lambda^p(V^*)$  на элемент пространства  $\Lambda^q(V^*)$ , в результате чего получается элемент  $\Lambda^{p+q}(V^*)$ . Если пользоваться базисными элементами, то внешнее умножение определяется формулой

$$(\varepsilon^{i_1} \wedge \varepsilon^{i_2} \wedge \dots \wedge \varepsilon^{i_p}) \wedge (\varepsilon^{j_1} \wedge \varepsilon^{j_2} \wedge \dots \wedge \varepsilon^{j_q}) = \varepsilon^{i_1} \wedge \dots \wedge \varepsilon^{i_p} \wedge \varepsilon^{j_1} \wedge \dots \wedge \varepsilon^{j_q}.$$

Это определение может быть продолжено на линейные комбинации базисных элементов так, чтобы сохранялась линейность по

<sup>2</sup>обычно называемого пространством тензоров типа  $(k, 0)$ , или  $k$  раз ковариантных. Таким образом,  $\Lambda^k(V^*)$  есть пространство кососимметрических тензоров типа  $(k, 0)$ . — Прим. ред.

<sup>3</sup>Это выражение представляет собой биномиальный коэффициент  $C_n^k$  (другой вариант обозначения —  $\binom{n}{k}$ ), а именно, коэффициент при  $x^k$  в разложении многочлена  $(1+x)^n$ , равный числу способов выбора неупорядоченного набора  $k$  элементов из  $n$ -элементного множества. — Прим. ред.

<sup>4</sup>Как и прежде, имеется в виду, что  $\varepsilon^1, \varepsilon^2, \dots, \varepsilon^n$  — некоторый базис пространства  $V^*$ . Доказательство линейной независимости указанных элементов и порождения ими всего  $\Lambda^k(V^*)$  оставляем читателю в качестве упражнения. — Прим. ред.

каждому множителю отдельно, т. е. должен выполняться дистрибутивный закон умножения. Например, если

$$\omega = 5\varepsilon^1 \wedge \varepsilon^7 + 3\varepsilon^2 \wedge \varepsilon^4 \quad \text{и} \quad \sigma = 6\varepsilon^1 \wedge \varepsilon^5 + 2\varepsilon^3 \wedge \varepsilon^5 - 9\varepsilon^3 \wedge \varepsilon^6$$

являются элементами пространства  $\Lambda^2(V^*)$ , то элемент  $\omega \wedge \sigma$  пространства  $\Lambda^4(V^*)$  вычисляется следующим образом:

$$\begin{aligned} \omega \wedge \sigma &= 5(\varepsilon^1 \wedge \varepsilon^7) \wedge \sigma + 3(\varepsilon^2 \wedge \varepsilon^4) \wedge \sigma \\ &= 30\varepsilon^1 \wedge \varepsilon^7 \wedge \varepsilon^1 \wedge \varepsilon^5 + 10\varepsilon^1 \wedge \varepsilon^7 \wedge \varepsilon^3 \wedge \varepsilon^5 - 45\varepsilon^1 \wedge \varepsilon^7 \wedge \varepsilon^3 \wedge \varepsilon^6 \\ &\quad + 18\varepsilon^2 \wedge \varepsilon^4 \wedge \varepsilon^1 \wedge \varepsilon^5 + 6\varepsilon^2 \wedge \varepsilon^4 \wedge \varepsilon^3 \wedge \varepsilon^5 - 27\varepsilon^2 \wedge \varepsilon^4 \wedge \varepsilon^3 \wedge \varepsilon^6. \end{aligned}$$

Первое слагаемое в правой части равно нулю, потому что в него элемент  $\varepsilon^1$  входит два раза. Остальные слагаемые надо преобразовать, чтобы превратить их в базисные элементы; при этом возможны изменения знаков. Например,

$$\varepsilon^1 \wedge \varepsilon^7 \wedge \varepsilon^3 \wedge \varepsilon^5 = \varepsilon^1 \wedge \varepsilon^3 \wedge \varepsilon^5 \wedge \varepsilon^7,$$

потому что  $\varepsilon^7$  надо переставлять на два места назад, а

$$\varepsilon^2 \wedge \varepsilon^4 \wedge \varepsilon^3 \wedge \varepsilon^6 = -\varepsilon^2 \wedge \varepsilon^3 \wedge \varepsilon^4 \wedge \varepsilon^6.$$

Итак,

$$\begin{aligned} \omega \wedge \sigma &= 10\varepsilon^1 \wedge \varepsilon^3 \wedge \varepsilon^5 \wedge \varepsilon^7 - 45\varepsilon^1 \wedge \varepsilon^3 \wedge \varepsilon^6 \wedge \varepsilon^7 \\ &\quad + 18\varepsilon^1 \wedge \varepsilon^2 \wedge \varepsilon^4 \wedge \varepsilon^5 - 6\varepsilon^2 \wedge \varepsilon^3 \wedge \varepsilon^4 \wedge \varepsilon^5 + 27\varepsilon^2 \wedge \varepsilon^3 \wedge \varepsilon^4 \wedge \varepsilon^6. \end{aligned}$$

Мы настоятельно рекомендуем читателю сделать побольше примеров внешнего умножения, чтобы почувствовать, как устроена эта операция. До конца книги внешнее умножение будет одним из основных вычислительных методов. Поэтому усилия, вложенные в приобретение подобных вычислительных навыков, окупятся сторицей в дальнейшем. Читатель должен убедиться в справедливости следующих правил:

Ассоциативный закон:  $(\omega \wedge \sigma) \wedge \tau = \omega \wedge (\sigma \wedge \tau).$

Дистрибутивный закон:  $\omega \wedge (\sigma + \tau) = \omega \wedge \sigma + \omega \wedge \tau.$

Антикоммутативность (сегодня это часто называют суперкоммутативностью) значит следующее<sup>5</sup>:

$$\omega \wedge \tau = (-1)^{pq} \tau \wedge \omega, \quad \text{если } \omega \in \Lambda^p(V^*) \text{ и } \tau \in \Lambda^q(V^*).$$

Эти правила ускоряют процедуру вычисления внешнего произведения.

С логической точки зрения наше определение не совсем удовлетворительно, потому что кажется, что есть зависимость от выбора базиса. Строго говоря, мы должны доказать, что если выбрать другой базис в пространстве  $V$ , соответственно получив другой базис пространства  $\Lambda^p(V^*)$ , то результат умножения элементов не изменится. Это можно показать непосредственным вычислением, что довольно утомительно. Гораздо лучше дать более абстрактное определение умножения, которое будет автоматически удовлетворять всем правилам, так что формула для умножения станет следствием этого определения. Мы сделаем это в приложении к главе 18, но если читателю это очень интересно, пусть прочитает это приложение прямо сейчас — оно независимо от остального материала. Гораздо важнее приобрести навык вычисления внешнего произведения.

В качестве примера рассмотрим четырехмерное пространство  $V$ . Имея в виду очевидную интерпретацию как пространства-времени, обозначим базисные элементы пространства  $V^*$  как  $dt$ ,  $dx$ ,  $dy$  и  $dz$ . В этом случае получим полный набор пространств  $\Lambda^k(V^*)$ :

- пространство  $\Lambda^0(V^*)$  одномерно; единственный базисный элемент назовем 1;
- пространство  $\Lambda^1(V^*)$  четырехмерно; это пространство  $V^*$ , его базисные элементы:  $\{dt, dx, dy, dz\}$ ;
- пространство  $\Lambda^2(V^*)$  шестимерно; базисные элементы:  $dt \wedge dx$ ,  $dt \wedge dy$ ,  $dt \wedge dz$ ,  $dx \wedge dy$ ,  $dx \wedge dz$ ,  $dy \wedge dz$ ;
- пространство  $\Lambda^3(V^*)$  четырехмерно; базисные элементы:  $dt \wedge dx \wedge dy$ ,  $dt \wedge dx \wedge dz$ ,  $dt \wedge dy \wedge dz$ ,  $dx \wedge dy \wedge dz$ ;

<sup>5</sup> Аккуратное доказательство этого свойства также является хорошим упражнением. — *Прим. ред.*

- пространство  $\Lambda^4(V^*)$  одномерно; базисный элемент  $dt \wedge dx \wedge dy \wedge dz$ .

## 15.2. $k$ -формы и оператор $d$

Чтобы перейти от пространства  $\Lambda^k(V^*)$  к пространству  $\Omega^k(V)$ , мы сделаем такое же расширение, как мы делали, переходя от дуального пространства  $V^*$  (обозначаемого также  $\Lambda^1(V^*)$ ) к пространству  $\Omega^1(V)$ . Элементом пространства  $\Omega^k(V)$  является функция, которая каждой точке приписывает элемент пространства  $\Lambda^k(V^*)$ , т. е. коэффициентами базисных элементов  $\Lambda^k(V^*)$  становятся функции, принимающие вещественные значения. Например, общий вид элемента пространства  $\Omega^2(\mathbb{R}^3)$  — это

$$a(x, y, z)dx \wedge dy + b(x, y, z)dx \wedge dz + c(x, y, z)dy \wedge dz,$$

где  $a, b, c$  — вещественнозначные функции. Что касается алгебраических операций, в том числе сложения и внешнего умножения, то элементы пространства  $\Omega^k(V)$  ведут себя так же, как элементы пространства  $\Lambda^k(V^*)$ . Элементы пространства  $\Omega^k(V)$ , называемые дифференциальными  $k$ -формами или просто  $k$ -формами, будут выступать как зародыши  $k$ -коцепей. Мы должны научиться дифференцировать и интегрировать их. Рассмотрим пространство  $V = \mathbb{R}^n$  и пусть  $dx^1, \dots, dx^n$  образуют базис пространства  $V^*$ . (Мы используем также обозначения  $dx, dy$  и т. д.)

Мы уже знакомы с дифференциальным оператором  $d$ , приписывающим 1-форму 0-форме. Если  $f(x, y, z)$  — дифференцируемая функция (0-форма), то

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz.$$

В конечном счете мы хотим построить  $k$ -форму как объект, который мы будем интегрировать по « $k$ -мерной гиперповерхности». Кроме того, мы хотим доказать теорему Стокса, которая говорит, что

$$\int_S dw = \int_{\partial S} \omega,$$

где  $S$  — это « $k$ -мерная гиперповерхность» и  $\partial S$  — ее граница. В главе 8 мы видели, что это уравнение, примененное к инфинитезимальным параллелограммам, давало определение оператора  $d$ , действующего на 1-формы. Тогда мы получили, что

$$d(f dg) = df \wedge dg.$$

Поскольку любая 1-форма является суммой таких выражений, то это давало нам определение  $d$  для 1-форм. Может быть, сейчас полезно вернуться назад и еще раз посмотреть на основные формулы главы 8, выписанные на с. 425.

Давайте в качестве примера проделаем алгебраические выкладки для 1-формы. Мы хотели бы получить

$$d(f dg \wedge dh \wedge \cdots) = df \wedge dg \wedge dh \wedge \cdots. \quad (*)$$

Это заставляет нас определить действие оператора  $d$  на  $k$ -форму в следующем виде:

$$\begin{aligned} \text{если } \tau &= f dx^{i_1} \wedge \cdots \wedge dx^{i_k} + g dx^{j_1} \wedge \cdots \wedge dx^{j_k} + \cdots, \\ \text{то } d\tau &= df \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} + dg \wedge dx^{j_1} \wedge \cdots \wedge dx^{j_k}. \end{aligned} \quad (15.1)$$

После действия оператора  $d$  бывает необходимо перегруппировать члены, чтобы собрать коэффициенты при каждом базисном элементе.

Рассмотрим пример действия  $d$  на 2-форму в пространстве  $\mathbb{R}^3$ . Пусть  $\tau = B_x dy \wedge dz - B_y dx \wedge dz + B_z dx \wedge dy$ , где  $B_x, B_y, B_z$  — дифференцируемые функции  $x, y, z$ , которые можно рассматривать как компоненты векторного поля. Тогда  $d\tau = dB_x \wedge dy \wedge dz - dB_y \wedge dx \wedge dz + dB_z \wedge dx \wedge dy$ . Подставляя дифференциалы и используя тот факт, что внешнее произведение равно нулю, если совпадают любые два множителя, мы получаем

$$d\tau = \left( \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} \right) dx \wedge dy \wedge dz.$$

Те, кто знаком с векторным анализом, могут заметить сходство между  $d : \Omega^2(\mathbb{R}^3) \rightarrow \Omega^3(\mathbb{R}^3)$  и операцией  $\operatorname{div}$ .

Оператор  $d$  обладает рядом полезных свойств, которые следует отметить. Каждое свойство достаточно проверить для случая

$k$ -формы вида  $\Omega = f dx \wedge dy \wedge \dots$ . Общий случай становится понятным в силу линейности оператора  $d$ .

Первое свойство связано с результатом действия оператора  $d$  на произведение функции и  $k$ -формы. Пусть  $f$  и  $g$  — дифференцируемые функции и  $\Omega = g dx \wedge dy \wedge \dots$ . Тогда  $d(f\Omega) = d(fg dx \wedge dy \wedge \dots) = d(fg) \wedge dx \wedge dy \wedge \dots$ . Но для дифференциалов существует правило  $d(fg) = f dg + g df$ . Отсюда следует, что  $d(f\Omega) = f dg \wedge dx \wedge dy \wedge \dots + g df \wedge dx \wedge dy \wedge \dots$  или  $d(f\Omega) = f d(g dx \wedge dy \wedge \dots) + df \wedge (g dx \wedge dy \wedge \dots)$ . Тогда мы получаем

$$d(f\Omega) = f d\Omega + df \wedge \Omega.$$

Второе свойство относится к действию оператора на произведение  $p$ -формы  $\omega$  и  $q$ -формы  $\Omega$ . Пусть  $\omega = f dx^1 \wedge \dots \wedge dx^p$  и  $\Omega = g dy^1 \wedge \dots \wedge dy^q$ , где  $f$  и  $g$  — дифференцируемые функции. Имеем  $d(\omega \wedge \Omega) = d(fg dx^1 \wedge \dots \wedge dx^p \wedge dy^1 \wedge \dots \wedge dy^q)$ . Поэтому  $d(\omega \wedge \Omega) = f(dg \wedge dx^1 \wedge \dots \wedge dx^p \wedge dy^1 \wedge \dots \wedge dy^q) + g df \wedge dx^1 \wedge \dots \wedge dx^p \wedge dy^1 \wedge \dots \wedge dy^q$ . В первом члене переставляем  $dg$  и множитель  $dx^1 \wedge \dots \wedge dx^p$ , в результате чего получается знак  $(-1)^p$ . Тогда

$$\begin{aligned} d(\omega \wedge \Omega) &= (-1)^p (f dx^1 \wedge \dots \wedge dx^p) \wedge (dg \wedge dy^1 \wedge \dots \wedge dy^q) \\ &\quad + (df \wedge dx^1 \wedge \dots \wedge dx^p) \wedge (g dy^1 \wedge \dots \wedge dy^q). \end{aligned}$$

Поскольку каждая  $p$ -форма есть сумма членов вида  $f dx^1 \wedge \dots \wedge dx^p$  и каждая  $q$ -форма есть сумма членов вида  $g dy^1 \wedge \dots \wedge dy^q$ , мы имеем

$$d(\omega \wedge \Omega) = (-1)^p \omega \wedge d\Omega + d\omega \wedge \Omega. \quad (15.2)$$

Третье свойство касается двукратного действия оператора  $d$ . Сначала рассмотрим  $d(df)$ , где  $f(x, y, \dots)$  — дважды дифференцируемая функция. В этом случае

$$d(df) = d\left(\frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \dots\right) = \frac{\partial^2 f}{\partial y \partial x} dy \wedge dx + \frac{\partial^2 f}{\partial x \partial y} dx \wedge dy + \dots.$$

Из равенства смешанных частных производных и соотношения  $dy \wedge dx = -dx \wedge dy$  следует, что  $d(df) = 0$ . Теперь рассмотрим более общий случай  $k$ -формы  $\omega = f dx \wedge dy \wedge \dots$ . В этом случае  $d\omega = df \wedge (dx \wedge dy \wedge \dots)$  и  $d(d\omega) = d(df \wedge (dx \wedge dy \wedge \dots)) + d(df) \wedge dx \wedge dy \wedge \dots$ .

Однако  $d(dx \wedge dy \wedge \dots) = 0$  и  $d(df) = 0$ . Отсюда следует, что в общем случае

$$d(d\omega) = 0. \quad (15.3)$$

Это согласуется со свойством кограничного оператора:

$$\boxed{d \circ d = 0.}$$

Конечно, уравнение (\*) теперь следует из (15.2) и (15.3).

Если мы считаем, что  $dx$  — это результат действия оператора  $d$  на координатную функцию  $x$ , то определение (15.1) становится следствием (15.2) и (15.3). Поэтому, когда Вы убедились в существовании оператора  $d$ , остается только запомнить следующие его свойства:

- (1)  $df = (\partial f / \partial x)dx + (\partial f / \partial y)dy + \dots$ , где  $x, y, \dots$  координатные функции;
- (2)  $d$  аддитивен, т.е.  $d(\omega + \sigma) = d\omega + d\sigma$ ;
- (3) как оператор  $d$  действует на произведение, т.е. (15.2);
- (4)  $d \circ d = 0$ .

С помощью этих четырех правил можно вычислить действие  $d$  на любую форму.

### 15.3. Интегрирование $k$ -форм

Чтобы закончить идентификацию  $k$ -форм в качестве *зародышей  $k$ -коцепей*, мы должны объяснить, как  $k$ -форма  $\Omega$  приписывает  $k$ -коцепь каждому комплексу с дифференцируемыми ячейками, находящемуся в  $n$ -мерном пространстве. Сначала рассмотрим случай  $k = 1$ . Пусть в пространстве  $\mathbb{R}^n$  имеется гладкая кривая, соединяющая точки  $A$  и  $B$ , и мы хотим вычислить интеграл от 1-формы  $\omega$  на этой кривой. Для этого выберем произвольную гладкую параметризацию кривой, т.е. произвольное гладкое отображение  $\alpha : \mathbb{R} \rightarrow \mathbb{R}^n$ , которое отображает отрезок  $[0, 1]$  в желаемую кривую с условиями  $\alpha(0) = A$  и  $\alpha(1) = B$ , как это показано на

рис. 15.3. Затем произведем *перенос* (*pullback*) 1-формы  $\omega$  и получим  $\alpha^*\omega$ , выражение типа  $g(t) dt$ . Определим интеграл по траектории от формы  $\omega$  вдоль кривой  $\alpha$  формулой  $\int_\alpha \omega = \int_0^1 (\alpha^*\omega)$ . Это определение полезно тем, что результат не зависит от способа параметризации кривой  $\alpha$ . В этом случае интеграл по кривой сводится к интегрированию по отрезку  $[0, 1]$ . Ключевым пунктом построения является процедура переноса для 1-формы: если  $\omega = f dx + g dy + \dots$ , то  $\alpha^*\omega = (\alpha^*f) d(\alpha^*x) + (\alpha^*g) d(\alpha^*y) + \dots$ . Очень важно доказать (с помощью цепного правила), что результат не зависит от параметризации.

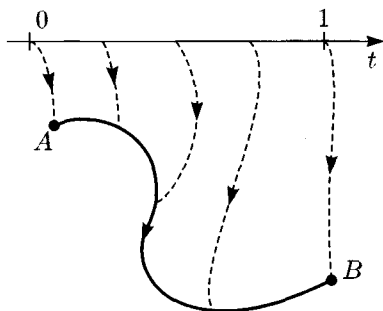


Рис. 15.3

В этом определении и в соответствующей теории двумерного интегрирования, которую мы обсуждали в главе 8, перенос играет чрезвычайно важную роль. Поэтому в первую очередь мы должны обобщить эту процедуру. Пусть  $\phi : \mathbb{R}^k \rightarrow \mathbb{R}^n$  — дифференцируемое отображение. Мы хотим определить операцию  $\phi^*$ , называемую переносом, которая любой дифференциальной форме  $\omega$  (любой степени) в пространстве  $\mathbb{R}^n$  приписывает дифференциальную форму  $\phi^*\omega$  (той же степени) в пространстве  $\mathbb{R}^k$ . При этом нужно, чтобы  $\phi^*$  сохраняла операции сложения и умножения, т. е. выполнялись равенства

$$\phi^*(\omega_1 + \omega_2) = \phi^*\omega_1 + \phi^*\omega_2 \quad (15.4)$$

и

$$\phi^*(\omega_1 \wedge \omega_2) = \phi^*\omega_1 \wedge \phi^*\omega_2. \quad (15.5)$$

Кроме того, мы хотим, чтобы  $\phi^*$  оставалось нашей прежней операцией переноса, когда она действует на 0-формы (т.е. на функции) и на линейные дифференциальные формы. Это значит, что если  $f \in \Omega^0(\mathbb{R}^n)$  — функция, то должны выполняться равенства

$$\phi^* f = f \circ \phi, \quad (15.6)$$

и

$$\phi^* df = d(\phi^* f). \quad (15.7)$$

Условия (15.4)–(15.7) полностью детерминируют построение. Например, рассмотрим отображение  $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ , определяемое следующим образом:

$$\phi \begin{pmatrix} r \\ \psi \\ \theta \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \text{где} \quad \begin{aligned} x &= r \sin \theta \cos \psi, \\ y &= r \sin \theta \sin \psi, \\ z &= r \cos \theta. \end{aligned}$$

Тогда, как следует из (15.6),

$$\phi^* x = r \sin \theta \cos \psi,$$

и так далее. Предположим, что нам надо вычислить  $\phi(dx \wedge dy \wedge dz)$ . Мы имеем

$$\begin{aligned} \phi^*(dx \wedge dy \wedge dz) &= \phi^* dx \wedge \phi^* dy \wedge \phi^* dz && \text{из (15.5)} \\ &= (d\phi^* x) \wedge (d\phi^* y) \wedge (d\phi^* z) && \text{из (15.7)} \\ &= [d(r \sin \theta \cos \psi)] \wedge [d(r \sin \theta \sin \psi)] \wedge [d(r \cos \theta)] && \text{из (15.6)}. \end{aligned}$$

Внутри каждого множителя этого тройного произведения мы пользуемся обычными правилами вычисления дифференциала от функции, так что  $d(r \sin \theta \cos \psi) = \sin \theta \cos \psi dr - r \sin \theta \sin \psi d\psi + r \cos \theta \cos \psi d\theta$  и т.д. Затем мы пользуемся правилами внешнего умножения, чтобы выразить тройное произведение через  $dr \wedge d\theta \wedge d\psi$ . В результате получаем

$$\phi^*(dx \wedge dy \wedge dz) = r^2 \sin \theta dr \wedge d\theta \wedge d\psi.$$

(Обычно эту формулу называют евклидовым объемом в сферических координатах<sup>6</sup>.)

В общем случае следствием условий (15.4)–(15.7) является определение

$$\phi^*(f dx^{i_1} \wedge dx^{i_2} \wedge \dots \wedge dx^{i_k}) = (f \circ \phi) d(x^{i_1} \circ \phi) \wedge \dots \wedge d(x^{i_k} \circ \phi), \quad (15.8)$$

где  $x^i$  ( $i = 1, 2, \dots, n$ ) — координаты в пространстве  $\mathbb{R}^n$ . В силу условия (15.4) наиболее общий вид  $\phi^*$  является суммой таких выражений. Если взять (15.8) и его обобщение на суммы в качестве определения процедуры переноса, то мы должны вернуться назад и проверить, что условия (15.4)–(15.7) выполняются. Это простое вычисление, и мы оставляем его читателю как упражнение. Из определения оператора  $d$  и свойств процедуры переноса следует, что

$$\begin{aligned} \phi^*[d(f dx^i \wedge dx^j \wedge \dots)] &= \phi^*[df \wedge dx^i \wedge \dots] \\ &= d(\phi^* f) \wedge d\phi^* x^i \wedge \dots \\ &= d[(\phi^* f) d\phi^* x^i \wedge \dots] \\ &= d[\phi^*(f dx^i \wedge \dots)]. \end{aligned}$$

Проделав такие вычисления с суммами выражений такого вида, мы получаем основную формулу

$$\phi^* d\omega = d\phi^* \omega. \quad (15.9)$$

Этот результат чрезвычайно важен. Он показывает, что если сначала подействовать оператором  $d$  и потом применить перенос, то получится то же самое, что и если сначала сделать перенос, а потом подействовать оператором  $d$ .

Есть еще одна важная формула для процедуры переноса, которую следует отметить. Пусть  $\psi : \mathbb{R}^p \rightarrow \mathbb{R}^k$ , так что мы можем рассмотреть композицию  $\phi \circ \psi : \mathbb{R}^p \rightarrow \mathbb{R}^n$ . Тогда мы знаем, что для любой функции  $f$  справедливо равенство

$$(\phi \circ \psi)^* f = f \circ (\phi \circ \psi) = (f \circ \phi) \circ \psi = \psi^* \phi^* f.$$

---

<sup>6</sup>он же — определитель матрицы Якоби отображения координатного пространства  $r, \theta, \psi$  в координатное пространство  $x, y, z$ ; смысл определителя в двумерной ситуации обсуждался в разделе 8.5 (с. 407). — *Прим. ред.*

Но тогда из полученных правил следует, что

$$(\psi \circ \phi)^* \omega = \phi^*(\psi^* \omega) \quad (15.10)$$

справедливо для любой дифференциальной формы  $\omega$ .

Давайте соберем вместе различные правила внешнего дифференциального исчисления.

*Алгебраические операции* — сложение и умножение дифференциальных форм:

- При сложении и умножении на скаляр дифференциальные формы (любой заданной степени) образуют векторное пространство.
- Произведение формы степени  $p$  на форму степени  $q$  образует форму степени  $p + q$ .
- Дистрибутивный закон:  $(\omega_1 + \omega_2) \wedge \sigma = \omega_1 \wedge \sigma + \omega_2 \wedge \sigma$ .
- Ассоциативный закон:  $(\omega_1 \wedge \omega_2) \wedge \omega_3 = \omega_1 \wedge (\omega_2 \wedge \omega_3)$ .
- Антисимметричность:  $\omega_1 \wedge \omega_2 = (-1)^{pq} \omega_2 \wedge \omega_1$ , если степень  $\omega_1$  равна  $p$  и степень  $\omega_2$  равна  $q$ .

*Оператор  $d$ :*

- Если  $\omega$  является формой степени  $p$ , то  $d\omega$  будет формой степени  $p + 1$ , где  $df = (\partial f / \partial x)dx + (\partial f / \partial y)dy + \dots$  в координатах  $(x, y, \dots)$ .

Кроме того,

$$\begin{aligned} d(\omega_1 + \omega_2) &= d\omega_1 + d\omega_2, \\ d(\omega \wedge \sigma) &= d\omega \wedge \sigma + (-1)^p \omega \wedge d\sigma, \quad \text{если степень } \omega \text{ равна } p, \end{aligned}$$

и

$$d(d\omega) = 0.$$

Перенос:

$$\begin{aligned}\phi^* f &= f \circ \phi \quad \text{для функции } f, \\ \phi^*(\omega_1 + \omega_2) &= \phi^*\omega_1 + \phi^*\omega_2, \\ \phi^*(\omega \wedge \sigma) &= \phi^*\omega \wedge \phi^*\sigma, \\ \phi^* d &= d\phi^*\end{aligned}$$

и

$$(\psi \circ \phi)^* = \phi^* \circ \psi^*.$$

Во всех рассуждениях предполагалось, что отображения  $\phi$ ,  $\psi$  и т. д. и формы  $\omega$ ,  $\sigma$  и т. д. определены во всем пространстве. Конечно, это не обязательно. Мы можем определить различные операции над формами, отображениями и так далее, когда они определены на различных открытых множествах. Надо только быть уверенным, что области определения соответствуют друг другу. Так, если  $\phi$  определено на некотором открытом множестве  $U$ , то достаточно, чтобы  $\omega$  было определено на открытом множестве  $O$ , которое содержит  $\phi(U)$ . Тогда мы сможем определить перенос  $\phi^*\omega$ . Если  $\psi$  определено только на открытом множестве  $O$ , то мы должны предположить, что  $\psi(O) \subset U$ , если хотим определить  $\phi \circ \psi$ .

**Определение интегрирования по областям в пространстве  $\mathbb{R}^k$ .** Сейчас мы хотим определить интеграл от  $k$ -формы по ориентированной области  $U$  в пространстве  $\mathbb{R}^k$ , т. е. мы рассматриваем ситуацию, когда степень формы равна размерности пространства. Предположим, что  $U$  — ограниченная область, и она «хорошая» в следующем смысле. Пусть на пространство  $\mathbb{R}^k$  наложена сетка с ячейками  $\varepsilon$ , так что  $\mathbb{R}^k$  оказывается составленным из неперекрывающихся кубиков с ребром  $\varepsilon$ . Пусть  $\text{Inn}_\varepsilon(U)$  обозначает полный объем всех тех кубиков, что целиком лежат в области  $U$ , а  $\text{Out}_\varepsilon(U)$  — полный объем всех кубиков, которые хотя бы частично лежат в данной области (см. рис. 15.4). Тогда можно составить разность

$$\text{Out}_\varepsilon(U) - \text{Inn}_\varepsilon(U)$$

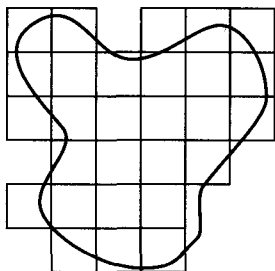


Рис. 15.4

и сделать ее достаточно малой, выбрав соответствующее малое значение  $\varepsilon$ . Это значит, что, вычисляя аппроксимирующие римановы суммы, мы не должны беспокоиться о кубиках вдоль границы. Например, очевидно, что любой ограниченный многогранник в этом смысле «хороший». Кроме того, если  $U$  — хорошая область и  $\psi$  — дифференцируемое отображение, определенное в области немного большей, чем  $U$ , то из теоремы о среднем значении следует, что  $\psi(U)$  тоже хорошее. (Вы можете доказать это самостоятельно или прочитайте главу 8 книги Loomis and Sternberg, *Advanced Calculus*, где подробно объясняются основные факты об интегралах Римана в  $n$ -мерном пространстве.) Предположим также, что ориентация  $U$  стандартная (задаваемая стандартным базисом и соответствующими координатами  $x^1, \dots, x^k$ ).

Пусть  $\Omega$  обозначает  $k$ -форму в пространстве  $\mathbb{R}^k$ . Тогда можно написать

$$\Omega = f dx^1 \wedge \dots \wedge dx^k,$$

где  $f$  — дифференцируемая функция. Для любой ячейки  $\varepsilon$  можно составить аппроксимирующую риманову сумму

$$\sum f(\mathbf{p}_i) \text{Vol}(\square_i) = \varepsilon^k \sum f(\mathbf{p}_i),$$

где  $\mathbf{p}_i \in \square_i$  и  $\square_i$  пробегает по всем кубикам сетки в области  $U$  (или хотя бы частично попадающим в нее). Непрерывность (равномерная)  $f$  гарантирует, что эта сумма стремится к пределу при  $\varepsilon \rightarrow 0$ , не зависящему от выбора как  $\mathbf{p}_i$  в кубиках, так и от самой сетки. Этот предел будем обозначать

$$\int_U \Omega.$$

Важно помнить, что в этом определении область  $U$  имеет стандартную ориентацию пространства  $\mathbb{R}^k$ . Приведем очевидные свойства этого интеграла. Предположим, что  $U = U_1 \cup \dots \cup U_p$  является конечным объединением хороших подобластей, имеющих стандартную ориентацию. Тогда

$$\int_U \Omega = \int_{U_1} \Omega + \int_{U_2} \Omega + \dots + \int_{U_p} \Omega.$$

**Игнорирование множеств меньшей размерности.** Будем говорить, что множество  $S$  имеет нулевую меру<sup>7</sup>, если для любого положительного числа  $\delta$  мы имеем  $\text{Out}_{\varepsilon}(S) < \delta$  при достаточно малом выбранном значении  $\varepsilon$ . (Другими словами,  $S$  имеет нулевую меру, если полный объем кубиков в сетке, которая пересекает  $S$ , может быть сделан сколь угодно малым, выбирая размер ячеек сетки достаточно малым.) Например, если  $S$  лежит в линейном подпространстве размерности  $< k$ , то очевидно, что  $S$  имеет нулевую меру. Множества нулевой меры не дают никакого вклада в интеграл, т.е. если  $U = U' \cup S$ , где  $S$  имеет нулевую меру, то  $\int_U \Omega = \int_{U'} \Omega$ . Отсюда, например, следует, что если мы разделим многогранник на подмногогранники (рис. 15.5), то можно при вычислении интегралов не беспокоиться о гранях меньшей размерности. Кроме того,

если  $U$  — единичный куб  $\{0 \leq x_i \leq 1, i = 1, \dots, k\}$ , то

$$\int_U \Omega = k\text{-кратному интегралу от } f = \int_0^1 \dots \int_0^1 f dx_1 \dots dx_k.$$

Доказательство точно такое же, как и для случая двух измерений (см. с. 389 в главе 8). Конечно, так же как и в случае двух измерений, вместо куба можно взять любую область, приводящую к  $k$ -кратному интегралу.

<sup>7</sup>Подразумевается определение меры в смысле Жордана, которая отличается от «стандартной» меры Лебега более узким классом измеримых множеств. Ограниченное множество в евклидовом пространстве измеримо по Жордану (или квадратуемо), если существует интеграл Римана (по объемлющей области) от его характеристической функции, при этом мера множества равна значению интеграла. — *Прим. ред.*

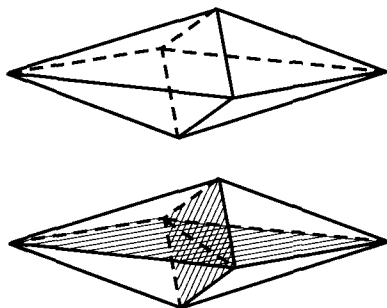


Рис. 15.5

**Формула замены переменных.** Предположим, что  $\phi : \mathbb{R}^k \rightarrow \mathbb{R}^k$  — взаимно однозначное, дифференцируемое, *сохраняющее ориентацию* отображение, для которого существует дифференцируемое обратное. Тогда для любой  $k$ -формы  $\omega$ , определенной в пространстве образов, и для любой хорошей области  $U$  в пространстве прообразов мы имеем

$$\int_{\phi(U)} \omega = \int_U \phi^* \omega.$$

(Как всегда,  $\phi$  должна быть определена в области, немного большей, чем  $U$ , а  $\omega$  должна быть определена в области немного большей, чем  $\phi(U)$ .)

*Доказательство формулы замены переменных.*

(i) Пусть  $A : \mathbb{R}^k \rightarrow \mathbb{R}^k$  — линейное отображение. Тогда для любой связной хорошей области  $\text{Vol}(AU)/\text{Vol}(U) = |\text{Det } A|$ .

Начнем с доказательства того, что левая часть этого равенства не зависит от  $U$ . Это делается так же, как и для двумерного пространства (см. параграф 1.9). Давайте обозначим левую часть равенства  $\text{Vol}(A)$ . Тем самым  $\text{Vol}(A)$  показывает, во сколько раз отображение  $A$  изменяет объем. Из определения следует, что

$$\text{Vol}(AB) = \text{Vol}(A) \cdot \text{Vol}(B).$$

Из главы 11 мы знаем, что

$$\text{Det}(AB) = \text{Det } A \cdot \text{Det } B.$$

Мы хотим доказать, что

$$\text{Vol}(A) = |\text{Det } A| \quad (*)$$

для всех  $k \times k$  матриц. Из предыдущих двух равенств можно сделать заключение, что если  $(*)$  справедливо для  $A$  и  $B$ , то оно верно и для  $AB$ . Если  $A$  — диагональная матрица, то легко увидеть, что  $(*)$  выполняется. Если у матрицы  $A$  все элементы под диагональю равны нулю, на диагонали стоят 1 и над диагональю только один элемент не равен нулю, то  $\text{Vol}(A) = 1$  (это по сути двумерное утверждение относительно преобразования сдвига). Из этого следует, что если  $A = U$  — верхне-треугольная матрицы с 1 на диагонали (под диагональю стоят только нули), то  $\text{Vol}(U) = \text{Det } U = 1$ . Аналогично для нижне-треугольных матриц. Итак, можно сделать вывод, что  $(*)$  верно для всех матриц  $A$ , что можно записать как

$$A = LDU,$$

где  $L$  — нижне-треугольная,  $D$  — диагональная и  $U$  — верхне-треугольная. Когда можно разложить матрицу  $A$  таким образом? Внимательное изучение процедуры редукции по строкам в параграфе 10.8 показывает, что такое разложение возможно тогда и только тогда, когда в редукции по строкам матрицы  $A$  не требуется делать транспонирование. А это происходит тогда и только тогда, когда ни один из главных миноров не обращается в нуль. (главные миноры — это определители квадратных матриц, возникающие на пересечении первых  $r$  строк и столбцов матрицы  $A$ .) Итак, должны выполняться условия:

$$a_{11} \neq 0, \quad (a_{11}a_{22} - a_{21}a_{12}) \neq 0, \quad \dots, \quad \text{Det } A \neq 0.$$

Однако, для любой матрицы  $A$  все эти неравенства можно обеспечить, делая сколь угодно малые изменения матричных элементов. Поскольку  $\text{Vol}(A)$  и  $\text{Det } A$  — непрерывные функции  $A$ , то равенство  $(*)$  выполняется для всех матриц.

Теперь приведем другое доказательство формулы  $(*)$ , основанное на другом способе разложения несингулярных матриц в

произведения. Для любой матрицы  $A$  матрица  $A^*A$  — самосопряженная, поскольку

$$(A^*A)^* = A^*A^{**} = A^*A.$$

Более того, если  $A$  несингулярна, то все собственные значения матрицы  $A^*A$  положительны, так как  $(A^*Ax, x) = (Ax, Ax) > 0$  для любого ненулевого вектора  $x$ . Поэтому  $A^*A$  имеет положительно определенный квадратный корень, т.е. матрицу  $S$ , которая симметрична, положительно определена и удовлетворяет условию  $S^2 = A^*A$ . (Действительно, мы можем написать, что  $A^*A = OLO^*$ , где  $L$  — диагональная матрица с положительными диагональными элементами и  $O$  — ортогональная матрица. Тогда  $S = OMO^*$ , где  $M$  — диагональная матрица, элементы которой равны квадратным корням из соответствующих элементов матрицы  $L$ .) Утверждается, что матрица  $AS^{-1}$  ортогональна. Действительно,

$$(AS^{-1}x, AS^{-1}y) = (S^{-1}x, A^*AS^{-1}y) = (S^{-1}x, S^2S^{-1}y) = (x, y),$$

поскольку  $S^{-1}$  — самосопряженная матрица. Тогда можно написать

$$A = SK,$$

где  $K = AS^{-1}$  ортогональна. Однако  $S = OMO^*$ . Поэтому

$$A = OMO^*K,$$

где  $O$  (и  $O^* = O^{-1}$ ) и  $K$  ортогональны, а  $M$  — диагональная матрица с положительными элементами. Для  $M$  мы имеем  $\text{Vol}(M) = \text{Det } M$ . Ортогональные матрицы сохраняют длину и поэтому сохраняют объем, т.е.  $\text{Vol}(O) = \text{Vol}(K) = 1$ . С другой стороны,  $\text{Det } O = \pm 1$  (потому что  $\text{Det } O \cdot \text{Det } O^* = \text{Det } OO^* = 1$  и  $\text{Det } O = \text{Det } O^*$ ). В этом и состоит другое доказательство формулы (\*).

(ii) Пусть  $\phi = A : \mathbb{R}^k \rightarrow \mathbb{R}^k$  — линейное отображение. Тогда

$$\phi^*(dx^1 \wedge \cdots \wedge dx^k) = (\text{Det } A)dx^1 \wedge \cdots \wedge dx^k. \quad (**)$$

Из определения процедуры переноса и внешнего произведения следует, что  $\phi^*x^1 = a_{11}x^1 + a_{12}x^2 + \dots + a_{1k}x^k$  и т. д. Поскольку  $a_{ij}$  постоянны, получаем

$$\begin{aligned}\phi^*dx^1 &= a_{11}dx^1 + \dots + a_{1k}dx^k, \\ &\vdots \\ \phi^*dx^k &= a_{k1}dx^1 + \dots + a_{kk}dx^k.\end{aligned}$$

Теперь  $\phi^*dx^1 \wedge \dots \wedge \phi^*dx^k$  будет кратно  $dx^1 \wedge \dots \wedge dx^k$ . Соответствующий множитель обозначим  $f(A)$ . Это некоторая числовая функция  $A$ , которая удовлетворяет всем аксиомам определителя. В силу единственности такой функции, доказанной в главе 11, мы считаем, что формула (\*\*) справедлива.

(iii) Пункты (i) и (ii) дают формулу замены переменных для случая, когда  $\phi$  — линейное отображение, т. е. для  $\phi = A$  при условии, что  $\text{Det } A > 0$ . Докажем это. Запишем  $\omega = g dx^1 \wedge \dots \wedge dx^k$ , где  $g$  — функция. В силу (ii)

$$\phi^*\omega = h \text{Det } A dx^1 \wedge \dots \wedge dx^k, \quad \text{где } h(x) = g(Ax).$$

Покроем область  $U$  сеткой с малыми ячейками, как это показано на рис. 15.6. Это значит, что область  $\phi(U)$  делится на малые параллелепипеды, причем объем каждого параллелепипеда отличается от объема соответствующего кубика множителем  $\text{Det } A$ .

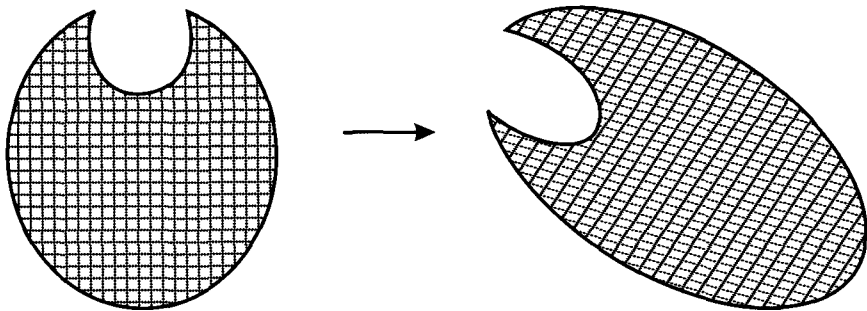


Рис. 15.6

Возьмем произвольное значение  $\varepsilon > 0$ . Мы можем сделать размер ячейки достаточно малым, чтобы в каждом параллелепипеде функция  $g$  менялась бы не больше, чем на  $\varepsilon$ :

$$|g(\mathbf{p}) - g(\mathbf{q})| < \varepsilon$$

при условии, что  $\mathbf{p}$  и  $\mathbf{q}$  лежат в одном и том же параллелепипеде. Это значит, что если  $\square$  обозначает один из внутренних кубиков, а  $\phi(\square)$  обозначает соответствующий параллелепипед, то

$$\left| \int_{\phi(\square)} \omega - g(\mathbf{p}) \text{Vol}(\phi(\square)) \right| < \varepsilon, \quad (***)$$

где  $\mathbf{p}$  — произвольная точка из  $\phi(\square)$ . Однако в силу (i)

$$\text{Vol}(\phi(\square)) = \text{Det } A \cdot \text{Vol}(\square).$$

Поэтому

$$\left| \int_{\phi(\square)} \omega - g(\mathbf{p}) \text{Det } A \cdot \text{Vol}(\square) \right| < \varepsilon. \quad (****)$$

Пусть  $U'$  обозначает объединение всех внутренних кубиков сетки. Суммируя неравенства (\*\*\*\*) по всем внутренним кубикам, получаем

$$\left| \int_{\phi(\square)} \omega - \sum g(\mathbf{Ar}) \text{Det } A \cdot \text{Vol}(\square) \right| < \varepsilon,$$

где  $\mathbf{r}$  — любая точка в кубике  $\square$ . Суммирование ведется по всем внутренним кубикам. Когда сетка измельчается, в пределе кубиками вдоль границы можно пренебречь, и тогда

$$\int_{\phi(U')} \omega \rightarrow \int_{\phi(U)} \omega,$$

а

$$\sum g(\mathbf{Ar}) \text{Det } A \cdot \text{Vol}(\square) \rightarrow \int \phi^* \omega.$$

(iv) Давайте посмотрим, что надо изменить в предыдущих рассуждениях, если  $\phi$  уже не предполагается линейным. Прежде

всего,  $\phi^*\omega = h \operatorname{Det}(\partial\phi/\partial x) dx^1 \wedge \dots \wedge dx^k$ . Поэтому постоянная  $\operatorname{Det} A$  заменяется функцией  $\operatorname{Det}(\partial\phi/\partial x)$ , где  $(\partial\phi/\partial x)$  обозначает матричную функцию

$$(\partial\phi/\partial x) = \begin{pmatrix} \partial\phi^1/\partial x^1 & \dots & \partial\phi^1/\partial x^k \\ \partial\phi^k/\partial x^1 & \dots & \partial\phi^k/\partial x^k \end{pmatrix}.$$

В этом случае аппроксимирующая риманова сумма становится равной

$$\sum g(A\mathbf{r}) \operatorname{Det}(\partial\phi/\partial x) \operatorname{Vol}(\square).$$

Неравенство (\*\*\*) все еще выполняется, если сетка достаточно мелкая. Но  $\phi(\square)$  уже не будет параллелепипедом, и у нас уже не будет точного равенства

$$\operatorname{Vol}(\phi(\square)) = \operatorname{Det}(\partial\phi/\partial x)(\mathbf{r}) \operatorname{Vol}(\square).$$

Но нам достаточно знать, что это равенство выполняется приблизительно, т. е. что

$$|\operatorname{Vol}(\phi(\square)) - \operatorname{Det}(\partial\phi/\partial x)(\mathbf{r}) \operatorname{Vol}(\square)| < \varepsilon \operatorname{Vol}(\square).$$

Это неравенство следует из теоремы о среднем значении. Детали доказательства достаточно просты. Их можно найти в книге Loomis and Sternberg, *Advanced Calculus*, глава 8, с. 343–344. Мы еще раз отсылаем читателя к этой главе для последовательного изучения теории интегрирования в пространстве  $\mathbb{R}^k$ . Этим мы завершаем обсуждение формулы замены переменных.

**Интегрирование  $k$ -форм по  $k$ -цепям комплекса.** Предположим, что  $\omega$  —  $k$ -форма в пространстве  $\mathbb{R}^n$ . Пусть  $\phi: \mathbb{R}^k \rightarrow \mathbb{R}^n$  — гладкое отображение, а  $U$  обозначает область в пространстве  $\mathbb{R}^k$ . Рассмотрим интеграл

$$\int_U \phi^*\omega.$$

Пусть  $K$  — комплекс, а  $\phi$  — отображение этого комплекса в пространство  $\mathbb{R}^n$ . Напомним: это значит, что  $\phi$  отображает каждую ячейку комплекса  $C$  в пространство  $\mathbb{R}^n$ , причем ограничение  $\phi$  на  $C$  — дифференцируемое отображение.

Это имеет смысл, потому что каждая ячейка  $C$  является выпуклым многогранником в некотором евклидовом пространстве. По существу, мы требуем, чтобы отображение  $\phi$  удовлетворяло более сильному (с технической точки зрения) условию: для каждой  $k$ -ячейки  $C$  комплекса существует область  $U$  (окрестность  $C$ ) в пространстве  $\mathbb{R}^k$  и такое дифференцируемое отображение  $\psi : U \rightarrow \mathbb{R}^n$ , что ограничения  $\phi$  и  $\psi$  на  $C$  совпадают. Каждая  $k$ -ячейка  $C$  комплекса ориентирована. Поэтому можно рассмотреть интеграл

$$\int_C \phi^* \omega$$

по ячейке  $C$ . Пусть  $\mathbf{c}$  — некоторая  $k$ -цепь. Тогда  $\mathbf{c}$  является линейной комбинацией  $k$ -ячеек:

$$\mathbf{c} = \sum r_j C_j,$$

где  $C_j$  —  $k$ -ячейки, а  $r_j$  — вещественные числа. Тогда интеграл от  $\phi^* \omega$  по цепи  $\mathbf{c}$  дается формулой

$$\int_{\mathbf{c}} \phi^* \omega = \sum r_j \int_{C_j} \phi^* \omega.$$

Очевидно, что это выражение линейно зависит от  $\mathbf{c}$ . Таким образом,  $k$ -форма  $\omega$  вместе с отображением  $\phi$  определяет  $k$ -коцепь на комплексе  $K$ .

Например, предположим, что наш комплекс  $K$  содержит грани, ребра и вершины тетраэдра. Для простоты пусть все грани ориентированы «согласованно». Под этим подразумевается, что

$$\text{если } \mathbf{c} = C_1 + C_2 + C_3 + C_4, \quad \text{то } \partial \mathbf{c} = 0.$$

Представим себе, что центр тетраэдра находится в начале координат пространства  $\mathbb{R}^3$ .

**Независимость от деления на ячейки.** Рассмотрим проекцию (рис. 15.7) из центра тетраэдра на поверхность единичной сферы  $S$  и назовем ее отображением  $\phi$ . Легко видеть, что  $\phi$  — гладкое отображение тетраэдра в пространство  $\mathbb{R}^3$ . Каждая грань тетраэдра отображается в часть сферы, а  $\phi$  будем рассматривать как

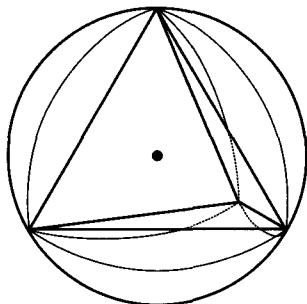


Рис. 15.7

отображение цепи  $\mathbf{c}$  на поверхность сферы (с определенным выбором ориентации). Если  $\omega$  — 2-форма в  $\mathbb{R}^3$ , то  $\int_{\mathbf{c}} \phi^* \omega$  — интеграл от  $\omega$  по сфере (также с выбором ориентации), что записывается в виде

$$\int_S \omega.$$

Это обозначение неявно содержит ряд правомерных допущений. Прежде всего, предполагается, что граничные кривые треугольных областей, на которые мы разделили сферу, можно игнорировать. Связано это с тем, что при вычислении двумерного интеграла определение через римановы суммы означает, что гладкие кривые не вносят вклада. Более важное предположение состоит в том, что конкретный выбор разбиения на треугольные области (комплекса  $K$  и отображения  $\phi$ ) не имеет значения. Вследствие теоремы о замене переменных нас интересует только поверхность сферы (покрываемая один раз) с определенной ориентацией. Давайте объясним это. Допустим, что мы рассматриваем другой комплекс  $L$ , состоящий из граней, ребер и вершин куба, с центром в начале координат. Обозначим  $\mu$  отображение комплекса  $L$  на поверхность сферы, задаваемое проекцией куба из центра (рис. 15.8). Предположим также, что грани комплекса  $L$  ориентированы согласованно, так что

$$\mathbf{e} = F_1 + F_2 + F_3 + F_4 + F_5 + F_6 \quad \text{удовлетворяет условию } \partial \mathbf{e} = 0,$$

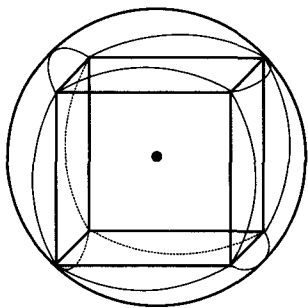


Рис. 15.8

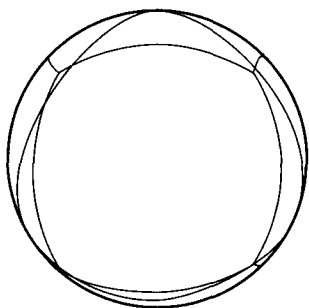


Рис. 15.9

где  $F_i$  ( $i = 1, 2, \dots, 6$ ) — грани куба. Утверждается, что

$$\int_{\mathbf{e}} \mu^* \omega = \pm \int_{\mathbf{c}} \phi^* \omega \quad (15.11)$$

со знаком  $+$ , если ориентации  $\mathbf{e}$  и  $\mathbf{c}$  «совпадают». (В процессе доказательства мы объясним, что это значит.)

Действительно, рассмотрим кривые на сфере, являющиеся образами ребер комплекса  $K$  (при отображении  $\phi$ ) и комплекса  $L$  (при отображении  $\mu$ ). Эти кривые делят поверхность сферы на много подобластей (рис. 15.9).

Каждая треугольная область на сфере состоит из нескольких таких подобластей, и точно так же каждая «квадратная область», происходящая от куба. Пусть  $W$  обозначает одну из таких подобластей. Тогда мы можем написать

$$W = \phi(U), \quad \text{где } U \text{ — подобласть одной из граней } K$$

и

$$W = \mu(V), \quad \text{где } V \text{ — подобласть одной из граней } L.$$

Рассмотрим отображение  $\psi : U \rightarrow V$ , заданное  $\psi = \mu^{-1} \circ \phi$ . Тогда

$$\phi = \mu \circ \psi,$$

откуда

$$\phi^* \omega = \psi^*(\mu^* \omega) \quad \text{и} \quad V = \psi(U).$$

(Легко проверить, что  $\psi$  — дифференцируемое отображение.)

Из формулы замены переменных имеем

$$\int_V \mu^* \omega = \pm \int_U \phi^* \omega$$

где знак  $+$  берется тогда и только тогда, когда  $\psi$  сохраняет ориентацию. Легко видеть, что отображение  $\psi$ , которое мы построили для подобласти  $W$ , будет сохранять ориентацию тогда и только тогда, когда соответствующие отображения для всех других под областей тоже сохраняют ориентацию. Если это происходит, то мы говорим, что  $s$  и  $e$  создают одинаковую ориентацию на сфере. Суммирование по всем под областям доказывает (15.11) и объясняет введенное обозначение. Мы уже кратко набросали доказательство справедливости нашего обозначения для случая сферы. Конечно, оно работает и в гораздо более общем случае. Для любой ориентированной поверхности в трехмерном пространстве (в общем случае для любого ориентированного подмногообразия  $M$  в пространстве  $\mathbb{R}^n$ ) интеграл  $\int_M \omega$  имеет смысл и не зависит от того, как мы «нарезали  $M$  на куски», чтобы записать его как объединение образов ячеек.

(Итак, мы определили интеграл от 2-формы по ориентированной 2-сфере, «разрезав ее на куски» и потом показав, что результат не зависит от способа разрезания. Существует другое (эквивалентное) определение, которое состоит в «покрытии сферы заплатами», а не в разрезании ее на куски. В конце следующего параграфа мы коротко расскажем об этом.)

## 15.4. Теорема Стокса

Теорема Стокса, которую мы докажем в этом параграфе в общем случае, говорит следующее. Пусть  $K$  — некоторый комплекс, а  $\phi$  — гладкое отображение  $K$  в пространство  $\mathbb{R}^n$ . Пусть, далее,  $\omega$  —  $(k-1)$ -форма в  $\mathbb{R}^n$ , а  $c$  —  $k$ -цепь. Тогда

$$\int_c \phi^* d\omega = \int_{\partial c} \phi^* \omega. \quad (15.12)$$

Если фиксированы  $\phi$  и  $\omega$ , то обе части этого равенства — линейные функции  $c$ . Поскольку любая цепь является линейной комбинацией ячеек, то равенство (15.12) достаточно доказать для

случая, когда  $c$  — ячейка. Доказывать будем так же, как это делалось в двумерном случае в главе 8, и только в самом конце будет небольшое отклонение.

В качестве первого шага рассмотрим случай, когда ячейка  $C$  является кубом. Можно предположить, что это единичный куб в пространстве  $\mathbb{R}^k$ . Пусть  $u^1, \dots, u^k$  — координаты в  $\mathbb{R}^k$ . Тогда можно написать

$$\begin{aligned} \phi^* \omega = a_1 du^2 \wedge \dots \wedge du^k + a_2 du^1 \wedge du^3 \wedge \dots \wedge du^k + \dots \\ + a_k du^1 \wedge \dots \wedge du^{k-1}, \end{aligned} \quad (15.13)$$

так что левая часть равенства (15.12) принимает вид

$$\begin{aligned} \int_C \phi^* d\omega = \int_C d\phi^* \omega \\ = \int_C \left( \frac{\partial a_1}{\partial u^1} - \frac{\partial a_2}{\partial u^2} + \dots \pm \frac{\partial a_k}{\partial u^k} \right) du^1 \wedge \dots \wedge du^k. \end{aligned} \quad (15.14)$$

Напомним правило выбора знаков для оператора  $\partial$ , действующего на ячейку комплекса. Вспоминаем, что ориентация ячейки  $C$  определяет, «правильно» или «неправильно» упорядочены базисные вектора. «Неправильное» множество базисных векторов преобразуется в «правильное» либо перестановкой двух векторов с соседними номерами, либо изменением знака одного вектора. Правильные и неправильные множества для  $k = 1, 2, 3$  показаны на рис. 15.10.

Граница ориентированной  $k$ -ячейки состоит из суммы ориентированных  $(k-1)$ -ячеек. Общее правило определения знака каждой ячейки на границе: в точке на границе ячейки  $C$  строим множество векторов, в котором первый вектор направлен из ячейки  $C$ , а остальные  $k-1$  векторов образуют правильно ориентированный базис граничной ячейки. Если получившееся множество  $k$  векторов является правильно ориентированным множеством для ячейки  $C$ , то граничная ячейка входит в  $\partial C$  со знаком плюс; в противном случае она входит со знаком минус.

На рис. 15.11 показано применение этого правила в случае  $k = 1$ . Вектор  $v_1$  направлен наружу, а других векторов нет. В

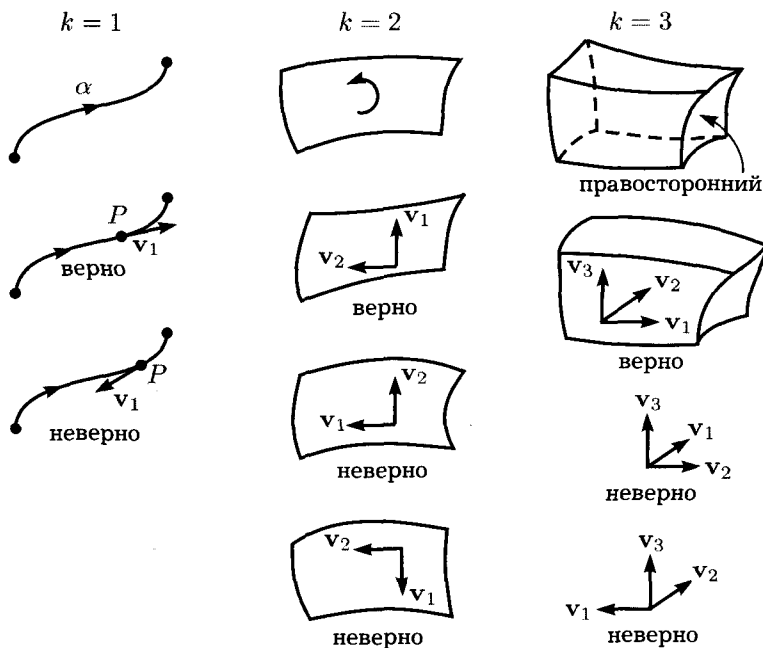


Рис. 15.10

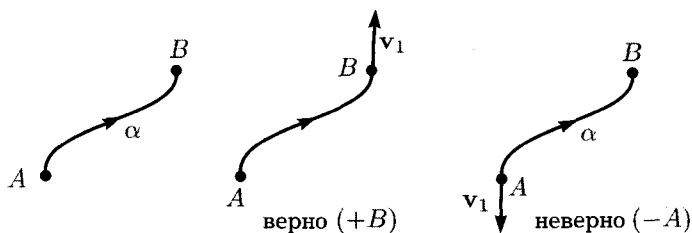


Рис. 15.11

точке  $B$  вектор  $v_1$  имеет правильную ориентацию относительно ветви  $\alpha$ , а в точке  $A$  — неправильную. Следовательно,  $\partial\alpha = B - A$ .

На рис. 15.12 показан случай  $k=2$  для треугольной ячейки  $C$ , ориентированной против часовой стрелки.

Вектор  $v_2$  выбран так, что он касателен к ветви, а его направление определяет ориентацию ветви. Вектор  $v_1$  всегда направлен из ячейки  $C$ . Для ветвей  $\alpha$  и  $\beta$  эта пара векторов правильно

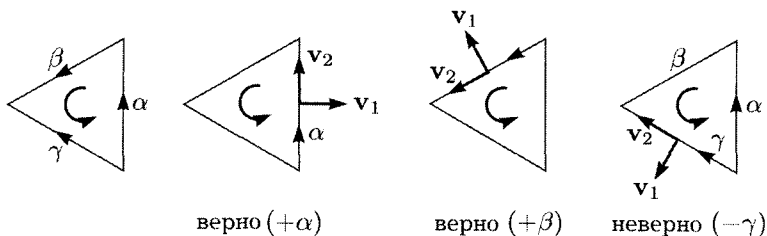


Рис. 15.12

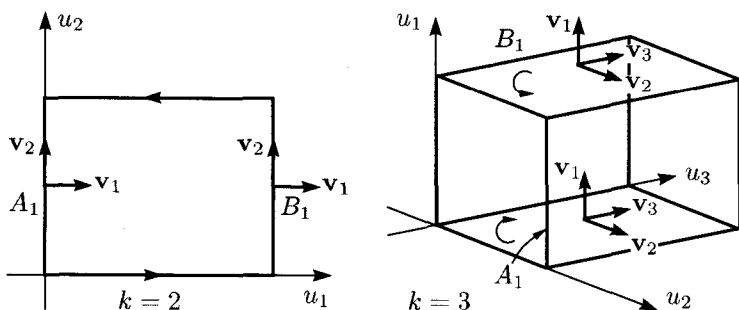


Рис. 15.13

ориентирована (в соответствии с ориентацией ячейки  $C$  против часовой стрелки), а вот для ветви  $\gamma$  вектора ориентированы неправильно. Следовательно,  $\partial C = \alpha + \beta - \gamma$ .

Вернемся к рассмотрению случая, когда ячейка  $C = I^k$  является кубом, ориентированным посредством стандартного базиса  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  (его векторы параллельны ребрам куба). Рассмотрим отдельные члены, получающиеся в результате подстановки (15.13) в  $\int_{\partial C} \phi^* \omega$ . Рассмотрим различные грани  $\partial C$ . Начнем с двух граней, для которых  $u^1 = \text{const}$ . На грани  $B_1$ , где  $u^1 = 1$ , вектор  $\mathbf{v}_1$  направлен *наружу* и остальные вектора  $\mathbf{v}_2, \dots, \mathbf{v}_k$  ориентированы правильно. На грани  $A_1$ , где  $u^1 = 0$ , вектор  $\mathbf{v}_1$  направлен *внутрь* и поэтому  $\mathbf{v}_2, \dots, \mathbf{v}_k$  ориентированы неправильно. Эта ситуация в случаях  $k = 2$  и  $k = 3$  показана на рис. 15.13.

Когда мы вычисляем  $\phi^* \omega$  на этих двух гранях с уравнениями  $u^1 = \text{const}$ , все члены в  $\tau$ , содержащие множитель  $du^1$ , дают нуль. Остается только  $a_1(u^1, u^2, \dots, u^k) du^2 \wedge \dots \wedge du^k$ . Вычисление этого члена на грани  $B_1$ , где  $\mathbf{v}_2, \dots, \mathbf{v}_k$  упорядочены правильно,

дает  $\int_{I^{k-1}} a_1(1, u^2, \dots, u^k) du^2 \dots du^k$ . А вычисление его на грани  $A_1$ , где вектора  $\mathbf{v}_2, \dots, \mathbf{v}_k$  упорядочены неправильно, дает  $-\int_{I^{k-1}} a_1(0, u^2, \dots, u^k) du^2 \dots du^k$ . Все вместе дает интеграл

$$\int_{I^{k-1}} [a_1(1, u^2, \dots, u^k) - a_1(0, u^2, \dots, u^k)] du^2 \dots du^k.$$

Из фундаментальной теоремы дифференциального исчисления имеем

$$a_1(1, u^2, \dots, u^k) - a_1(0, u^2, \dots, u^k) = \int_0^1 \frac{\partial a_1}{\partial u^1}(u^1, u^2, \dots, u^k) du^1.$$

Поэтому общий вклад от двух граней с  $u_1 = \text{const}$  равен

$$\int_{I^k} \frac{\partial a_1}{\partial u^1} du^1 \dots du^k,$$

где теперь интеграл вычисляется по единичному кубу в пространстве  $\mathbb{R}^k$ .

Выполнив такую же процедуру для граней с  $u^2 = \text{const}$ , заметим, что множество векторов  $\mathbf{v}_2, \mathbf{v}_1, \mathbf{v}_3, \dots, \mathbf{v}_k$  упорядочено *неправильно*. На грани, где  $u^2 = 1$ , вектор  $\mathbf{v}_2$  направлен наружу и поэтому остальные вектора  $\mathbf{v}_1, \mathbf{v}_3, \dots, \mathbf{v}_k$  упорядочены неправильно. На грани, где  $u^2 = 0$ , вектор  $\mathbf{v}_2$  направлен внутрь, а остальные вектора  $\mathbf{v}_1, \mathbf{v}_3, \dots, \mathbf{v}_k$  упорядочены правильно. При вычислении  $\tau$  вклад дает только член  $a_2(u^1, u^2, \dots, u^k) du^1 \wedge du^3 \wedge \dots \wedge du^k$ . Поэтому мы получаем

$$\begin{aligned} \int_{I^{k-1}} [-a_2(u^1, 1, \dots, u^k) + a_2(u^1, 0, \dots, u^k)] du^1 du^3 \dots du^k \\ = - \int_{I^k} \frac{\partial a_2}{\partial u^2} du^1 du^2 \dots du^k. \end{aligned}$$

Аналогичные соображения показывают, что вклад от граней, где  $u^j = \text{const}$ , равен  $\pm \int_{I^k} (\partial a_j / \partial u^j) du^1 \dots du^k$ , причем берется знак  $+$  для нечетных  $j$ , и знак  $-$  для четных. Если просуммировать вклады от всех  $2k$  граней в  $\partial I^k$ , то получится (15.12).

Итак, мы доказали теорему Стокса для куба. Теперь надо рассмотреть ячейки более общего вида. Как мы показали в главе 8, на плоскости любой многоугольник может быть разложен

на треугольники. Действительно, любой многоугольник можно разложить на выпуклые многоугольники, которые, в свою очередь, разлагаются на треугольники, если выбрать точку внутри и соединить ее со всеми вершинами.

Аналогично в трехмерном пространстве любой многогранник может быть разложен на тетраэдры. Действительно, сделаем предварительное разложение и предположим, что  $C$  — выпуклый многогранник. Предположим, что все его грани уже разложены на треугольники. Опять выбираем точку  $p$  внутри  $C$ . Тогда этот многогранник разлагается на тетраэдры, основанием каждого из которых является треугольник из разложения грани, а вершина находится в точке  $p$ . Таким образом, в случае 2-форм теорему Стокса достаточно доказать для тетраэдра.

С помощью индукции мы можем проделать эту процедуру и в пространстве  $k$  измерений. Пусть  $S \in \mathbb{R}^k$  будет множеством, определенным формулой

$$S = \left\{ x \mid x_i \geq 0 \text{ для всех } i \text{ и } \sum x_i \leq 1 \right\}.$$

$k$ -симплекс — это подмножество пространства  $\mathbb{R}^k$ , являющееся образом  $S$  при обратимом аффинном отображении. Таким образом 1-симплекс — это отрезок, 2-симплекс — это треугольник и 3-симплекс — это тетраэдр.

Предположим, что мы разложили  $C$  на симплексы  $[\Delta_i]$ . Утверждается, что если мы знаем теорему Стокса для всех симплексов, то мы знаем ее для  $C$ . Действительно, рассмотрим комплекс, который разлагается на ячейки — симплексы. Тогда комплекс  $C$  можно идентифицировать с цепью в нем, т. е. можно написать

$$C = \sum \Delta_i, \quad \partial C = \sum \partial \Delta_i.$$

Поэтому

$$\begin{aligned} \int_{\partial C} \omega &= \sum \int_{\partial \Delta_i} \omega \\ &= \sum \int_{\Delta_i} d\omega \quad (\text{в силу теоремы Стокса для } \Delta_i) \\ &= \int_C d\omega. \end{aligned}$$

Итак, теорему Стокса достаточно доказать для симплекса.

Воспользуемся инвариантностью интеграла относительно процедуры переноса и сведем доказательство к случаю куба. Здесь мы рассмотрим доказательство для треугольника.

Чтобы доказать теорему Стокса для треугольника, достаточно доказать ее для равностороннего треугольника  $\Delta$ , поскольку для любого треугольника можно найти аффинное преобразование, которое преобразует его в равносторонний. Предположим, что длина стороны треугольника равна 1. Пусть линейная дифференциальная форма  $\tau$  тождественно равна нулю за пределами круга, центр которого находится в одной из вершин и который не достигает противоположной стороны (рис. 15.14).

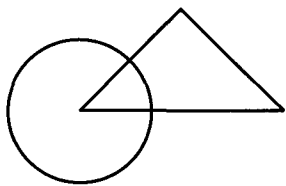


Рис. 15.14

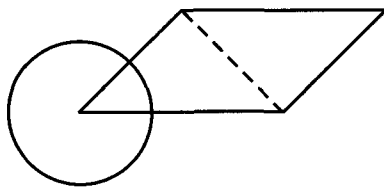


Рис. 15.15

Тогда  $d\tau$  также тождественно равняется нулю за пределами круга. В силу теоремы Стокса в этой ситуации вклад в интеграл возникает только от внутренней части круга, так что мы можем без ущерба нарастить «внешнюю» часть треугольника, построив его до параллелограмма (рис. 15.15).

Таким образом, если  $\square$  обозначает параллелограмм, а  $\Delta$  обозначает треугольник, то поскольку  $\tau$  и  $d\tau$  равны нулю за пределами круга, мы имеем

$$\int_{\partial\square} \tau = \int_{\partial\Delta} \tau \quad \text{и} \quad \int_{\square} d\tau = \int_{\Delta} d\tau.$$

Однако, с точностью до замены переменных параллелограмм — это просто квадрат. Следовательно,  $\int_{\partial\square} \tau = \int_{\square} d\tau$  и поэтому  $\int_{\partial\Delta} \tau = \int_{\Delta} d\tau$ .

Теорема Стокса будет доказана для  $\Delta$ , если мы установим следующее. Каждая гладкая дифференциальная форма (определенная в окрестности  $\Delta$ ) может быть записана в виде суммы трех

членов

$$\tau = \tau_1 + \tau_2 + \tau_3,$$

где каждый член  $\tau_i$  тождественно равен нулю вне круга  $D_i$ . Причем центр круга находится в вершине треугольника, а радиус  $R$  изменяется<sup>8</sup> в пределах  $2/3 < R < 3/4$  (так что три круга покрывают весь треугольник, но ни один круг не достигает противоположной стороны треугольника, см. рис 15.16).

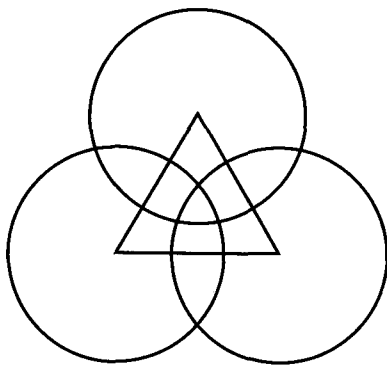


Рис. 15.16

Чтобы это доказать, потребуется свойство функции одной переменной.

**Лемма.** *Функция*

$$f(u) = \begin{cases} e^{-1/u}, & u \geq 0, \\ 0, & u \leq 0 \end{cases}$$

*бесконечно дифференцируема во всех точках.*

**Доказательство.** Очевидно, что для  $u \neq 0$  функция  $f$  имеет производные всех порядков и  $f^{(k)}(u) \equiv 0$  для  $u < 0$ . Поэтому мы должны доказать, что  $f^{(k)}(u)/u \rightarrow 0$  при  $u \rightarrow 0$  для любого  $k$ . Мы имеем  $f^{(k)}(u) = P_k(1/u)e^{-1/u}$ , где  $P_k$  — некоторый полином<sup>9</sup>,

<sup>8</sup>Более последовательный подход — зафиксировать радиус  $R = 1/\sqrt{2}$ . Причина объясняется в сноске на с. 786. — *Прим. ред.*

<sup>9</sup>Полезное упражнение — вычислить его в явном виде индукцией по  $k$ . — *Прим. ред.*

и поэтому

$$\lim_{u \rightarrow 0} (1/u) \cdot f^{(k)}(u) = \lim_{s \rightarrow \infty} s P_k(s) e^{-s} = 0,$$

потому что  $e^s$  стремится к бесконечности быстрее любого полинома.

Заметим, что  $f$  тождественно равна нулю для  $u \leq 0$  и строго положительна для  $u > 0$ .

Пусть  $r_i$  обозначает расстояние от  $i$ -й вершины. Определим функцию  $g_i$  формулой

$$g_i(x) = \begin{cases} f(R - r_i(x)), & r_i(x) \leq R, \\ 0, & r_i(x) > R. \end{cases}$$

Тогда  $g_i$  бесконечно дифференцируема и

$$g_i(x) \begin{cases} > 0, & r_i(x) < R, \\ = 0, & r_i(x) \geq 0. \end{cases}$$

Пусть

$$g = g_1 + g_2 + g_3.$$

Поскольку каждая точка в  $\Delta$  — внутренняя для по крайней мере одного из кругов  $D_i$ , т. е. поскольку  $r_i(x) < R$  хотя бы для одного из  $i = 1, 2, 3$ , то

$$g(x) > 0$$

для каждого  $x \in \Delta$ . Поэтому мы можем разделить на  $g$  и определить

$$\phi_i = \frac{g_i}{g}, \quad L = 1, 2, 3.$$

Тогда

$$\phi_i(x) \begin{cases} > 0, & r_i(x) < R, \\ = 0, & r_i(x) \geq 0 \end{cases}$$

и

$$\phi_1(x) + \phi_2(x) + \phi_3(x) \equiv 1 \quad \text{для всех } x \in \Delta.^{10}$$

Умножая последнее уравнение на  $\tau$ , получаем

$$\tau = \phi_1 \tau + \phi_2 \tau + \phi_3 \tau.$$

<sup>10</sup> Построенный набор функций  $\phi_i$  принято называть *разбиением единицы*, подчиненным данному покрытию многообразия (треугольника) множествами  $D_i$  (тремя кругами). Это понятие играет важную роль в анализе на многообразиях. Подробнее см. в конце этого раздела на с. 791. — *Прим. ред.*

Если мы положим  $\tau_i = \phi_i \tau$ , то получим желаемое разложение

$$\tau = \tau_1 + \tau_2 + \tau_3.$$

Очевидно, что это доказательство можно провести и в случае  $n$  измерений. Только вместо 3 будет соответственно  $n+1$  (по числу вершин симплекса) шаров  $D_i$ , функций  $\phi_i$  и т. д., а радиус шаров можно выбрать  $R = 1/\sqrt{2}$ .<sup>11</sup> Итак, теорема Стокса доказана.

На самом деле мы не должны быть настолько педантичными. Надо просто покрыть симплекс конечным числом шаров  $D_i$ , так чтобы пересечение  $D_i$  с симплексом имело вид части куба, разбить на куски и свести к случаю куба.

## Пример

В качестве применения теоремы Стокса и вычисления форм на ячейках рассмотрим 2-форму  $\tau = (x^2 + y^2 + z^2) dx \wedge dy$  в пространстве  $\mathbb{R}^3$  и ячейку  $C$ , являющуюся сплошным полушарием радиуса  $R$ , обладающую правосторонней ориентацией и изображенную на рис. 15.17. В явном виде мы вычислим  $\int_{\partial C} \tau$  и  $\int_C d\tau$ .

Сплошное полушарие ограничено двумя ячейками: полусферической поверхностью  $A$  и диском  $B$  в экваториальной плоскости. Пусть эти ячейки ориентированы против часовой стрелки, если смотреть на них снаружи, тогда  $\partial C = A + B$ .

<sup>11</sup>Происхождение данной величины, вероятно, следует пояснить, учитывая небрежность авторов в этом месте оригинала. Если симплекс правильный (как  $n$ -мерный многогранник) и длина его ребра равна 1, то его высота

$$h_n = \sqrt{\frac{n+1}{2n}},$$

а расстояние от вершины до центра симплекса

$$d_n = \frac{n}{n+1} h_n = \sqrt{\frac{n}{2(n+1)}}.$$

То, что шары радиуса  $R = 1/\sqrt{2}$  с центрами в вершинах симплекса покрывают симплекс и не достигают противоположных граней, вытекает из очевидных неравенств  $d_n < 1/\sqrt{2} < h_n$ , причем обе последовательности сходятся к  $1/\sqrt{2}$  при  $n \rightarrow \infty$ . Геометрически это выглядит так: с ростом  $n$  прямоугольный треугольник «вершина симплекса – центр противоположной грани – вершина грани» с гипотенузой 1 и катетами  $h_n$  и  $d_{n-1}$  стремится к равнобедренному.

— Прим. ред.

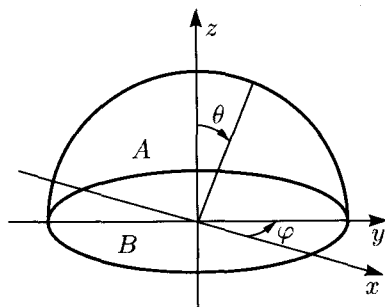


Рис. 15.17

В качестве параметров для описания полусферической поверхности воспользуемся сферическими координатами  $\theta$  и  $\phi$ . Двумерное параметрическое пространство изображено на рис. 15.18. В принципе можно ввести параметры  $u = 2\theta/\pi$  и  $v = \phi/2\pi$ , так что областью интегрирования будет единичный квадрат, но прямоугольник не менее удобен для вычисления.

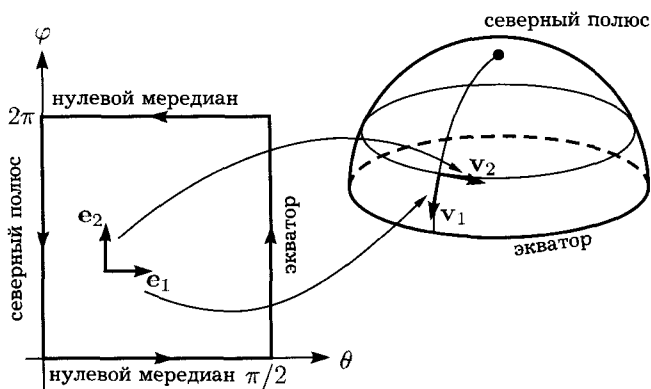


Рис. 15.18

Перенос  $x, y, z$  при указанной параметризации  $\alpha$  имеет вид:  $\alpha^*x = R \sin \theta \cos \phi$ ,  $\alpha^*y = R \sin \theta \sin \phi$ ,  $\alpha^*z = R \cos \theta$ . Тогда

$$\alpha^*dx = R \cos \theta \cos \phi d\theta - R \sin \theta \sin \phi d\phi,$$

$$\alpha^*dy = R \cos \theta \sin \phi d\theta + R \sin \theta \cos \phi d\phi.$$

Следовательно,

$$\begin{aligned}\alpha^*\tau &= \alpha^*(x^2 + y^2 + z^2)(\alpha^*dx) \wedge (\alpha^*dy) \\ &= R^2(R^2 \sin \theta \cos \theta \cos^2 \phi d\theta \wedge d\phi - R^2 \sin \theta \cos \theta \sin^2 \phi d\phi \wedge d\theta) \\ &= R^4 \sin \theta \cos \theta d\theta \wedge d\phi.\end{aligned}$$

Чтобы проверить, что  $\alpha$  сохраняет ориентацию, посмотрим на образы упорядоченного множества векторов  $\mathbf{e}_1$  и  $\mathbf{e}_2$ . Образы  $\mathbf{v}_1$  и  $\mathbf{v}_2$  соответствуют ориентации полусферы  $A$ , что показано на рис. 15.18.

Для вычисления 2-формы  $\tau$  на полусфере  $A$  мы просто вычислим двойной интеграл  $\int \alpha^*\tau = \int_{\theta=0}^{\pi/2} \int_{\phi=0}^{2\pi} R^4 \sin \theta \cos \theta d\theta d\phi = \pi R^4$ .

Для диска  $B$  удобно выбрать параметры  $r^2 = x^2 + y^2$  и угол  $\phi$ . Чтобы диск  $B$  имел правильную ориентацию, выбираем порядок  $\phi, r$ . Соответствующее параметрическое пространство изображено на рис. 15.19. Тогда образы  $\mathbf{e}_1$  и  $\mathbf{e}_2$  правильно упорядочены для ориентации  $B$  против часовой стрелки, если смотреть на диск снизу, и по часовой стрелке, если смотреть на него сверху.

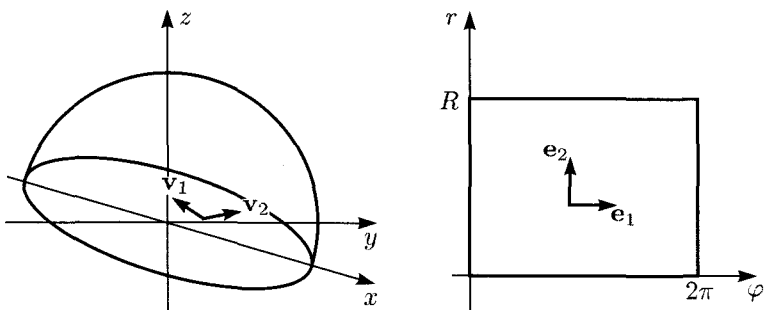


Рис. 15.19

Для параметризации  $\beta$  перенос имеет вид:  $\beta^*x = r \cos \phi$ ,  $\beta^*y = r \sin \phi$ ,  $\beta^*z = 0$ , следовательно,  $\beta^*\tau = r^2(\cos \phi dr - r \sin \phi d\phi) \wedge (\sin \phi dr + r \cos \phi d\phi) = r^3 dr \wedge d\phi = -r^3 d\phi \wedge dr$ . Таким образом, значение  $\tau$  для ориентированного диска  $B$  равно

$$\int_{r=0}^R \int_{\phi=0}^{2\pi} -r^3 d\phi dr = -\frac{1}{2}\pi R^4.$$

Собрав вместе результаты для двух ячеек на границе, мы получаем

$$\int_{\partial C} = \int_A \tau + \int_B \tau = \pi R^4 - \frac{1}{2}\pi R^4 = \frac{1}{2}\pi R^4.$$

Теперь вычислим  $\int_C d\tau$ . Поскольку  $\tau = (x^2 + y^2 + z^2) dx \wedge dy$ ,  $d\tau = 2z dz \wedge dx \wedge dy = 2z dx \wedge dy \wedge dz$ . Мы параметризуем ячейку  $C$ , используя сферические координаты  $r, \theta, \phi$ . Перенос, соответствующий данной параметризации:  $\psi^*x = r \sin \theta \cos \phi$ ,  $\psi^*y = r \sin \theta \sin \phi$ ,  $\psi^*z = r \cos \theta$ . После вычислений мы получаем  $\psi^*(d\tau) = 2r^3 \sin \theta \times \cos \theta dr \wedge d\theta \wedge d\phi$ . Для нахождения значения  $d\tau$  на  $C$  интегрируем по прямоугольному параллелепипеду в параметрическом пространстве:

$$\int_C d\tau = \int_{r=0}^R \int_{\theta=0}^{\pi/2} \int_{\phi=0}^{2\pi} 2r^3 \sin \theta \cos \theta dr d\theta d\phi = \frac{1}{2}\pi R^4.$$

Давайте покажем кратко, как можно использовать *доказательство* теоремы Стокса, приведенное выше, чтобы дать альтернативное *определение* интегрирования  $k$ -формы по  $k$ -мерному подмногообразию. Это альтернативное определение будет совпадать с определением интегрирования по  $k$ -цепи, которое мы дали в случае, когда подмногообразие является образом  $k$ -цепи. Потом более детально покажем, как доказать, что интеграл по  $k$ -цепи не зависит от способа разложения на ячейки. Сделаем это так же, как уже сделали в конце предыдущего параграфа для сферы. При первом чтении материал с этого места и до конца параграфа можно опустить.

## Заплаты

Доказывая теорему Стокса, мы использовали набор функций  $\{\phi_1, \dots, \phi_n\}$ , причем каждая из этих функций непрерывно дифференцируема, неотрицательна и

$$\phi_1 + \dots + \phi_n \equiv 1 \quad \text{в рассматриваемой области (симплексе).}$$

Кроме того, каждая функция  $\phi_i$  за пределами шара  $D_i$  обращается в нуль. Шар  $D_i$  выбран так, чтобы выполнялись желаемые

свойства. (При доказательстве теоремы Стокса мы хотели, чтобы с помощью дифференцируемого отображения пересечение  $D_i$  с симплексом могло быть преобразовано в подмножество куба.)

А теперь вспомним определение подмногообразия  $M$  в пространстве  $\mathbb{R}^n$ . Подмножество  $M$  в пространстве  $\mathbb{R}^n$  называлось подмногообразием размерности  $k$ , если вокруг каждой точки  $p$  подмногообразия  $M$  мы можем найти шар  $D$  и его диффеоморфизм  $f$  на открытое подмножество  $U$  в пространстве  $\mathbb{R}^n$ , так что  $f(D \cap M) = U \cap \mathbb{R}^k$ , где  $\mathbb{R}^k$  идентифицируется с  $k$ -мерным подпространством пространства  $\mathbb{R}^n$ , заданным условием, что последние  $n - k$  координат равны нулю. Предположим, что мы можем покрыть подмногообразие  $M$  конечным числом таких шаров  $D_1, \dots, D_r$ . (Это всегда можно сделать в случае, если  $M$  — компактное, т. е. замкнутое и ограниченное подмножество в пространстве  $\mathbb{R}^n$ .) На каждом  $D_i$  определен свой диффеоморфизм  $f_i$  и

$$f_i(D_i \cap M) = U_i \cap \mathbb{R}^k.$$

Это подмножество обозначим  $W_i$ . Тогда

$$W_i = U_i \cap \mathbb{R}^k$$

будет открытым подмножеством пространства  $\mathbb{R}^k$ . Для любого перекрытия  $D_i \cap D_j$  мы можем рассмотреть ограничение либо  $f_i$ , либо  $f_j$ . Поэтому  $f_i(D_i \cap D_j)$  есть открытое подмножество  $U_i$  и  $f_j(D_i \cap D_j)$  — открытое подмножество  $U_j$ . Тогда

$$f_i \circ f_j^{-1} \text{ отображает } f_j(D_i \cap D_j) \text{ в } f_i(D_i \cap D_j)$$

и переводит

$$W_j \cap f_j(D_i \cap D_j) \text{ в } W_i \cap f_i(D_i \cap D_j).$$

Итак, наша ситуация выглядит следующим образом: мы покрыли  $M$  множествами  $O_i = D_i \cap M$ . Пусть  $g_i$  обозначает ограничение  $f_i$  на  $O_i$ . Тогда  $g_i$  будет взаимно-однозначным отображением  $O_i$  на открытое подмножество  $W_i$  в пространстве  $\mathbb{R}^k$ .<sup>12</sup>

<sup>12</sup>Описанный набор областей  $O_i \subset M$ , покрывающий многообразие  $M$ , вместе с диффеоморфизмами  $g_i : O_i \rightarrow W_i \subset \mathbb{R}^k$  принято называть *атласом карт* данного (гладкого) многообразия. — Прим. ред.

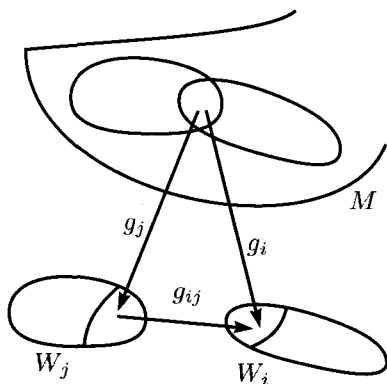


Рис. 15.20

Далее, отображение

$$g_{ij} = g_i \circ g_j^{-1} \quad \text{из} \quad W_j \cap g_j(O_i \cap O_j) \quad \text{в} \quad W_i \cap g_i(O_i \cap O_j)$$

дифференцируемо и имеет дифференцируемое обратное, определенное формулой  $g_{ji} = g_j \circ g_i^{-1}$ . См. рис. 15.20.

Мы можем вообразить, что многообразие  $M$  покрыто «заплатами». Отображения  $g_i^{-1}$  говорят, как  $W_i$  покрывают многообразие, а отображения  $g_{ij}$  говорят, как  $W_i$  и  $W_j$  сшиваются вместе. Каждое отображение  $g_{ij}$  может либо сохранять ориентацию, либо менять ее на обратную. Допустим, что в нашем случае все отображения  $g_{ij}$  сохраняют ориентацию<sup>13</sup>. (Это определяет «ориентацию» на  $M$ . Интуитивно ясно, что если  $M$  связно, то на нем существуют только две ориентации.) Пусть теперь набор функций  $\phi_1, \dots, \phi_r$  обладает свойствами:

$$0 \leq \phi_i \leq 1, \quad \phi_i = 0 \quad \text{вне} \quad D_i \quad \text{и} \quad \phi_1 + \dots + \phi_r \equiv 1 \quad \text{на} \quad M.$$

Теперь для каждой  $k$ -формы  $\omega$  мы можем написать

$$\omega = \phi_1 \omega + \dots + \phi_r \omega.$$

<sup>13</sup> Следовательно, многообразие  $M$  ориентируемо. Возможность выбора атласа, для которого все отображения склейки  $g_{ij}$  сохраняют ориентацию, и есть критерий ориентируемости многообразия. — Прим. ред.

Все слагаемые справа обращаются в нуль вне шара  $D_i$ . Тогда можно было бы определить

$$\int_M \omega = \sum \int_{W_i} (g_i^{-1})^* \phi_i \omega.$$

Другими словами, вместо разрезания  $M$  на отдельные куски, мы можем записывать любую форму в виде суммы небольших кусков, каждый из которых помещается внутрь одной-единственной координатной заплатки на  $M$ . Эту форму можно интегрировать посредством переноса на подмножество пространства  $\mathbb{R}^k$ . Повторное применение формулы замены переменных показывает, что это определение интеграла не зависит от выбора функций  $\phi_i$  или от заплаток (т. е. от выбора  $O_i$  и  $g_i$ ), но зависит от выбора ориентации  $M$ . Для детального изучения этого определения мы опять отсылаем читателя к книге Loomis and Sternberg, *Advanced Calculus*, главы 9–12.

## 15.5. Дифференциальные формы и когомологии\*

Для произвольного комплекса, расположенного в пространстве  $\mathbb{R}^n$ ,  $k$ -формы, которые мы определили, будут зародышами  $k$ -коцепей, а дифференциальный оператор  $d$  совместим с кограничным оператором  $\delta$ . Аналогично подпространствам, которые мы определили через образ и ядро кограничного оператора, сейчас мы можем построить подпространства (бесконечномерных) пространств  $\Omega^{(k)}(\mathbb{R}^n)$  с помощью дифференциального оператора  $d$ .

Если  $k$ -форма  $\omega$  удовлетворяет условию  $d\omega = 0$ , она называется *замкнутой*. Такая  $k$ -форма после интегрирования по ячейкам становится  $k$ -коцепью  $W$ , удовлетворяющей условию  $dW = 0$ , и поэтому она принадлежит пространству  $k$ -коциклов  $Z^k$ . Другими словами, *замкнутая  $k$ -форма* является зародышем  $k$ -коцикла.

Если  $k$ -форма  $\omega$  может быть записана в виде  $\omega = d\tau$ , то она называется *точной*. В этом случае после интегрирования  $\tau$  по

---

\*При первом чтении этот параграф можно опустить.

ячейкам комплекса получается  $(k-1)$ -коцепь  $T$ , а  $\omega$  порождает  $k$ -коцепь  $W$ , так что  $W = dT$ . Поэтому коцепь  $W$  лежит в подпространстве  $k$ -кограниц  $B^k$ . Другими словами, любая точная  $k$ -форма является зародышем кограницы.

Мы знаем, что для любого комплекса пространство  $B^k$  является подпространством  $Z^k$ . Соответствующее утверждение относительно  $k$ -форм состоит в том, что любая точная  $k$ -форма (зародыш элемента пространства  $B^k$ ) также замкнута (зародыш элемента пространства  $Z^k$ ). Доказательство очень простое: если  $\omega$  точная, то  $\omega = d\tau$  и поэтому  $d\omega = d(d\tau) = 0$ . Следовательно,  $\omega$  замкнута.

Для любого комплекса факторпространства (пространства когомологий)  $H^k = Z^k/B^k$  зависят лишь от основного пространства и не зависят способа его разрезания в комплекс. Поэтому разумно ожидать, что пространства  $H^k$  можно получить, рассматривая дифференциальные формы: от бесконечномерного пространства замкнутых форм ( $d\omega = 0$ ) мы переходим к факторпространству по пространству точных форм ( $\omega = d\tau$ ). Образовавшиеся пространства определяют так называемые *когомологии де Рама* основного пространства, на котором определены дифференциальные формы.

Для построения базиса пространств  $H^k$  вспомним, что для любого комплекса пространство  $H^k$  дуально гомологическому пространству  $H_k$ . Поэтому разумно ожидать, что базисный элемент для пространства  $H^k$  будет определяться его значениями на  $k$ -цепях, которые образуют базис  $H_k$  в любом удобном комплексе пространства.

Давайте рассмотрим несколько совсем простых примеров для иллюстрации этих абстрактных рассуждений.

**Пример 1А.** Пусть основное пространство — отрезок прямой (рис. 15.21). В этом пространстве построим 0-формы, которые будут дифференцируемыми функциями  $f(t)$ , и 1-формы, имеющие вид  $\omega = g(t) dt$ .

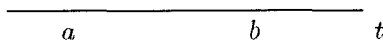


Рис. 15.21

В этом случае замкнутые 0-формы — это постоянные функции, определяющие одномерное подпространство. Точных 0-форм не бывает, поэтому факторпространство  $H^0$  замкнутых 0-форм по точным одномерно. Так отражается тот факт, что комплекс, полученный из отрезка, всегда *связен* и имеет  $\dim H_0 = 1$ .

При рассмотрении 1-форм мы обнаруживаем, что каждая 1-форма точная: для заданной  $\omega = g(t) dt$  мы образуем первообразную функцию  $G(t) = \int g(t) dt$  такую, что  $\omega = dG$ . Следовательно, в этом случае пространство  $H^1$  имеет нулевую размерность.

**Пример 1В.** Пусть основное пространство состоит из двух непересекающихся отрезков прямой (рис. 15.22). В этом случае пространство  $H^0 = Z^0$  двумерное: оно состоит из функций, которые имеют одно постоянное значение на отрезке  $[a, b]$  и другое, может быть отличное от первого, постоянное значение на отрезке  $[c, d]$ . Это отражает тот факт, что комплекс, построенный на этом пространстве, будет иметь две связанных компоненты.

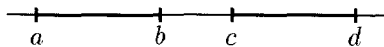


Рис. 15.22

В общем случае при любой размерности основного пространства замкнутые 0-формы будут функциями, принимающими постоянное значение на каждой связной компоненте. Они будут образовывать пространство  $H^0 = Z^0$  с размерностью, равной числу связных компонент в основном пространстве. С этого момента и дальше мы будем рассматривать только связные пространства, и нам не нужно будет возвращаться к пространству  $H^0$ .

**Пример 2А.** Основное пространство — прямоугольник на плоскости (рис. 15.23). Это пространство связно, поэтому  $H^0$  одномерно. Изучая 1-формы заметим, что любая 1-форма  $\omega$  точная. Мы доказали это, рассматривая криволинейные интегралы. Поэтому пространство  $H^1$  имеет нулевую размерность.

Рассмотрев теперь 2-формы, замечаем, что каждая 2-форма точная. Общий вид 2-формы — это  $\tau = f(x, y) dx \wedge dy$ . образуем

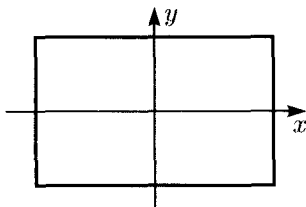


Рис. 15.23

функцию

$$F(x, y) = \int_0^x f(t, y) dt \quad \text{такую, что} \quad \frac{\partial F}{\partial x} = f(x, y).$$

Тогда

$$d(Fdy) = f(x, y) dx \wedge dy = \tau.$$

Отсюда следует, что  $H^2$  имеет нулевую размерность.

**Пример 2В.** Пусть опять основное пространство будет прямоугольником на плоскости, но с одной выкинутой точкой — началом координат (рис. 15.24). И на этот раз пространство  $H^0$  одномерно, потому что основное пространство связно. А пространство  $H^2$  имеет нулевую размерность, потому что каждая 2-форма точная. Однако, если мы рассмотрим 1-формы, то появится новое явление: теперь существуют замкнутые 1-формы, которые не будут точными. Так что в этом случае пространство  $H^1$  нетривиально.

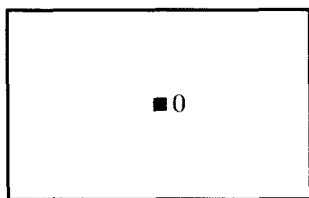


Рис. 15.24

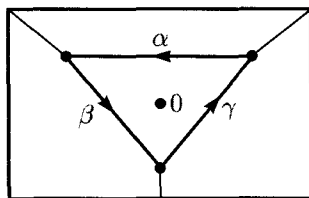


Рис. 15.25

Чтобы получить базис пространства  $H^1$ , сначала рассмотрим пространство гомологий  $H_1$  для 2-комплекса, построенного на этом пространстве (см. рис. 15.25). Поскольку начало координат не принадлежит пространству, то в комплексе не может быть

ячейки, включающей начало. В результате цикл типа  $\alpha + \beta + \gamma$ , окружающий начало координат, не может быть границей. Класс эквивалентности такого цикла (по модулю пространства границ  $B_1$ ) образует базис одномерного пространства  $H_1$ . Для того же комплекса пространство когомологий  $H^1$  дуально  $H_1$ . Его базисный элемент будет классом эквивалентности 1-коцикла  $W$ , который не является *кограницей*. Поскольку  $H^1$  дуально  $H_1$ , мы ожидаем, что  $W$  принимает фиксированное ненулевое значение для любого цикла типа  $\alpha + \beta + \gamma$ , обходящего начало один раз против часовой стрелки. Чтобы получить базис для пространства когомологий де Рама  $H^1$ , мы должны найти зародыш коцикла для  $W$ : замкнутую 1-форму  $\omega_0$  с фиксированным ненулевым значением для любой кривой, окружающей начало один раз. Такую 1-форму можно получить, рассмотрев функцию  $\theta = \arctan(y/x)$ , которая определена везде, кроме начала координат, и которая может быть сделана непрерывной везде, кроме одной прямой, выходящей из начала координат (обычно это ось  $x$  в отрицательном направлении, так что  $-\pi < \theta \leq \pi$ ). Тогда напишем дифференциал

$$\omega_0 = d\theta = \frac{x dy - y dx}{x^2 + y^2}.$$

Эта 1-форма определена и непрерывна везде, кроме начала координат. Кроме того, она замкнута:  $d\omega_0 = 0$ , что можно проверить прямым вычислением. С другой стороны, она не точная: нет такой непрерывной функции  $f(x, y)$ , чтобы  $\omega_0 = df$ . Для любой кривой  $\alpha$ , обходящей начало координат  $n$  раз против часовой стрелки, мы имеем

$$\int_{\alpha} \omega_0 = 2\pi n.$$

Поэтому базисом пространства  $H^1$  является класс эквивалентности

$$\overline{\omega_0} = \frac{x dy - y dx}{x^2 + y^2} + dg,$$

где  $g(x, y)$  — любая дифференцируемая функция.

**Пример 3А.** Пусть основное пространство — прямоугольный параллелепипед в пространстве  $\mathbb{R}^3$  (рис. 15.26). Без доказательства

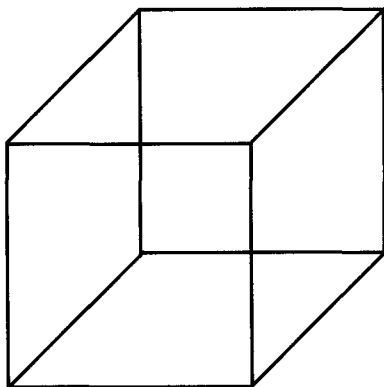


Рис. 15.26

утверждаем, что

$H^0$  одномерно (связная область);

$H^1 = \{0\}$  (каждая замкнутая 1-форма точна);

$H^2 = \{0\}$  (каждая замкнутая 2-форма точна);

$H^3 = \{0\}$  (каждая 3-форма точна).

Эти результаты для пространств  $H^1$ ,  $H^2$ ,  $H^3$  являются следствиями общей теоремы, известной как *лемма Пуанкаре*. Грубо говоря, лемма утверждает, что на любом пространстве, которое можно непрерывно стянуть в точку, все замкнутые дифференциальные формы точные. Интересные случаи возникают, если пространство имеет «дырки» и поэтому его нельзя стянуть в точку. Конечно, в этих случаях существуют циклы, не являющиеся границами, и поэтому дуальные пространства  $H_1, H_2, \dots$  ненулевые.

**Пример 3В.** Пусть основное пространство опять будет прямоугольным параллелепипедом в пространстве  $\mathbb{R}^3$ , но на этот раз выкинуто начало координат. Результаты для  $H^0$ ,  $H^1$  и  $H^3$  не меняются. Однако теперь пространство  $H^2$  одномерно. Дифференциальная форма

$$\tau_0 = \frac{x \, dy \wedge dz + y \, dz \wedge dx + z \, dx \wedge dy}{(x^2 + y^2 + z^2)^{3/2}}$$

определена везде, кроме начала координат. Она замкнута, что можно проверить, проделав громоздкие вычисления. Однако,  $\int_S \tau_0 = 4\pi$  для любой замкнутой поверхности  $S$ , окружающей начало координат. Конечно, такая поверхность является циклом, но не границей. Не удивительно, что она принадлежит классу эквивалентности, который дуален  $H^2$ .

**Пример 3С.** Пусть основное пространство будет прямоугольным параллелепипедом в пространстве  $\mathbb{R}^3$ , из которого выкинута ось  $z$ . В этом случае  $H^1$  одномерно, и его базис

$$\overline{\omega}_0 = \frac{x dy - y dx}{x^2 + y^2} + dg.$$

При этом пространство  $H^2$  имеет нулевую размерность.

**Пример 3D.** Пусть основное пространство — поверхность тора, расположенного в пространстве  $\mathbb{R}^3$  (см. рис. 15.27). В этом случае  $H^1$  двумерно. Один из его базисных элементов имеет вид

$$\overline{\omega}_0 = \frac{x dy - y dx}{x^2 + y^2} + dg,$$

и определяет, сколько раз замкнутая кривая обходит ось  $z$ . Другой базисный элемент

$$\overline{\omega}_1 = \frac{(R-1) dz - z dR}{z^2 + (R-1)^2} + dg, \quad R = \sqrt{(x^2 + y^2)}.$$

Значение этой формы на любой замкнутой кривой определяет сколько раз кривая пройдет по окружности единичного радиуса в плоскости  $xy$ . Таким образом, базисные элементы  $\overline{\omega}_0$  и  $\overline{\omega}_1$  дуальны двум типам циклов для любого комплекса на торе.

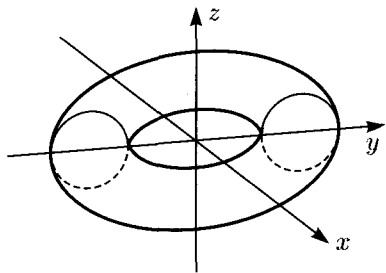


Рис. 15.27

А сейчас мы докажем лемму Пуанкаре, чтобы обосновать утверждения, сделанные в этих примерах. Мы получим точный рецепт, как строить  $(k-1)$ -форму  $\phi$ , обладающую свойством  $d\phi = \omega$ , где  $\omega$  — заданная замкнутая  $k$ -форма. Лемма будет доказана для случая, когда  $\omega$  определена на открытом множестве в пространстве  $\mathbb{R}^n$ , обладающим тем свойством, что каждая точка множества может быть соединена с началом координат отрезком прямой, целиком принадлежащим множеству. Такое множество называется *звездным*. Позднее мы обобщим доказательство на случай множества, которое с помощью гладкого отображения может быть взаимно-однозначно преобразовано к звездному. На рис. 15.28 показаны примеры таких множеств.

Для данной звездной области  $Q$  с координатами  $x^1, \dots, x^n$  мы можем построить отрезок прямой, соединяющий начало координат с произвольной точкой  $p$ . Поскольку  $Q$  звездна, то этот отрезок целиком лежит в области  $Q$ . Определим функцию  $\beta$ , которая для каждой точки  $p$  отображает  $[0, 1]$  в этот отрезок. Таким образом, если точка  $p$  имеет координаты  $x^1, \dots, x^n$ , то прямолинейный отрезок задается параметрически формулой

$$\begin{pmatrix} u^1 \\ \vdots \\ u^n \end{pmatrix} = \begin{pmatrix} tx^1 \\ \vdots \\ tx^n \end{pmatrix}, \quad 0 \leq t \leq 1.$$

Если воспользоваться терминологией переноса, то мы имеем  $\beta^*(u^1) = tx^1$  и т. д.

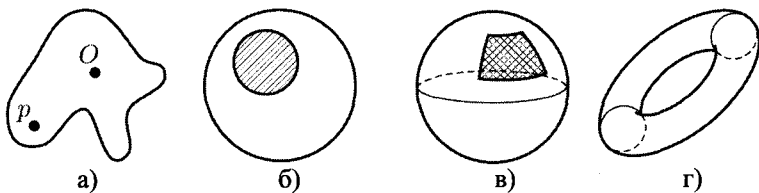


Рис. 15.28. а) Звездность относительно точки  $O$  (но не точки  $p$ ). б) Внутри области дырка, поэтому нет звездности относительно какой-либо точки. в) Область на поверхности сферы не звездна, но это образ звездной области. г) Поверхность всего тора не является образом никакой звездной области

Любая функция  $g$ , определенная в области  $Q$ , с помощью переноса может быть преобразована к области  $Q \times [0, 1]$ . Например, функция  $g(u^1, \dots, u^n)$  в результате переноса преобразуется в функцию  $\beta^*g = g(tx^1, \dots, tx^n)$ . По существу, это значит, что  $\beta^*g$  является функцией  $p = (x^1, \dots, x^n)$  и  $t$ ; при изменении  $t$  от 0 до 1 она пробегает значения  $g$  вдоль отрезка, соединяющего  $p$  с началом координат. Конечно, если  $p$  меняется, то значения  $g$  берутся вдоль другого прямолинейного отрезка. Внимательно посмотрите на рис. 15.29, чтобы осознать смысл  $\beta^*g$ .

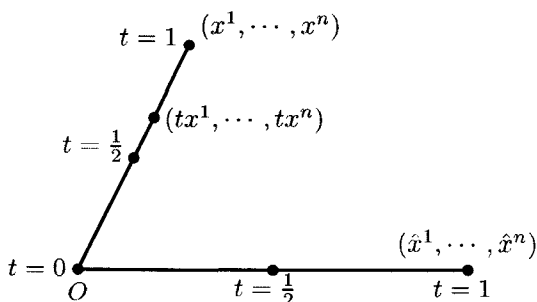


Рис. 15.29

Согласно обычным правилам, мы можем проделать перенос с любой базисной 1-формой любой  $k$ -формы:

$$\beta^* du^i = d(tx^i) = x^i dt + t dx^i.$$

Произвольная  $k$ -форма  $\omega$ , определенная на  $Q$ , может быть записана в координатах  $u^1, \dots, u^n$ :

$$\omega = \sum_{i_1 < \dots < i_k} a_{i_1, \dots, i_k}(u^1, \dots, u^n) du^{i_1} \wedge \dots \wedge du^{i_k}.$$

Проделав перенос с такой формой, получим

$$\begin{aligned} \beta^* \omega = \sum_{i_1 < \dots < i_k} a_{i_1, \dots, i_k}(tx^1, \dots, tx^n) \\ \times (t dx^{i_1} + x^{i_1} dt) \wedge \dots \wedge (t dx^{i_k} + x^{i_k} dt). \end{aligned}$$

Эта  $k$ -форма  $\beta^*\omega$  является суммой членов вида

$$\tau_1(t) = A(t, x^1, \dots, x^n) dx^{i_1} \wedge \dots \wedge dx^{i_k}$$

или

$$\tau_2(t) = B(t, x^1, \dots, x^n) dt \wedge dx^{i_1} \wedge \dots \wedge dx^{i_{k-1}}.$$

Определим линейный оператор  $L$ , действующий на эти формы следующим образом:

$$\begin{aligned} L\tau_1 &= 0, \\ L\tau_2 &= \left( \int_0^1 B(t, x^1, \dots, x^n) dt \right) dx^{i_1} \wedge \dots \wedge dx^{i_{k-1}}. \end{aligned}$$

Этот оператор обладает замечательными свойствами:

$$\begin{aligned} dL\tau_1 + Ld\tau_1 &= \tau_1(1) - \tau_1(0), \\ dL\tau_2 + Ld\tau_2 &= 0. \end{aligned}$$

Они доказываются прямым вычислением.

Начнем с  $\tau_1$ . Конечно,  $dL\tau_1 = 0$ . Кроме того,

$$d\tau_1 = \frac{\partial A}{\partial t} dt \wedge dx^{i_1} \wedge \dots \wedge dx^{i_k} + \text{члены без } dt.$$

Поэтому

$$\begin{aligned} Ld\tau_1 &= \int_0^1 \left( \frac{\partial A}{\partial t} dx^{i_1} \wedge \dots \wedge dx^{i_k} \right) dt \\ &= (A(1, x^1, \dots, x^n) - A(0, x^1, \dots, x^n)) dx^{i_1} \wedge \dots \wedge dx^{i_k}. \end{aligned}$$

Для  $\tau_2$  мы имеем выражение

$$L\tau_2 = \left( \int_0^1 B(t, x^1, \dots, x^n) dt \right) dx^{i_1} \wedge \dots \wedge dx^{i_{k-1}},$$

которое не зависит от  $t$ . Вычисляя  $dL\tau_2$ , мы дифференцируем под знаком интеграла и получаем

$$dL\tau_2 = \sum_{j=1}^n \left( \int_0^1 \frac{\partial B}{\partial x^j} dt \right) dx^j \wedge dx^{i_1} \wedge \dots \wedge dx^{i_{k-1}}.$$

С другой стороны,

$$d\tau_2 = \sum_{j=1}^n \frac{\partial B}{\partial x^j} dx^j \wedge dt \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}}$$

или

$$d\tau_2 = - \sum_{j=1}^n \frac{\partial B}{\partial x^j} dt \wedge dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}},$$

так что

$$Ld\tau_2 = - \sum_{j=1}^n \left( \int_0^1 \frac{\partial B}{\partial x^j} dt \right) dx^j \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_{k-1}},$$

откуда следует, что  $dL\tau_2 + Ld\tau_2 = 0$ .

Итак, если мы рассмотрим  $dL(\beta^*\omega) + Ld(\beta^*\omega)$ , то можно игнорировать все члены вида  $\tau_2$ , содержащие множитель  $dt$ . Для всех  $k > 0$  мы имеем  $\tau_1(0) = 0$ , поскольку множитель  $t$  стоит при каждом  $dx^i$ . Положив  $t = 1$ , мы получаем

$$dL(\beta^*\omega) + Ld\beta^*\omega = \sum_{i_1 < \cdots < i_k} a_{i_1 \dots i_k}(x^1, \dots, x^n) dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \omega.$$

Однако нам известно, что  $d\beta^*\omega = \beta^*d\omega$ , поэтому

$$dL\beta^*\omega + L\beta^*d\omega = \omega.$$

Таким образом, оператор  $S = L\beta^*$  удовлетворяет тождеству

$$dS\omega + Sd\omega = \omega.$$

Если  $\omega$  замкнута, т. е.  $d\omega = 0$ , то

$$dS\omega = \omega.$$

Итак, мы доказали, что  $\omega$  точная.

До сих пор мы предполагали, что  $\omega$  определена в звездной области  $Q$ . Сейчас давайте распространим наше доказательство на случай, когда  $\omega$  определена в области  $\Psi(Q)$ , которая является образом звездной области при гладком взаимно-однозначном

отображении  $\Psi$ . Если  $\omega$  замкнута, то  $\Psi^*\omega$  тоже замкнута, потому что  $d\Psi^*\omega = \Psi^*d\omega = 0$ . Поэтому можно написать  $\Psi^*\omega = d(S\Psi^*\omega)$ . Поскольку  $\Psi$  обратима, то можно применить обратный перенос  $(\Psi^*)^{-1}$  и получить

$$\omega = (\Psi^*)d(S\Psi^*\omega) \quad \text{или} \quad \omega = (\Psi^*)^{-1}S\Psi^*\omega.$$

Итак, для формы  $\omega$ , определенной в области  $D$ , мы имеем семейство неопределенных интегралов  $(\Psi^*)^{-1}S\Psi^*\omega$ , соответствующих различным способам выражения  $D$  в качестве образа звездной области:  $D = \Psi(Q)$ . Есть и другой взгляд на эту ситуацию. В области  $D$  мы можем ввести различные системы координат, так чтобы область в *координатном пространстве*, отображаемая в  $D$ , была звездной. Тогда оператор  $(\Psi^*)^{-1}S\Psi^*$  соответствует интегрированию вдоль прямых линий в координатном пространстве. Если, например,  $D$  — поверхность Мексиканского залива, то мы получаем  $S$ , интегрируя вдоль прямых линий, соединяющих точки области  $D$  с центром Земли (или с другим началом координат). Введя сферические координаты, мы построим другой неопределенный интеграл  $(\Psi^*)^{-1}S\Psi^*$ , когда каждая точка залива соединяется с началом координат линией, представляемой прямой на карте с проекцией Меркатора.

Давайте сформулируем процедуру получения неопределенного интеграла от дифференциальной формы

$$\omega = \sum_{i_1 < \dots < i_k} a_{i_1 \dots i_k}(x^1, \dots, x^n) dx^{i_1} \wedge \dots \wedge dx^{i_k}.$$

*Шаг 1.* Построим перенос  $\beta^*\omega$ , сделав замену  $x^i \rightarrow tx^i$  в аргументах всех функций так, чтобы

$$a(x^1, \dots, x^n) \rightarrow a(tx^1, \dots, tx^n),$$

и заменив

$$dx^i \rightarrow x^i dt + t dx^i.$$

*Шаг 2.* Выбросим все члены, в которые не входит  $dt$ . Во всех остающихся членах  $dt$  сдвигаем влево, внимательно следя за знаками.

*Шаг 3.* Обращаемся с  $dt$  как в обычном интеграле и интегрируем по  $t$  от 0 до 1. Эта процедура в явном виде показана на следующих примерах.

**Пример 1.**  $\omega = y^2 dx + 2xy dy$  (замкнутая 1-форма).

*Шаг 1.*  $\beta^*\omega = t^2 y^2 (x dt + t dx) + 2t^2 xy (y dt + t dy)$ .

*Шаг 2.*  $\beta^*\omega = (t^2 xy^2 + 2t^2 xy^2) dt + \text{другие члены}$ .

*Шаг 3.*  $S\omega = xy^2 \int_0^1 3t^2 dt = xy^2$ .

*Проверка.*  $d(xy^2) = y^2 dx + 2xy dy$ .

**Пример 2.**  $\omega = \sin x dx \wedge dy$ .

*Шаг 1.*  $\beta^*\omega = \sin(tx)(x dt + t dx) \wedge (y dt + t dy)$ .

*Шаг 2.*  $\beta^*\omega = xt \sin tx dt \wedge dy - yt \sin tx dt \wedge dx + \text{члены, не содержащие } dt$ .

*Шаг 3.*

$$\begin{aligned} S\omega &= (x dy - y dx) \int_0^1 t \sin tx dt \\ &= (x dy - y dx) \left( \frac{\sin x - x \cos x}{x^2} \right) \\ &= \left( \frac{\sin x - x \cos x}{x} \right) dy - y \left( \frac{\sin x - x \cos x}{x^2} \right) dx \end{aligned}$$

или

$$S\omega = -\cos x dy + y \left( \frac{x \cos x - \sin x}{x^2} \right) dx + \frac{\sin x}{x} dy.$$

Ответ отличается от «очевидной» первообразной  $-\cos x dy$  на точную 1-форму

$$d\left(\frac{y}{x} \sin x\right) = \frac{x \cos x - \sin x}{x^2} y dx + \frac{\sin x}{x} dy.$$

**Пример 3.**  $\beta^*\omega = (y^2 - x^2)z dx \wedge dy + (x^2 - z^2)y dz \wedge dx + (z^2 - y^2)x dy \wedge dz$ .

*Шаг 1.*

$$\begin{aligned} \beta^*\omega &= t^3(y^2 - x^2)z(x dt + t dx) \wedge (y dt + t dy) \\ &\quad + t^3(x^2 - z^2)y(z dt + t dz) \wedge (x dt + t dx) \\ &\quad + t^3(z^2 - y^2)x(y dt + t dy) \wedge (z dt + t dz). \end{aligned}$$

Шаг 2.

$$\begin{aligned}\beta^*\omega &= t^4(-(y^2 - x^2)zy + (x^2 - z^2)yz) dt \wedge dx \\ &\quad + ((y^2 - x^2)zx - (z^2 - y^2)xz) dt \wedge dy \\ &\quad + (-(x^2 - z^2)yx + (z^2 - y^2)xy) dt \wedge dz \\ &\quad + \text{члены, не содержащие } dt.\end{aligned}$$

Шаг 3.

$$\begin{aligned}S\omega &= \int_0^1 t^4 dt (2x^2yz - y^3z - yz^3) dx \\ &\quad + (2y^2zx - x^3z - xz^3) dy + (2z^2xy - xy^3 - x^3y) dz\end{aligned}$$

или

$$S\omega = \frac{2}{5}(x^2yz dx + y^2zx dy + z^2xy dz).$$

Мы только что показали, что ядро  $S : \Omega^k \rightarrow \Omega^{k-1}$  является подпространством образа  $S : \Omega^{k+1} \rightarrow \Omega^k$ . На самом деле это весь образ. Предположим, например, что  $k$ -форма удовлетворяет условию  $S\omega = 0$ . Тогда, поскольку  $\omega = Sd\omega + dS\omega$ , мы знаем, что  $\omega = Sd\omega$ , т. е. что  $\omega$  является образом  $S$ . Итак,  $S\omega = 0$  тогда и только тогда, когда  $\omega = S\phi$  для некоторого  $\phi$ .

## Резюме

### А. Внешняя алгебра и дифференциальное исчисление

Вы должны уметь определять пространства  $\Lambda^k(V^*)$  для произвольного  $n$ -мерного пространства  $V$  и записывать базис для каждого из них.

Необходимо знать и уметь применять свойства оператора  $d$  для дифференциальных форм произвольной степени, включая процедуру переноса.

### В. Интегрирование дифференциальных форм

Вы должны уметь вычислять интеграл от  $k$ -формы по  $k$ -ячейке, представленной в виде образа единичного  $k$ -куба.

Вы должны знать, уметь доказывать и применять теорему Стокса.

### С. Дифференциальные формы и когомологии

Вы должны знать, как строить коцепи посредством интегрирования дифференциальных форм; должны уметь находить формы, определяющие базисные элементы пространства  $H^k$  для 2-комплекса и 3-комплекса.

Для данной дифференциальной  $k$ -формы  $\phi$ , удовлетворяющей условию  $d\phi = 0$  и определенной в звездной области, надо уметь строить  $(k-1)$ -формы  $S\phi$ , обладающие свойством  $dS\phi = \phi$ .

### Задачи

- 15.1. (а) Пусть  $\tau(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$  — кососимметричная трилинейная функция (т.е. 3-форма). Не обращаясь к свойствам определителей, докажите, что если  $\mathbf{v}_1, \mathbf{v}_2$  и  $\mathbf{v}_3$  линейно зависимы, то  $\tau(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3) = 0$ .
- (б) Пусть  $\omega^1, \omega^2, \omega^3$  — три элемента пространства  $V^*$ . Не используя свойства определителей, докажите, что если  $\omega^1, \omega^2, \omega^3$  линейно зависимы, то  $\omega^1 \wedge \omega^2 \wedge \omega^3 = 0$ .
- 15.2. Пусть в четырехмерном пространстве-времени дана 2-форма  $E = dt \wedge (E_x dx + E_y dy + E_z dz)$ . Пусть  $B = B_z dx \wedge dy - B_y dx \wedge dz + B_x dy \wedge dz$ . Вычислите  $E \wedge B$ .
- 15.3. Предположим, что детерминант линейного преобразования  $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$  имеет вид

$$\text{Det } A = dx^1 \wedge dx^2 \wedge \cdots \wedge dx^n [Ae_1, Ae_2, \dots, Ae_n],$$

где  $\{dx^1, dx^2, \dots, dx^n\}$  дуальны  $\{e_1, e_2, \dots, e_n\}$ .

- (а) Докажите, что если  $A$  — единичная матрица, то  $\text{Det } A = 1$ .
- (б) Докажите, что если  $A^*$  обозначает матрицу, сопряженную к  $A$ , то

$$\text{Det } A = A^* dx^1 \wedge A^* dx^2 \wedge \cdots \wedge A^* dx^n (e_1, e_2, \dots, e_n).$$

- 15.4. Вычислите следующие определители, используя результаты предыдущего упражнения:

(a)

$$\text{Det} \begin{pmatrix} 2 & -1 & 3 \\ -1 & 3 & 2 \\ 3 & 2 & -1 \end{pmatrix} = dx \wedge dy \wedge dz \left[ \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \right].$$

(b)

$$\text{Det} \begin{pmatrix} a & -b & 0 & 0 \\ b & a & -b & 0 \\ 0 & b & a & -b \\ 0 & 0 & b & a \end{pmatrix}.$$

15.5. Пусть  $\Lambda^3(V)$  обозначает пространство *всех кососимметричных трилинейных* функций из  $V \times V \times V$  в  $\mathbb{R}$ . Допустим, что  $V$  — четырехмерное пространство и пусть  $\varepsilon^1, \varepsilon^2, \varepsilon^3, \varepsilon^4$  образуют базис дуального пространства  $V^*$ . Используя обозначения  $\wedge$ , напишите базис  $\Lambda^3(V)$  и напишите формулу, как базисный элемент действует на тройку векторов  $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ .

15.6. Пусть  $\omega = f dx + dy$  и  $\tau = g dx + dz$ , где  $f$  и  $g$  — дифференцируемые функции  $x, y$  и  $z$ . Вычислите  $d(\omega \wedge \tau)$ , записав ответ как множитель при  $dx \wedge dy \wedge dz$ .

15.7. Пусть  $\beta$  обозначает дифференцируемое отображение из аффинного пространства  $A$  (размерности  $m$ ) в пространство  $B$  (размерности  $n$ ). Пусть  $\tau(q, \mathbf{w})$  — 1-форма в пространстве  $B$ , где  $q$  является точкой в этом пространстве,  $\mathbf{w}$  — вектор, направленный из этой точки. Тогда  $d\tau$  можно *определить* как

$$d\tau(q, \mathbf{w}_1, \mathbf{w}_2) = \text{наилучшее линейное приближение к} \\ \tau(q + \mathbf{w}_1, \mathbf{w}_2) - \tau(q, \mathbf{w}_2) - \tau(q + \mathbf{w}_2, \mathbf{w}_1) + \tau(q, \mathbf{w}_1).$$

Перенос 1-формы *определяется* следующим образом:

$$(\beta^* \tau)(q, \mathbf{v}) = \tau(\beta q, d\beta(\mathbf{v})),$$

а для 2-формы  $\sigma$ :

$$(\beta^* \sigma)(q, \mathbf{v}_1, \mathbf{v}_2) = \sigma(\beta q, d\beta(\mathbf{v}_1), d\beta(\mathbf{v}_2)).$$

(a) Воспользовавшись этими определениями (и при необходимости цепным правилом), докажите, что

$$d\beta^* \tau = \beta^*(d\tau).$$

- (b) В пространстве  $A$  введите аффинные координаты  $u^1, \dots, u^m$ , а в пространстве  $B$  — координаты  $x^1, \dots, x^n$ . Тогда  $\beta$  можно задать с помощью дифференцируемых функций переноса

$$\begin{aligned}\beta^* x^1 &= F^1(u^1, \dots, u^m), \\ \beta^* x^2 &= F^2(u^1, \dots, u^m), \\ &\vdots \\ \beta^* x^n &= F^n(u^1, \dots, u^m),\end{aligned}$$

а

$$\tau = \sum_{j=1}^n G^j(x^1, \dots, x^n) dx^j,$$

где  $G^j$  — дифференцируемые функции. Вычислите явно  $d\beta^*\tau$  и  $\beta^*(d\tau)$  и покажите, что они равны.

- (с) Пусть  $\tau$  и  $\lambda$  — 1-формы. Покажите, что если  $\Omega = \tau \wedge \lambda$ , то  $\beta^*\Omega = (\beta^*\tau) \wedge (\beta^*\lambda)$ .

15.8. Сферические координаты  $r, \theta, \phi$  связаны с декартовыми координатами  $x, y$  и  $z$  формулами

$$\alpha^* x = r \sin \theta \cos \phi, \quad \alpha^* y = r \sin \theta \sin \phi, \quad \alpha^* z = r \cos \theta,$$

где  $\alpha$  — отображение, переводящее точку  $\begin{pmatrix} r \\ \theta \\ \phi \end{pmatrix}$  в пространстве

сферических координат в точку  $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$  с координатами  $r, \theta$  и  $\phi$ .

Вычислите переносы следующих дифференциальных форм (т.е. выразите их в сферических координатах):

- (a)  $\alpha^*(x dy - y dx)$ ;  
 (b)  $\alpha^* \frac{x dx + y dy + z dz}{(x^2 + y^2 + z^2)^3}$ ;  
 (c)  $\alpha^* \frac{x dy \wedge dz + y dz \wedge dx + z dx \wedge dy}{(x^2 + y^2 + z^2)^{3/2}}$ ;  
 (d)  $\alpha^* dx \wedge dy \wedge dz$ .

- 15.9. (a) Пусть  $f$  — дифференцируемая функция в пространстве  $\mathbb{R}^3$ . Пусть  $A$  обозначает 2-форму

$$A = \frac{\partial f}{\partial x} dy \wedge dz + \frac{\partial f}{\partial y} dz \wedge dx + \frac{\partial f}{\partial z} dx \wedge dy.$$

Покажите, что  $dA = \Delta f dx \wedge dy \wedge dz$ , где  $\Delta$  обозначает лапласиан  $\partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ .

- (b) Пусть  $F_x, F_y, F_z$  — дифференцируемые функции в пространстве  $\mathbb{R}^3$ , удовлетворяющие условию

$$\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} = 0.$$

Определим

$$B = \left( \frac{\partial F_z}{\partial y} - \frac{\partial F_y}{\partial z} \right) dx + \left( \frac{\partial F_x}{\partial z} - \frac{\partial F_z}{\partial x} \right) dy + \left( \frac{\partial F_y}{\partial x} + \frac{\partial F_x}{\partial y} \right) dz.$$

Покажите, что  $dB = \Delta F_x dy \wedge dz + \Delta F_y dz \wedge dx + \Delta F_z dx \wedge dy$ .

- 15.10. Тело вращения вокруг оси  $z$  ограничено: снизу — диском  $x^2 + y^2 \leq 1$ ,  $z = 0$ ; сбоку — поверхностью цилиндра  $x^2 + y^2 = 1$ ,  $0 < z < 2$ ; сверху — параболоидом  $z = 1 + x^2 + y^2$ , задаваемым параметрически в виде

$$\beta^* x = r \cos \theta, \quad \beta^* y = r \sin \theta, \quad \beta^* z = 1 + r^2.$$

Если смотреть снаружи, то все поверхности ориентированы против часовой стрелки, как это показано на рис. 15.30.

- (a) Вычислите интеграл от 2-формы  $\tau = x dy \wedge dz - y dx \wedge dz$  по поверхности верхнего параболоида.  
 (b) Вычислите интеграл от  $\tau$  по поверхности цилиндра.  
 (c) С помощью теоремы Стокса выразите *объем* этого твердого тела через только что полученные интегралы.

- 15.11. Пусть  $\Omega = x dy \wedge dz$ .

- (a) Вычислите  $\int \Omega$  по диску  $x^2 + y^2 \leq R^2$ ,  $z = 0$ .  
 (b) Рассмотрите полусферу  $x^2 + y^2 + z^2 = R^2$ ,  $z \geq 0$ , которую параметризуем сферическими координатами (см. рис. 15.31):

$$z = R \cos \theta, \quad x = R \sin \theta \cos \phi, \quad y = R \sin \theta \sin \phi.$$

Выразите  $\Omega$  через  $R$ ,  $\theta$  и  $\phi$ . Потом вычислите  $\int \Omega$  по этой полусфере.

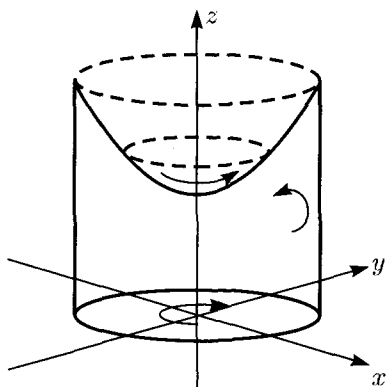


Рис. 15.30

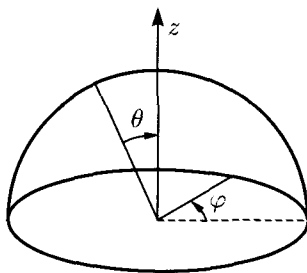


Рис. 15.31

- (с) Воспользуйтесь теоремой Стокса и запишите интеграл от 3-формы, который равняется разности двух поверхностных интегралов, вычисленных в пунктах (b) и (a). Выясните геометрический смысл этого интеграла и вычислите его.
- 15.12. (a) Рассмотрите твердое тело с однородной плотностью  $\rho$ , которое занимает область  $Q$  в пространстве  $\mathbb{R}^3$ . Покажите, что интеграл от 2-формы

$$\tau = \frac{1}{3}(x^3 dy \wedge dz - y^3 dx \wedge dz)$$

по границе  $Q$  дает момент инерции  $I_{zz}$  этого твердого тела.

- (b) Придумайте 2-форму  $\sigma$ , интеграл от которой по  $\partial Q$  дает момент инерции  $I_{xy}$ . Почему ответ в этой задаче не единственный?
- (с) Используйте результат пункта (a) для вычисления момента инерции сферы радиуса  $a$  относительно ее диаметра. Воспользуйтесь сферическими координатами.
- 15.13. (a) Пусть  $W$  обозначает область в пространстве  $\mathbb{R}^3$ , занятую твердым телом с однородной плотностью  $\rho$ . Покажите, что координата  $z$  центра масс этого тела  $\bar{z}$  может быть определена как частное от двух поверхностных интегралов, вычисленных по границе тела  $\partial W$ , т. е.

$$\bar{z} = \int_{\partial W} \frac{1}{2} \rho r^2 dx \wedge dy \bigg/ \int_{\partial W} z dx \wedge dy,$$

где  $r^2 = x^2 + y^2 + z^2$ .

- (b) Переход сферическим координатам дается формулами:

$$\alpha^*x = r \sin \theta \cos \phi, \quad \alpha^*y = r \sin \theta \sin \phi, \quad \alpha^*z = r \cos \theta.$$

Вычислите  $\alpha^*(dx \wedge dy)$ .

- (c) Используя результат пункта (b), вычислите

$$\int_{\partial W} \frac{1}{2} r^2 dx \wedge dy$$

для случая, когда  $W$  является внутренностью полусферы радиуса  $a$  с правосторонней ориентацией.

- 15.14. Пусть  $C$  обозначает половину цилиндра, ограниченного плоскостями  $z = 0$  «низ»,  $z = 1$  «верх», плоскостью  $x = 0$  «плоскость» и боковой поверхностью  $x^2 + y^2 = a^2$  «кривая».  $C$  имеет правостороннюю ориентацию, так что вектора  $e_x, e_y, e_z$  образуют правильно упорядоченный базис, как это показано на рис. 15.32. Для каждой грани  $C$  имеется упорядоченный базис, который определяет ее ориентацию.

- (a) Напишите выражение для  $\partial C$ , следя за знаками.
- (b) Имеется 2-форма  $\tau = x^2 dy \wedge dz - 2xz dx \wedge dy$ . Вычислите перенос  $\tau$  при отображении, заданном следующим образом:  $\beta^*x = a \cos u$ ,  $\beta^*y = a \sin u$ ,  $\beta^*z = v$ . Вычислите интеграл от  $\tau$  по боковой поверхности «кривая».
- (c) Пусть задано отображение:  $\alpha^*x = r \cos \theta$ ,  $\alpha^*y = r \sin \theta$ ,  $\alpha^*z = 1$ . Вычислите интеграл от  $\tau$  по поверхности «верх».
- (d) Чему равен интеграл от  $\tau$  по другим двум поверхностям? Объясните.
- (e) Вычислите  $d\tau$ . Используя теорему Стокса, объясните, с помощью этого результата проверить ответы в пунктах (b), (c) и (d).

- 15.15. Рассмотрите 2-комплекс, изображенный на рис. 15.33.

- (a) Пусть  $W$  — 1-коцепь. Выразите 2-коцепь  $dW$  через  $W$ , т. е. выразите  $\int_{S_1} dW$  и  $\int_{S_2} dW$  через  $W_\alpha$  и др.
- (b) Рассмотрите 1-форму  $\omega = y dx + x dy = d(xy)$ . Вычислите 1-коцепь  $W$ , определенную интегрированием  $\omega$  по ветвям. Проверьте, что  $dW = 0$ .

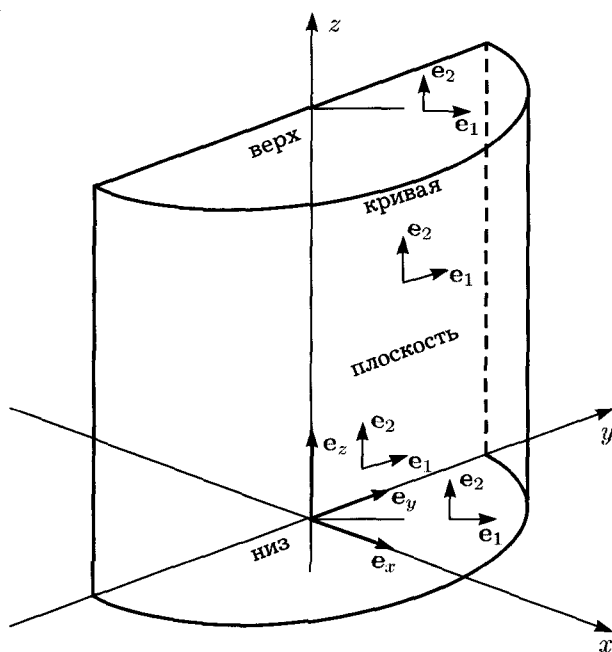


Рис. 15.32

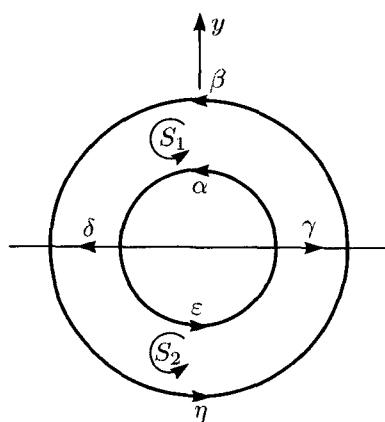


Рис. 15.33

- (с) Сделайте то же самое для 1-формы  $\omega = (x dy - y dx)/r^2$ . Покажите, что она определяет коцепь  $W$ , которая является коциклом, но не кограницей. Будет ли  $W$  коциклом, если  $S_3$ , внутренний круг, ограниченный  $\alpha$  и  $\varepsilon$ , тоже включить в комплекс?

15.16. Единичный куб, изображенный на рис. 15.34, определяет 3-комплекс с шестью гранями:

$$\begin{array}{ll} S_1: x = 0, & 0 \leq y, z \leq 1; & S_4: y = 1, & 0 \leq y, z \leq 1; \\ S_2: x = 1, & 0 \leq y, z \leq 1; & S_5: z = 0, & 0 \leq x, y \leq 1; \\ S_3: y = 0, & 0 \leq x, z \leq 1; & S_6: z = 1, & 0 \leq x, y \leq 1. \end{array}$$

Каждая грань ориентирована против часовой стрелки, если смотреть на нее снаружи. Внутренняя область куба  $R$  имеет границу  $\partial R = S_1 + S_2 + S_3 + S_4 + S_5 + S_6$ .

- (а) Пусть  $T$  — произвольная 2-коцепь. Получите выражение для  $dT$ .
- (б) Имеется 2-форма  $\tau = x^2 dy \wedge dz + xy dz \wedge dx + xz dx \wedge dy$ . Постройте соответствующую 2-коцепь  $T$ , интегрируя  $\tau$  по каждой грани. Вычислите  $d\tau$  и постройте соответствующую 3-коцепь. Проверьте, что она равна  $dT$ .
- (с) Имеется 1-форма  $\omega = x^2 y dz$ . Вычислите  $d\omega$ . Получите 1-коцепь  $W$ , соответствующую  $\omega$  на ветвях  $\alpha, \beta, \gamma, \delta, \varepsilon, \eta$ . Получите 2-коцепь  $T$ , соответствующую  $d\omega$  на гранях  $S_2$

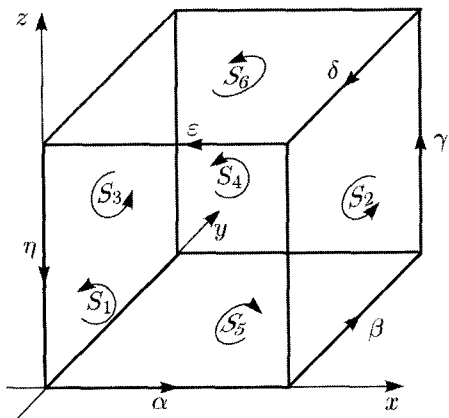


Рис. 15.34

и  $S_3$ . Проверьте, что  $T$ , полученная для  $S_2 + S_3$ , равна  $W$ , полученной для границы  $S_2 + S_3$ .

- 15.17. (а) В пространстве  $\mathbb{R}^3$  имеется 2-форма  $\tau = a \, dx \wedge dy + b \, dx \wedge dz + c \, dy \wedge dz$ , где  $a, b, c$  — дифференцируемые функции  $x, y, z$ . Пусть  $\phi$  будет отображением из  $\mathbb{R}^2$  (координаты  $u, v$ ) в  $\mathbb{R}^3$  (координаты  $x, y, z$ ). Оно задано формулами

$$\phi^*x = F(u, v), \quad \phi^*y = G(u, v), \quad \phi^*z = H(u, v).$$

Проделав вычисления в явном виде, проверьте, что  $\phi^*(d\tau) = d(\phi^*\tau)$ . Конечно, придется использовать цепное правило.

- (b) Пусть  $\mathbf{v}_1$  и  $\mathbf{v}_2$  вектора в пространстве  $\mathbb{R}^2$ . Тогда перенос  $\phi^*\tau$  определяется в виде

$$\phi^*\tau(\mathbf{v}_1, \mathbf{v}_2) = \tau[d\phi(\mathbf{v}_1), d\phi(\mathbf{v}_2)]$$

(имейте ввиду, что в каждой точке  $d\phi$  представляется матрицей Якоби  $3 \times 2$ ). Используя это определение и цепное правило, докажите, что  $\phi^*(d\tau) = d(\phi^*\tau)$ .

- 15.18. Используя оператор  $S$ , описанный в параграфе 15.4, сделайте следующее.

- (а) Найдите такую функцию  $f(x, y, z)$ , чтобы

$$df = ye^{yz} \, dx + (x + xyz)e^{yz} \, dy + xy^2e^{yz} \, dz.$$

- (b) Найдите такую 1-форму  $\omega$ , чтобы

$$d\omega = x \, dy \wedge dz + y \, dx \wedge dz.$$

- (c) Найдите такую 2-форму  $\tau$ , чтобы

$$d\tau = xy^2z^3 \, dx \wedge dy \wedge dz.$$

- 15.19. Пусть  $B = z^2 \, dx \wedge dy + yz \, dx \wedge dz - xz \, dy \wedge dz$ .

- (а) Пусть  $W$  — любая область в пространстве  $\mathbb{R}^3$ . Покажите, что

$$\int_{\partial W} B = 0.$$

- (b) Найдите такую 1-форму  $A$ , чтобы  $dA = B$ .

- 15.20. Рассмотрите 2-комплекс, изображенный на рис. 15.35. Он имеет четыре узла, пять ветвей и одну 2-ячейку  $S$ , заштрихованную на рисунке. Пусть

$$\omega = d\theta = \frac{x dy - y dx}{x^2 + y^2}.$$

Определите 1-цепь  $W$ , которая соответствует  $\omega$ . Вычислите  $W^\alpha$ ,  $W^\beta$ ,  $W^\gamma$ ,  $W^\delta$ ,  $W^\varepsilon$ . Покажите, что  $W$  является коциклом, но не кограницей.

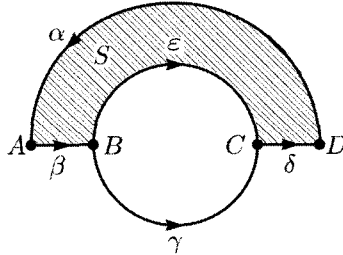


Рис. 15.35

- 15.21. Пусть  $\mathbb{R}^{1,3}$  обозначает четырехмерное пространство-время с упорядоченными аффинными координатами  $t, x, y, z$ .

- Напишите базис для  $\Lambda^2(\mathbb{R}^{1,3})$ .
- Рассмотрите 2-форму

$$W = A(x, y, t) dt \wedge dx + B(x, y, t) dt \wedge dy + C(x, y, t) dx \wedge dy.$$

Если  $dW = 0$ , то какие отношения между частными производными  $A, B, C$  должны выполняться?

## Глава 16

---

# Электростатика в пространстве $\mathbb{R}^3$

---

Глава 16 посвящена электростатике. Мы считаем, что диэлектрические свойства вакуума дают непрерывный аналог цепей с емкостями и что именно эти диэлектрические свойства определяют евклидову геометрию в трехмерном пространстве. В этой главе представлены основные положения теории потенциалов.

### 16.1. От дискретности к непрерывности

Пусть  $A^0 = \Omega^0(\mathbb{R}^3)$  обозначает пространство гладких функций в  $\mathbb{R}^3$  (мы их считаем формами нулевой степени). Пусть, далее,  $A^1 = \Omega^1(\mathbb{R}^3)$  обозначает пространство гладких линейных дифференциальных форм,  $A^2$  — пространство гладких форм второй степени, а  $A^3$  — пространство гладких форм третьей степени. Мы будем рассматривать  $A^i$  как *зародыши коцепи* степени  $i$ , т. е. если в пространстве  $\mathbb{R}^3$  у нас есть комплекс, то каждая форма степени  $i$  в результате интегрирования определяет коцепь степени  $i$ . Таким образом  $i$ -формы определяют правила, приписывающие цепям соответствующие числа. На эту картину можно взглянуть и с другой стороны. Если мы фиксируем определенную цепь с степени  $i$ , то считаем, что  $c$  определяет линейную функцию на пространстве  $A^i$ : каждой форме  $\omega$  степени  $i$  мы приписываем

число  $\int_c \omega$ . Если есть две формы  $\omega_1, \omega_2$  степени  $i$ , а  $a$  и  $b$  — вещественные числа, то из свойств операции интегрирования следует, что

$$\int_c (a\omega_1 + b\omega_2) = a \int_c \omega_1 + b \int_c \omega_2;$$

другими словами, интеграл  $\int_c \omega$  является линейной функцией формы  $\omega$ .

В качестве простейшей иллюстрации рассмотрим 0-цепь с. Пусть эта 0-цепь состоит из единственной точки  $P \in \mathbb{R}^3$  с ориентацией  $+$ . В этом случае интегрирование сводится к простому вычислению: цепь  $c$  порождает линейную функцию на пространстве  $A^0$ , которая каждой функции  $f \in A^0$  приписывает значение  $f(P)$ . Можно полагать, что  $f(P)$  является энергией, соответствующей такому распределению зарядов, когда в точке  $P$  находится единичный заряд и создается потенциал  $f$ . Такая «обращенная» точка зрения, когда  $P$  считается линейной функцией  $f$ , имеет определенные преимущества. Допустим, что мы хотим дискретное распределение зарядов, сконцентрированное в точке  $P$ , заменить на «размазанное», или непрерывное распределение зарядов с плотностью  $\rho$ . Предполагается, что плотность  $\rho$  — гладкая функция с компактным носителем. (Последнее означает, что  $\rho$  обращается в нуль за пределами некоторого ограниченного множества.) Мы можем поступить следующим образом. Рассмотрим 3-форму (с компактным носителем)  $\rho dx \wedge dy \wedge dz$ . Каждую функцию  $f \in A^0$  мы умножаем на  $\rho dx \wedge dy \wedge dz$  и получаем 3-форму  $f \rho dx \wedge dy \wedge dz$ , которая является 3-формой с компактным носителем в пространстве  $\mathbb{R}^3$ . Имея фиксированную ориентацию в  $\mathbb{R}^3$ , интегрируем эту 3-форму и получаем число. Таким образом, форма  $\rho dx \wedge dy \wedge dz$  определяет линейную функцию в пространстве  $A^0$ . Эта линейная функция каждой  $f \in A^0$  приписывает число  $\int_{\mathbb{R}^3} f \rho dx \wedge dy \wedge dz$ . (Заметим, что если полный интеграл от функции  $\rho$  равен 1 и если  $\rho$  становится все более сконцентрированной вокруг точки  $P$  — другими словами, если мы устремим непрерывное распределение зарядов к точечному вокруг  $P$ , — то значение интеграла будет стремиться к  $f(P)$ ). Пусть  $A_3$  обозначает пространство 3-форм с компактным носителем в пространстве  $\mathbb{R}^3$ . Мы уже показали, что каждая  $\rho dx \wedge dy \wedge dz \in A_3$  определяет

линейную функцию на пространстве  $A^0$ . Тогда элемент пространства  $A_3$  оказывается размазанной 0-цепью.

На самом деле электрические заряды, как известно, дискретны. Но сейчас мы наметим, как можно *аппроксимировать дискретное распределение непрерывным*. Предположим, что есть плотно упакованное распределение небольших зарядов  $c(P_i)$ . Тогда получаем линейную функцию

$$f \rightarrow \sum c_i f(P_i) \rightsquigarrow \begin{array}{c} \bullet \\ \bullet \quad \bullet \\ \bullet \end{array} \quad \text{Рис. 16.1}$$

от функций. Эта линейная функция может быть хорошо аппроксимирована (для широкого класса функций  $f$ ) *размазанным распределением зарядов*  $\rho dx \wedge dy \wedge dz$  в том смысле, что

$$\int f \rho dx \wedge dy \wedge dz \quad \begin{array}{c} \bullet \bullet \bullet \bullet \bullet \\ \bullet \bullet \bullet \bullet \bullet \\ \bullet \bullet \bullet \bullet \bullet \end{array} \quad \text{Рис. 16.2}$$

близко к

$$\sum c_i f(P_i).$$

Введем *непрерывные аппроксимации* для 1-цепей, 2-цепей и т. д. Пусть  $A_2$  обозначает пространство гладких 2-форм с компактным носителем. Утверждается, что элемент  $\Omega$  пространства  $A_2$  можно рассматривать как своего рода размазанную 1-цепь, т. е. что  $\Omega$  определяет линейную функцию на пространстве  $A^1$ . Чтобы это увидеть, рассмотрим  $\omega$  (элемент пространства  $A^1$ ) и образуем  $\omega \wedge \Omega$ . Это 3-форма с компактным носителем, которую можно интегрировать по пространству  $\mathbb{R}^3$  и получить число. Другими словами,  $\Omega \in A_2$  определяет линейную функцию на пространстве  $A^1$ , которая каждой  $\omega$  приписывает число  $\int_{\mathbb{R}^3} \omega \wedge \Omega$ . Аналогично, если  $A_1$  обозначает пространство гладких 1-форм с компактным носителем, то каждый элемент пространства  $A_1$  определяет линейную функцию в пространстве  $A^2$  и может рассматриваться как размазанная 2-цепь. Если  $A_0$  обозначает пространство 0-форм (т. е. гладких функций) с компактным носителем, то каждый элемент  $A_0$  определяет линейную функцию в пространстве  $A^3$  и может

рассматриваться как размазанная 3-цепь. Итак суммируем полученные результаты:

$A^0$  спаривается с  $A_3$ ,

$A^1$  спаривается с  $A_2$ ,

$A^2$  спаривается с  $A_1$ ,

$A^3$  спаривается с  $A_0$ ,

в том смысле, что если  $\omega \in A^i$  и  $\Omega \in A_{3-i}$ , то мы получаем число  $\int_{\mathbb{R}^3} \omega \wedge \Omega$ . Для фиксированной  $\Omega$  это линейная функция  $\omega$ , а для фиксированной  $\omega$  это линейная функция  $\Omega$ . Кроме того, это спаривание инвариантно относительно любого сохраняющего ориентацию взаимно-однозначного отображения  $\phi$  пространства  $\mathbb{R}^3$  на себя, имеющего дифференцируемое обратное. Если  $\phi$  — именно такое отображение, то для любой 3-формы  $\tau$  с компактным носителем из формулы замены переменных в интеграле получаем  $\int_{\mathbb{R}^3} \phi^* \tau = \int_{\mathbb{R}^3} \tau$ . С другой стороны,  $\phi^*(\omega \wedge \Omega) = \phi^* \omega \wedge \phi^* \Omega$ . Поэтому

$$\int \phi^* \omega \wedge \phi^* \Omega = \int \omega \wedge \Omega.$$

Другими словами, спаривание инвариантно относительно сохраняющей ориентацию замены координат.

## 16.2. Граничный оператор

Поскольку у нас есть отображение  $d : A^i \rightarrow A^{i+1}$ , то мы имеем также сопряженное отображение  $\partial$ , которое каждой линейной функции  $c$  на пространстве  $A^{i+1}$  приписывает линейную функцию  $\partial c$  на пространстве  $A^i$  по формуле  $(\partial c)(\omega) = c(d\omega)$ . Это означает, что мы должны уметь найти такое отображение  $\partial$  из  $A_{i+1}$  в  $A_i$ , чтобы для любого  $\sigma \in A_{i+1}$  и любого  $\omega \in A^i$  было справедливо равенство

$$\int_{\mathbb{R}^3} \omega \wedge \partial \sigma = \int_{\mathbb{R}^3} d\omega \wedge \sigma.$$

Чтобы найти такой оператор  $\partial$ , заметим, что

$$d(\omega \wedge \sigma) = d\omega \wedge \sigma + (-1)^i \omega \wedge d\sigma$$

и по теореме Стокса (благодаря тому, что  $\omega \wedge \sigma$  имеет компактный носитель, так что интегрирование по  $\mathbb{R}^3$  можно заменить интегрированием по некоторой конечной области, на границе которой форма  $\omega \wedge \sigma$  обращается в нуль) интеграл  $\int_{\mathbb{R}^3} d(\omega \wedge \sigma)$  становится равным нулю. Тогда

$$(-1)^{i+1} \int_{\mathbb{R}^3} \omega \wedge d\sigma = \int_{\mathbb{R}^3} d\omega \wedge \sigma.$$

Отсюда мы видим, что

$$\partial = (-1)^{i+1} d.$$

Другими словами,

$$\begin{aligned} \partial &= -d && \text{в пространстве } A_0, \\ \partial &= d && \text{в пространстве } A_1, \\ \partial &= -d && \text{в пространстве } A_2, \\ \partial &= d = 0 && \text{в пространстве } A_3. \end{aligned}$$

Таким образом, мы получили последовательность отображений

$$\begin{aligned} A^0 &\xrightarrow{d} A^1 \xrightarrow{d} A^2 \xrightarrow{d} A^3, \\ A_3 &\xleftarrow{\partial} A_2 \xleftarrow{\partial} A_1 \xleftarrow{\partial} A_0, \end{aligned}$$

где  $\partial = (-1)^{i+1} d$  в пространстве  $A_{i+1}$ .

Иногда оператор  $\partial$  называют *формальным транспонированием*  $d$ . Причина для появления слова «формальный» состоит в том, что оператор  $\partial$  может быть выражен через  $d$ , и формула

$$\int \omega \wedge \partial \sigma = \int d\omega \wedge \sigma$$

справедлива только тогда, когда интегрирование ведется по области, где границы не вносят вклада. Мы обеспечили это, проинтегрировав по всему пространству при условии, что  $\sigma$  имеет компактный носитель. Эта формула выполняется также, если  $\omega$  имеет компактный носитель, т.е. мы считаем, что  $d$  является отображением из  $A_i$  в  $A_{i+1}$ , а  $\partial$  — отображением из  $A^i$  в  $A^{i+1}$ . Таким образом, мы можем рассматривать формы с компактным носителем как фундаментальные объекты и считать, что отображение  $\partial$  определено на пространстве линейных функций  $A_i$ .

### 16.3. Телесный угол

В качестве иллюстрации вышеуказанной обращенной точки зрения, давайте сделаем ряд вычислений. Пусть дана точка  $P$  в пространстве  $\mathbb{R}^3$  с координатами  $x_P, y_P, z_P$ . Пусть  $\tau_P$  обозначает *телесный угол*, наблюдаемый из точки  $P$ , определяемый 2-формой

$$\tau_P = \frac{(x - x_P) dy \wedge dz - (y - y_P) dx \wedge dz + (z - z_P) dx \wedge dy}{r_P^3}.$$

Здесь  $r_P$  обозначает расстояние до точки  $P$ . Форма  $\tau_P$  определена не на всем пространстве  $\mathbb{R}^3$ , а только в пространстве  $\mathbb{R}^3 - \{P\}$ , т. е. в трехмерном пространстве, из которого удалена точка  $P$ . И тем не менее форма  $\tau_P$  определяет линейную функцию в пространстве  $A_1$ . Действительно, пусть  $\omega = a dx + b dy + c dz$  обозначает линейную дифференциальную форму с гладкими коэффициентами и компактным носителем. Тогда

$$\omega \wedge \tau_P = r_P^{-3} [a(x - x_P) + b(y - y_P) + c(z - z_P)] dx \wedge dy \wedge dz.$$

Коэффициент  $dx \wedge dy \wedge dz$  не определен в точке  $P$ , но особенность в  $P$  порядка  $r_P^{-2}$  (поскольку в точке  $P$  выражение в квадратных скобках имеет нуль первого порядка). Следовательно, при вычислении интеграла точка  $P$  не создает проблем. Поскольку  $\omega$  имеет компактный носитель, то интеграл по всему пространству определен корректно. Чтобы вычислить этот интеграл, удобно перейти к сферическим координатам с центром в точке  $P$ . Если  $r_P, \theta, \phi$  — такие координаты, то

$$\tau_P = \sin \theta d\theta \wedge d\phi.$$

(Таким образом,  $\tau_P$  дает нам телесный угол, заматаемый поверхностью, по которой мы интегрируем, рис. 16.3.) Предположим, что в этих сферических координатах  $\omega = F dr_P + G d\theta + H d\phi$ . Тогда

$$\omega \wedge \tau_P = F \sin \theta dr_P \wedge d\theta \wedge d\phi.$$

Итак,

$$\int_{\mathbb{R}^3} \omega \wedge \tau_P = \int_0^{2\pi} \int_0^\pi \sin \theta \left\{ \int_0^\infty F(r_P, \theta, \phi) dr_P \right\} d\theta d\phi.$$

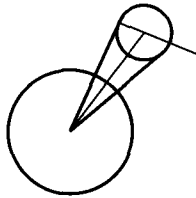


Рис. 16.3

Давайте рассмотрим случай, когда  $\omega = du$  для некоторой гладкой функции  $u$  с компактным носителем. Тогда в предыдущей формуле в сферических координатах  $F = \partial u / \partial r_P$ , и внутренний интеграл в правой части сводится к постоянному значению  $-u(P)$ , так что интегрирование по  $\theta$  и  $\phi$  сводится к умножению на площадь поверхности единичной сферы  $4\pi$ . Таким образом, мы доказали базисную формулу

$$\int du \wedge \tau_P = -4\pi u(P). \quad (16.1)$$

Пусть  $\delta_P$  обозначает 0-цепь, которая каждой функции  $u$  приписывает значение  $u(P)$ , т. е.  $\delta_P$  представляет единичный заряд, сосредоточенный в точке  $P$ .<sup>1</sup> Тогда правая часть уравнения (16.1) есть значение 0-цепи на функции  $u$ , равное  $-4\pi\delta_P$ . А левая часть уравнения есть значение  $\tau_P$ , вычисленное для  $du$ . Тогда уравнение (16.1) можно записать в виде

$$\partial \tau_P = -4\pi \delta_P. \quad (16.2)$$

(Заметим, что в предыдущем уравнении  $\partial$  нельзя приравнять к  $d$ . Форма  $\tau_P$  не определена в точке  $p$ , поэтому  $d\tau_P$  там тоже не определена. Во всех точках, где  $\tau_P$  определена, мы имеем  $d\tau_P = 0$ . И тем не менее уравнение (16.2) справедливо как равенство 0-цепей, т. е. линейных функции от функций.)

Предположим, что у нас есть конечное число точек  $P_1, P_2, P_3$  и т. д. Заряд  $c_1$  помещаем в  $P_1$ ,  $c_2$  в  $P_2$  и т. д. Значит, мы

<sup>1</sup>Принято называть  $\delta_P$  дельта-функцией Дирака, сосредоточенной в точке  $P$ . Она относится к так называемым обобщенным функциям, которые играют фундаментальную роль во многих задачах анализа и математической физики. — Прим. ред.

рассматриваем 0-цепь

$$Q = c_1 \delta_{P_1} + c_2 \delta_{P_2} + \dots.$$

Пусть

$$D = -(c_1 \tau_{P_1} + c_2 \tau_{P_2} + \dots).$$

Тогда  $D$  корректно определяет линейную функцию в пространстве  $A_1$ , и мы имеем формулу

$$\partial D = -4\pi Q.$$

Если вместо дискретного распределения конечного числа зарядов, нам дано распределение зарядов с гладкой функцией плотности  $\rho$ , имеющей компактный носитель, то предыдущее определение  $D$  через сумму заменяется на соответствующий интеграл, т. е. мы теперь пишем  $D = \int \rho(P) \tau_P dP$  и  $Q = \rho dx \wedge dy \wedge dz$ , при этом выполняется предыдущее уравнение. В этом случае непрерывного распределения несложно показать (мы это сделаем очень скоро), что  $D$  — гладкая 2-форма, определенная во всем пространстве.

## 16.4. Напряженность электрического поля и электрическое смещение

Сейчас мы уже можем дать определения фундаментальных понятий теории электростатики. *Напряженность электрического поля*  $E$  является линейной дифференциальной формой, проинтегрировав которую по любой траектории, мы получаем падение напряжения на ней. Величина  $E$  имеет физическую размерность напряжение/длину. (Поскольку размерностью напряжения является энергия/заряд, а размерностью силы является энергия/длину, то можно также считать, что  $E$  измеряет силу/заряд.) Основное уравнение, которому удовлетворяет  $E$ , имеет вид

$$dE = 0.$$

Это уравнение локально эквивалентно существованию потенциала  $u$ , т. е. условию

$$E = -du. \quad (16.3)$$

Мы будем пользоваться этим уравнением для большинства рассматриваемых областей.

Будем рассматривать  $E$  как зародыш 1-цепи, а  $u$  — как зародыш 0-цепи. Кроме того нам понадобится объект, представляющий размазанную 1-цепь. Это будет 2-форма  $D$ , называемая *электрическим смещением*. Она должна представлять размазанный вариант 1-цепи, дающей заряды ветвей. Мы также хотим иметь размазанный вариант 0-цепи, представляющей заряды узлов. Это будет 3-форма  $\rho dx \wedge dy \wedge dz$ . В нашей модели емкостной цепи мы имели

$$\partial(\text{заряды вервей}) = -\text{заряды узлов}.$$

Поэтому мы ожидаем, что

$$\partial D = -\rho dx \wedge dy \wedge dz.$$

Стандартное определение единиц измерения  $D$  и  $\rho$  (в системе СГС) таково, что это соотношение верно с точностью до множителя  $4\pi$ :

$$\boxed{\partial D = -4\pi \rho dx \wedge dy \wedge dz.} \quad (16.4)$$

Это называется *законом Гаусса*. Вследствие теоремы Стокса и того, что для 2-форм выполняется условие  $\partial = -d$ , мы можем написать

$$\int_{\partial W} D = 4\pi \int_W \rho dx \wedge dy \wedge dz,$$

что связывает интеграл от  $D$  по поверхности, ограничивающей область  $W$ , с полным зарядом внутри нее.

Конечно, уравнение  $\partial D = -4\pi \rho dx \wedge dy \wedge dz$  не определяет  $D$  однозначно. Прибавление к  $D$  любой замкнутой формы не влияет на уравнение. Частично 2-форма  $D$  дает нам информацию об электростатической системе. В принципе она может быть измерена следующим образом. Напишем

$$D = D_x dy \wedge dz + D_y dz \wedge dx + D_z dx \wedge dy.$$

Чтобы измерить  $D_z$ , поместим в исследуемую среду две небольшие параллельные плоскости  $xy$  металлические пластины

(рис. 16.4). Соединим пластины вместе, а потом вновь разделим их. Пластины приобретут заряды  $\pm Q$ . Тогда

$$D_z = 4\pi \lim_{\text{площадь} \rightarrow 0} \frac{\text{заряд верхней пластины (со стороны } +z)}{\text{площадь пластин}}.$$

( $4\pi$  появляется за счет системы СГС.) Этот метод работает в любом диэлектрике, даже если мы не знаем его диэлектрическую проницаемость  $\epsilon$ , и при этом *нигде не упоминается электрическое поле*.

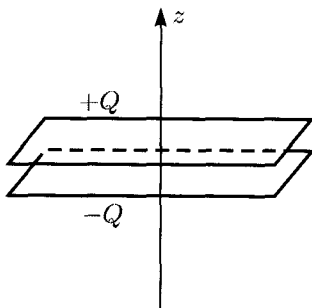


Рис. 16.4

По существу это определение  $D$  можно обобщить и дать определение  $D$ , не связанное с системой координат. Для любой пары векторов  $\mathbf{v}_1$  и  $\mathbf{v}_2$  образуем плоскопараллельный конденсатор (рис. 16.5), пластинами которого являются параллелограммы, построенные на векторах  $h\mathbf{v}_1$  и  $h\mathbf{v}_2$ . Тогда

$$D(\mathbf{v}_1, \mathbf{v}_2) = 4\pi \lim_{h \rightarrow 0} \frac{\text{заряд на верхней пластине}}{h^2}.$$

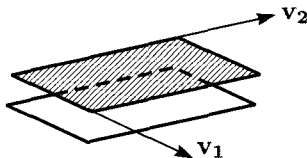


Рис. 16.5

Какую пластину считать верхней, зависит от ориентации<sup>2</sup>, поскольку ориентация изменяется при перестановке  $\mathbf{v}_1$  и  $\mathbf{v}_2$ :

$$D(\mathbf{v}_2, \mathbf{v}_1) = -D(\mathbf{v}_1, \mathbf{v}_2).$$

Поскольку заряд аддитивен, то  $D$  — билинейная функция  $\mathbf{v}_1$  и  $\mathbf{v}_2$ . Поэтому, по определению,  $D$  является 2-формой.

Итак, мы обобщили топологические уравнения для сетей с емкостями на случай электростатики.

$\mathbf{V} = -d\Phi$	заменяется	$E = -du;$
$\partial Q = -\rho$	заменяется	$\partial D = -4\pi\rho dx \wedge dy \wedge dz.$

Если  $D$  — гладкая 2-форма, то  $dD = +4\pi\rho dx \wedge dy \wedge dz$ .

Сейчас мы должны вспомнить, что  $E$  может рассматриваться (при выборе, например, стандартной ориентации) как линейная функция от  $D$ , которая приписывает  $D$  число

$$\int_{\mathbb{R}^3} E \wedge D.$$

В этой формуле мы все еще считаем, что либо  $D$ , либо  $E$  имеет компактный носитель. Но по существу этот интеграл может быть определен для другой пары  $E$  и  $D$ , у которой произведение  $E \wedge D$  достаточно быстро стремится к нулю на  $\infty$ . Например, если произведение стремится к нулю как  $r^{-4}$  или быстрее, то интеграл сходится. Мы не должны забывать об этой возможности.

Теперь у нас есть непрерывные размазанные версии топологической части нашей теории цепей с емкостями. Но нам еще надо иметь вариант матрицы  $C$ , удовлетворяющей соотношению  $Q = CV$ . Поэтому мы хотим иметь отображение

$$D = C(E),$$

определяющее электрическое смещение по напряженности поля. В следующем параграфе мы будем подробно изучать такое отображение  $C$ . Здесь мы получим ряд следствий из описанных ниже предположений относительно  $C$ .

<sup>2</sup>Очевидно, имеется в виду, что векторы  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{w}$  образуют положительно ориентированный базис  $\mathbb{R}^3$ , если  $\mathbf{w}$  направлен от нижней пластины к верхней.

— Прим. ред.

Вспомним, что в теории цепей отображение  $C$  было диагональной матрицей относительно напряжений ветвей. В частности, если на какой-то ветви напряжение равно нулю, то обращается в нуль и заряд этой ветви. В качестве слабой аналогии с этим условием предположим, что

- (а) отображение  $C$  локально в том смысле, что если  $E$  тождественно равно нулю внутри какой-то области, то  $D = C(E)$  тоже там тождественно равно нулю.

По аналогии с теорией цепей предположим, что

- (b) отображение  $C$  симметрично, т. е. если есть два электрических поля  $E_1$  и  $E_2$ , то

$$\int E_1 \wedge C(E_2) = \int E_2 \wedge C(E_1).$$

В следующем параграфе мы опишем отображение  $C$  с позиций его реальной физической интерпретации. Мы увидим, что существует тесная связь между отображением  $C$  для вакуума и геометрией Евклида. Во *всех* средах выполняются условия (а) и (b). А сейчас мы получим некоторые следствия из них.

По определению, проводящий материал не оказывает сопротивления при движении электрического заряда. Другими словами, в проводнике форма  $E$ , которая измеряет работу, выполненную при движении электрического заряда, должна обращаться в нуль. Тогда, используя свойство (а), мы видим, что

*Внутри любого проводника формы  $E$  и  $D$  должны обращаться в нуль. Следовательно, функция  $\rho$  должна быть постоянна на каждой системе соединенных проводящих тел.*

Так как внутри проводника  $D = 0$ , то и  $\rho = 0$ . Поэтому внутри проводников не может быть электрических зарядов (в рамках электростатики). Или, говоря словами Фарадея,

*все заряды проводника располагаются на его поверхности.*

Конечно, плотность этого поверхностного заряда определяется формой  $D$ . Полный заряд  $\rho_i$  отдельного проводника получается интегрированием  $D$  по его поверхности.

Давайте предположим, что заряды есть только на проводниках. Мы считаем, что вне проводников  $dD = 0$ . Тогда для двух потенциалов  $u$  и  $\hat{u}$  с соответствующими  $D$  и  $\hat{D}$  мы имеем

$$\begin{aligned} d(u\hat{D}) &= du \wedge \hat{D} + u d\hat{D} \\ &= -E \wedge \hat{D} \end{aligned}$$

вне проводников, так как там  $d\hat{D} = 0$ . Тогда по теореме Стокса

$$(E, \hat{E}) \int_{\text{по области вне проводников}} E \wedge \hat{D} = \int_{\text{по поверхностям всех проводников}} u \hat{D}.$$

На поверхности  $i$ -го проводника  $u$  принимает постоянное значение  $\Phi_i$ . Тогда  $u$  можно вынести из поверхностного интеграла и получить

$$(E, \hat{E}) = \sum_{\text{проводникам}} \Phi \hat{\rho}.$$

Это теорема взаимности Грина, рассмотренная в параграфе 13.8.

Третье свойство отображения  $C$  состоит в том, что скалярное произведение  $(\cdot, \cdot)$  положительно определено. Таким образом, если в предыдущем уравнении взять  $E = \hat{E}$ , то получим

$$(E, E) = \sum_{\text{проводникам}} \Phi \rho.$$

Если заряды всех проводников равны нулю, то правая часть этого уравнения равна нулю. Вследствие положительной определенности это значит, что  $E = 0$  и следовательно все потенциалы  $\Phi$  равны. Если мы рассматриваем один проводник (скажем, воображаемый проводник на бесконечности) как заземленный, то заряды  $\rho$  определяют потенциал  $\Phi$  единственным образом. Аналогично, потенциалы однозначно определяют заряды. Тогда мы имели бы отображение из пространства потенциалов всех проводников  $\Phi$  на пространство зарядов всех проводников  $\rho$ , и это дало бы нам возможность использовать результаты главы 13.

При этом предполагается, что для *любого* заданного распределения потенциалов  $\Phi$  по различным проводникам мы можем получить соответствующие  $E$  и  $D$  (единственным образом). В этом вопросе взгляды математика и физика расходятся.

С точки зрения математика поставлена интересная и нетривиальная *задача Дирихле*. Это вопрос теоремы существования в теории уравнений с частными производными. Решение ее потребовало огромных усилий многих талантливых математиков последней трети девятнадцатого века. В конце концов она была решена положительно с помощью различных методов, каждый из которых вызвал к жизни интересное математическое развитие, часто с важными следствиями в физике. Соответствующая проблема для более общих уравнений с частными производными более общего вида все еще активно исследуется. В следующих параграфах мы еще поговорим об этой математической проблеме, но ее подробное обсуждение не входит в задачу этой книги.

С точки зрения физика вопрос о существовании не столь важен и относится к логической последовательности математической идеализации физической модели. Физик знает, что гипотеза о распределении статического заряда является лишь приближением к реальности. Если модель работает, т. е. если идеализация распределения статического заряда на поверхности проводника при постоянных потенциалах и соответствующих электрических полях дает разумное описание ситуации, то, конечно, теорема существования должна выполняться, потому что каждый проводник мы можем снабдить любым зарядом, каким захотим. Если же теорема существования нарушается, то либо потому, что математики придумали немыслимые проводники, которые нельзя построить на практике (или другие «физически бессмысленные» условия), либо потому, что есть фундаментальная ошибка в теории. Гораздо проще вывести логические следствия из предполагаемого существования решения и проверить на опыте, выполняются ли физические предсказания. Если они выполняются, то нарушения теоремы существования, вероятно, представляют что-то экзотическое и не имеет большого физического значения.

В дальнейшем этот параграф излагается с физической точки зрения. Мы предполагаем существование решений для данного

распределения зарядов или напряжений на проводнике. Предположим также, что решения зависят от заданных граничных условий непрерывно.

А теперь получим несколько важных заключений.

**Принцип электрической экранировки.** Допустим, что некоторые из проводников, скажем, с первого по седьмой, полностью окружены другим — восьмым — проводником (рис. 16.6). Допустим, что на окруженных проводниках нет зарядов. Рассмотрим произвольное распределение зарядов на остальных проводниках. Тогда возникнет электрическое поле  $E = -du$  с соответствующим электрическим смещением  $D$  и потенциалами, индуцированными на всех проводниках. В результате действия зарядов восьмой проводник приобретет постоянный потенциал  $\Phi_8$ .

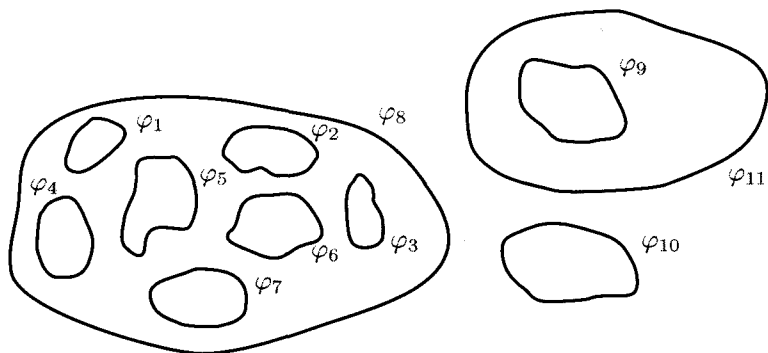


Рис. 16.6

Априори мы не знаем, каковы индуцированные заряды и потенциалы на внутренних проводниках. Предположим, что мы рассматриваем новую функцию  $u' = u$  вне и  $u' \equiv \Phi_8$  внутри полости восьмого проводника. Это является решением нашей задачи, причем внутри полости  $E \equiv D \equiv 0$ , и на всех внутренних проводниках заряды нулевые. В силу единственности оно совпадает с исходным решением, потому что ни один заряд не попал на внутренние проводники, и то же самое было бы и при изменении потенциалов внешних проводников. Короче, внутренняя полость восьмого проводника экранирована от всей электростатической системы вне этого проводника. Этот принцип используется

при создании электростатических измерительных инструментов, чтобы внешние электрические помехи не влияли на наблюдаемый эффект. (Если мы хотим иметь возможность видеть внутренность полости, то можно оставить маленькое отверстие или заменить часть проводника проволочной сеткой, которая экранирует почти так же хорошо, как и сплошной проводник.) Помимо практической важности, принцип электрического экранирования очень полезен при вычислениях, когда он применяется вместе со следующим принципом.

**Замена эквипотенциальной поверхности проводящим слоем.** Предположим, что для некоторой системы зарядов или потенциалов у нас есть решение  $E = -du$ ,  $D = CE$ . Пусть  $S$  обозначает поверхность, на которой потенциал  $u$  постоянен и равен  $\Phi$ . (Тогда  $S$  называется эквипотенциальной поверхностью.) Допустим, что поверхность  $S$  заменили тонким проводящим слоем, на котором поддерживается потенциал  $\Phi$ . Природа отображения  $C$  такова, что оно фактически не изменяется вставкой этого проводника. Поэтому вставка проводника практически не производит эффекта. Внутри этого тонкого проводящего слоя  $E$  и  $D$  обращаются в нуль, заряды аккумулируются на его двух поверхностях, но все остальное остается без изменения.

Например, рассмотрим среду, инвариантную относительно вращений. Тогда потенциал заряда  $Q$ , помещенного в начало координат, должен иметь вид

$$u = Af(r).$$

(В следующем параграфе мы увидим, что если система инвариантна относительно всех евклидовых движений, то

$$f(r) = \frac{c}{r},$$

где  $c$  — некоторая постоянная.) Теперь вставим тонкий проводящий сферический слой радиуса  $a$ . Поля внутри и вне этой сферы не изменятся, а  $D$  внутри обратится в нуль. Если внутри проводящего слоя сферы нарисовать еще одну сферу, то полный заряд внутри нее должен быть равен нулю. Таким образом, на внутренней поверхности проводящей сферы аккумулируется полный

заряд  $-Q$  (а на внешней — полный заряд  $Q$ ). Мы можем теперь разрядить внутреннюю поверхность (с помощью провода, соединяющего заряд в начале координат и внутренней поверхности проводящей сферы). В силу принципа электрического экранирования, вне сферы ничего не произойдет, но потенциал внутри сферы станет постоянным. Таким образом, потенциал поля, создаваемого заряженной сферой, имеет вид:

$$\begin{aligned} Qf(r), \quad r &\geq a, \\ Qf(a), \quad r &\leq a, \end{aligned}$$

где  $f(r)$  обозначает потенциал единичного заряда, находящегося в начале координат. Этот метод позволяет нам вычислять поле и емкость сферического конденсатора, пластинами которого являются два концентрических заряженных проводника (рис. 16.7).

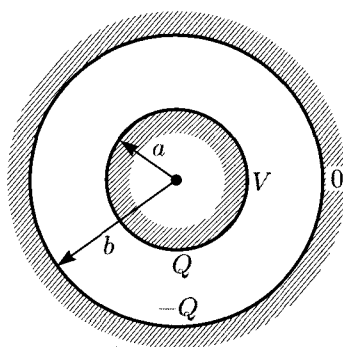


Рис. 16.7

Давайте опять поместим заряд  $Q$  в начало координат, а затем окружим его двумя сферическими проводниками. Заряды соберутся на внутренних и внешних поверхностях проводящих сфер. Разрядим внутреннюю поверхность внутренней сферы, соединив ее проводом с началом координат. А внешнюю поверхность внешней сферы разрядим, заземлив ее. Поля внутри и вне обеих сфер обратятся в нуль. Функция  $u$  между двумя сферами остается без изменения, при этом остается заряд  $Q$  на внешней поверхности внутренней сферы и заряд  $-Q$  на внутренней поверхности внешней сферы.

Рассмотрим еще один пример. Предположим, что заряд  $Q$  помещен в точку  $\begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$  на оси  $x$ , а заряд  $-Q$  в точку  $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ . Тогда получится потенциал

$$Q(f(r_+) - f(r_-)),$$

где  $r_-$  и  $r_+$  обозначают расстояния до этих двух точек. При условии инвариантности относительно евклидовых преобразований  $f(r) = c/r$ . Эквипотенциальные поверхности в этой ситуации изображены на рис. 16.8.

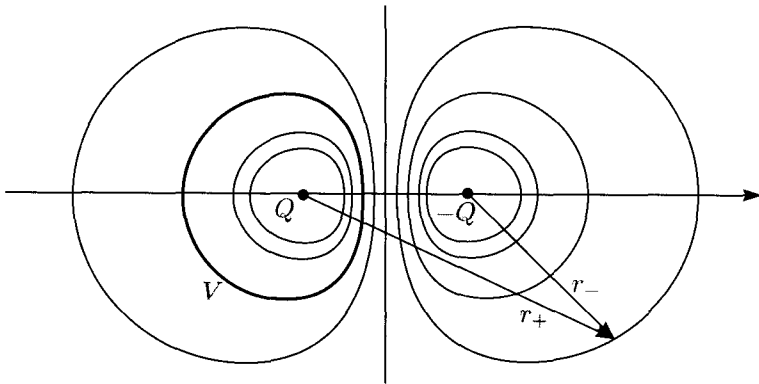


Рис. 16.8

Вдоль одной из эквипотенциальных поверхностей вставим проводящий слой. Поле с одной стороны этого проводника убираем, соединив его с зарядом  $-Q$ . Тогда мы получаем поле, создаваемое точечным зарядом  $Q$  и оставшейся заряженной поверхностью вставленного проводника (рис. 16.9). Например, если в качестве проводящей поверхности мы возьмем плоскость  $x = 0$  с нулевым потенциалом (рис. 16.10), то получим решение для задачи с точечным зарядом в присутствии заземленного плоского проводника. На поверхности проводника, обращенной к заряду  $Q$ , индуцируется заряд  $-Q$ . Поле с другой стороны проводника обращается в нуль, а поле со стороны заряда будет таким же, как если бы на другой стороне проводника (в симметричной точке)

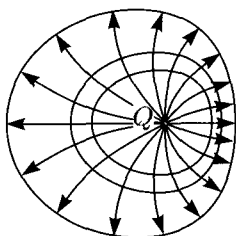


Рис. 16.9

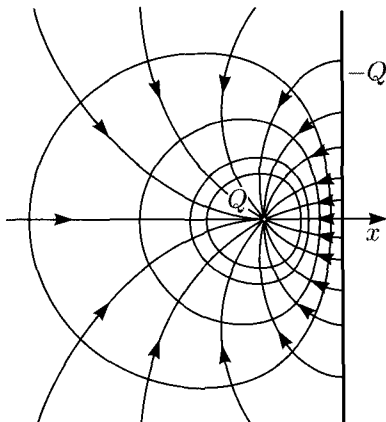


Рис. 16.10

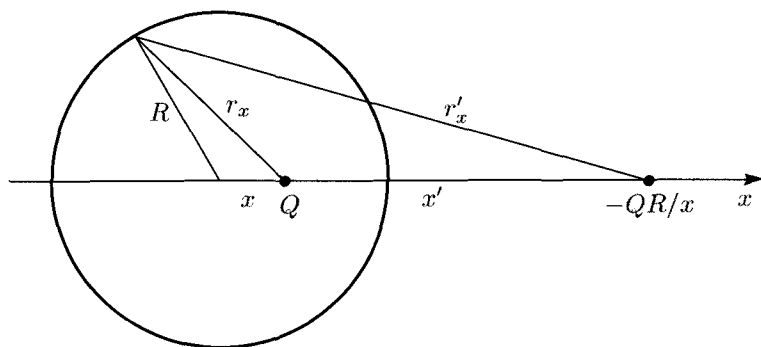


Рис. 16.11

находился заряд  $-Q$ . Это один из примеров применения метода образов.

И наконец, еще один пример, который мы используем позднее. Рассмотрим сферу радиуса  $R$ . Заряд  $Q$  поместим в точку  $\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix}$ . Пусть  $x'$  таково, что  $xx' = R^2$ . Из подобных треугольников (рис. 16.11) следует, что для любой точки на поверхности сферы

$$\frac{x'}{r_{x'}} = \frac{R}{r_x} \quad \text{или} \quad \frac{1}{r_{x'}} = \frac{R}{x' r_x}.$$

Поместив заряд  $-Q(R/x)$  в точку  $\begin{pmatrix} x' \\ 0 \\ 0 \end{pmatrix}$ , мы получаем, что потенциал на сфере равен нулю. Это позволяет вычислить потенциал для точечного заряда, находящегося в любой точке внутри заземленной проводящей сферы.

## 16.5. Диэлектрическая проницаемость

Рассмотрим задачу нахождения электростатического аналога отображения из  $C^1$  в  $C_1$ , определяемого матрицей емкостей, т. е. отображения  $\mathbf{Q} = \mathbf{C}\mathbf{V}$ .

Вспоминаем, что матрица  $\mathbf{C}$  *диагональна*, если в качестве базиса взяты ветви. Другими словами, вклад распределения зарядов в распределение напряжений в любой ветви локализован на этой ветви. Примем аналогичную гипотезу в электростатике: мы хотим получить такое линейное отображение  $E \mapsto D$ , чтобы в любой точке значение  $D$  зависело от значения  $E$  в этой же точке. Другими словами, если

$$E = E_x dx + E_y dy + E_z dz$$

и

$$D = D_x dy \wedge dz + D_y dz \wedge dx + D_z dx \wedge dy$$

(где коэффициенты  $E_x, \dots, D_z$  являются функциями), то соотношение между  $E$  и  $D$  определяется матрицей из функций

$$\varepsilon = (\varepsilon_{ij}).$$

Тогда

$$D_x = \varepsilon_{11}E_x + \varepsilon_{12}E_y + \varepsilon_{13}E_z,$$

$$D_y = \varepsilon_{21}E_x + \varepsilon_{22}E_y + \varepsilon_{23}E_z,$$

$$D_z = \varepsilon_{31}E_x + \varepsilon_{32}E_y + \varepsilon_{33}E_z.$$

Каждый матричный элемент  $\varepsilon_{ij}$  является функцией, и матрица  $\varepsilon = (\varepsilon_{ij})$  определяет *диэлектрическую проницаемость* среды. Связь между  $D$  и  $E$  можно записать в форме

$$D = \varepsilon E.$$

Теперь электростатические уравнения принимают вид

$$E = -du, \quad \partial D = -4\pi\rho dx \wedge dy \wedge dz$$

и

$$D = \epsilon E.$$

Можно собрать эти уравнения вместе, введя оператор Лапласа

$$(\Delta u) dz \wedge dy \wedge dz = -\partial \epsilon du$$

и получить уравнение Пуассона

$$\Delta u = -4\pi\rho.$$

Чтобы продвинуться дальше, надо иметь больше информации относительно  $\epsilon$ .

Для произвольной среды мы знаем только, что  $\epsilon$  — симметричная матрица. Причины те же, что и для матрицы  $C$ . Больше ничего сказать нельзя. Существуют, например, кристаллы под давлением, для которых  $\epsilon$  — переменная симметричная матрица.

Среда называется *однородной*, если  $\epsilon$  инвариантна относительно трансляций, и в этом случае она постоянна. Среда называется *изотропной*, если соотношение между  $E$  и  $D$  инвариантно относительно вращений. Далее мы рассмотрим однородные и изотропные среды и посмотрим, как выглядит  $\epsilon$  в этих случаях.

## 16.6. Звездный оператор в трехмерном евклидовом пространстве

Начнем с линейного отображения

$$\star : \Lambda^1(\mathbb{R}^{3*}) \rightarrow \Lambda^2(\mathbb{R}^{3*}),$$

которое инвариантно относительно вращений. По определению

$$\star dx = dy \wedge dz,$$

$$\star dy = dz \wedge dx,$$

$$\star dz = dx \wedge dy.$$

Далее нужно продолжить оператор по линейности, т. е. определить  $\star : \Lambda^1(\mathbb{R}^{3*}) \rightarrow \Lambda^2(\mathbb{R}^{3*})$  правилом

$$\star(a\,dx + b\,dy + c\,dz) = a\,dy \wedge dz + b\,dz \wedge dx + c\,dx \wedge dy.$$

Пусть

$$\omega = a\,dx + b\,dy + c\,dz$$

и

$$\sigma = A\,dx + B\,dy + C\,dz.$$

Заметим, что

$$\star\omega \wedge \sigma = (aA + bB + cC)\,dx \wedge dy \wedge dz.$$

Последнее уравнение запишем по-другому. Скалярное произведение в евклидовом пространстве  $\mathbb{R}^3$  определяет скалярное произведение в  $\mathbb{R}^{3*}$  (в пространстве 1-форм) согласно формуле

$$(\omega, \sigma) = aA + bB + cC.$$

Скалярное произведение с учетом *ориентации* выделяет конкретную форму объема  $dx \wedge dy \wedge dz$ . Тогда последнее уравнение принимает вид

$$\star\omega \wedge \sigma = (\omega, \sigma)\,dx \wedge dy \wedge dz. \quad (16.5)$$

Если немного подумать, то становится понятным, что это уравнение *определяет* отображение  $\star$  единственным образом. Однако, справа в уравнении содержится только скалярное произведение и учитывается ориентация (как функция  $\omega$  и  $\sigma$ ). Любое вращение сохраняет и то, и другое. Следовательно,

*Оператор  $\star$  инвариантен относительно вращений.*

Утверждается, что с точностью до скалярного множителя  $\star$  является единственным отображением  $\star : \Lambda^1(\mathbb{R}^{3*}) \rightarrow \Lambda^2(\mathbb{R}^{3*})$ , инвариантным относительно вращений. Таким образом, мы утверждаем, что если  $r : \Lambda^1(\mathbb{R}^{3*}) \rightarrow \Lambda^2(\mathbb{R}^{3*})$  — другое отображение, инвариантное относительно вращений, то  $r = a\star$  для некоторого скаляра  $a$ . Сначала отметим, что либо  $r = 0$ , либо  $r$  имеет

нулевое ядро. Действительно, ядро  $r$  является подпространством  $\mathbb{R}^3$ . Если  $r$  инвариантно относительно вращений, то это подпространство тоже должно быть инвариантно относительно вращений. Но в пространстве  $\mathbb{R}^3$  нет подпространств, инвариантных относительно вращений, кроме тривиальных подпространств  $\{0\}$  и  $\mathbb{R}^3$ . Таким образом, или  $r = 0$ , или  $r$  является изоморфизмом. Если  $r = 0$ , то доказывать нечего. Если же  $r$  — изоморфизм, то рассмотрим отображение

$$l = r^{-1}\star, \quad l: \mathbb{R}^{3*} \rightarrow \mathbb{R}^{3*}.$$

Согласно предположению,  $l$  инвариантно относительно вращений, т. е.

$$l(R\omega) = Rl(\omega)$$

для любого вращения  $R$ . Запишем

$$l\omega = a\omega + L\omega,$$

где  $L\omega$  перпендикулярно  $\omega$ . Другими словами, разложим  $l\omega$  на компоненты вдоль и перпендикулярно  $\omega$ . В силу вращательной инвариантности коэффициент  $a$  определяется формулой

$$(l\omega, \omega) = a\|\omega\|^2$$

и не зависит от  $\omega$ . Утверждается, что  $L\omega = 0$ . Действительно, для любого вращения  $R$  мы должны иметь  $LR\omega = RL\omega$ . Выберем поворот  $R$ , который оставляет на месте  $\omega$ , но является нетривиальным вращением в перпендикулярной  $\omega$  плоскости. Тогда  $R\omega = \omega$ , но  $RL\omega \neq L\omega$ , если  $L\omega \neq 0$ . Значит,  $L\omega = 0$ . Тогда

$$l\omega = a\omega, \quad \text{или} \quad r\omega = a\star\omega,$$

что и требовалось доказать.

Таким образом, с точностью до скалярного множителя  $\star$  — это единственное отображение из  $\Lambda^1(\mathbb{R}^{3*})$  в  $\Lambda^2(\mathbb{R}^{3*})$ , инвариантное относительно вращений.

(Стоит отметить и обратное утверждение. Оператор  $\star$  однозначно определяет скалярное произведение  $(\cdot, \cdot)$ , стоящее в правой части уравнения (16.5), следовательно, оператор  $\star$  в пространстве  $\mathbb{R}^{3*}$  определяет скалярное произведение и ориентацию.)

Пусть

$$\omega = a dx + b dy + c dz$$

— дифференциальная форма, т.е. мы считаем, что  $a$ ,  $b$  и  $c$  — функции. Как и раньше, определим оператор  $\star$  формулой

$$\star\omega = a dy \wedge dz + b dz \wedge dx + c dx \wedge dy.$$

Опять

$$\star\omega \wedge \tau = (\omega, \tau) dx \wedge dy \wedge dz,$$

где  $\omega$  и  $\tau$  — дифференциальные формы.

Мы предположили, что работаем в однородной и изотропной среде. Это значит, что

$$\epsilon = \epsilon \star,$$

где  $\epsilon$  — константа, называемая диэлектрической постоянной среды. Например, в вакууме  $\epsilon = \epsilon_0 \star$ , где  $\epsilon_0$  — константа. И обратно, поскольку оператор  $\star$  из  $\Lambda^1(\mathbb{R}^3)$  в  $\Lambda^2(\mathbb{R}^3)$  детерминирует скалярное произведение, мы можем сказать, что *диэлектрические свойства вакуума определяют евклидову геометрию пространства*.

В дальнейшем мы предполагаем, что единицы измерения  $E$ ,  $D$  и  $\epsilon$  таковы, что  $\epsilon$  поглощается оператором  $\star$ . Тогда уравнения электростатики принимают вид

$$E = -du, \quad D = \star E, \quad dD = 4\pi\rho dx \wedge dy \wedge dz,$$

откуда

$$\Delta u = -4\pi\rho,$$

где непосредственное вычисление показывает, что

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

Важно заметить, что оператор  $\star$  не инвариантен относительно дифференцируемых отображений. Поэтому, вообще говоря, мы не имеем равенства

$$\phi^*(\star\omega) = \star\phi^*(\omega).$$

Равенство будет *только в случае*, если  $\phi$  является сохраняющим ориентацию евклидовым преобразованием. Тем не менее, мы можем вычислять оператор  $\star$  в более общей системе координат, используя его определение. Например, мы хотим вычислить оператор  $\star$  в сферических координатах, определяемых формулами

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \theta = \arctan(\sqrt{x^2 + y^2}/z), \quad \phi = \arctan(y/x).$$

Вычислим дифференциалы от этих функций и запишем их в базисе  $dx, dy, dz$ , а координаты для удобства выразим через  $r, \theta$  и  $\phi$ . Тогда получим

$$\begin{aligned} dr &= \sin \theta \cos \phi dx + \sin \theta \sin \phi dy + \cos \theta dz, \\ d\theta &= \frac{1}{r} (\cos \theta \cos \phi dx + \cos \theta \sin \phi dy - \sin \theta dz), \\ d\phi &= \frac{1}{r \sin \theta} (-\sin \phi dx + \cos \phi dy). \end{aligned}$$

Прямое вычисление с использованием ортонормированности  $dx, dy, dz$  показывает, что в каждой точке  $(r, \theta, \phi)$  пространства  $\mathbb{R}^{3*}$   $dr, d\theta$  и  $d\phi$  являются ортогональными элементами, и что

$$(dr, dr) = 1, \quad (d\theta, d\theta) = \frac{1}{r^2}, \quad (d\phi, d\phi) = \frac{1}{r^2 \sin^2 \theta}.$$

Чтобы подвести итоги этих вычислений, заметим, что  $dr, r d\theta, r \sin \theta d\phi$  образуют ортонормированный базис пространства  $\Lambda^1(\mathbb{R}^3)$ , следовательно, эти три дифференциала можно вычислять так же, как мы это делали с  $dx, dy$  и  $dz$ . В результате получаем

$$\begin{aligned} \star dr &= r^2 \sin \theta d\theta \wedge d\phi, \\ \star(r d\theta) &= -r \sin \theta dr \wedge d\phi, \\ \star(r \sin \theta d\phi) &= r dr \wedge d\theta. \end{aligned}$$

Воспользовавшись линейностью, мы можем затем вычислить  $\star d\theta$  или  $\star d\phi$ .

## 16.7. Формулы Грина

В этом параграфе мы распространим формулы Грина, выведенные в главе 13, на непрерывный случай. Пусть  $U$  обозначает ограниченную область в пространстве  $\mathbb{R}^3$  с границей  $\partial U$ , и две линейные дифференциальные формы

$$E = E_x dx + E_y dy + E_z dz \quad \text{и} \quad F = F_x dx + F_y dy + F_z dz$$

определены в  $U$ . Их *скалярное произведение*  $(E, F)_U$  вычисляется по формуле

$$(E, F)_U = \int_U E \wedge \star F = \int_U (E_x F_x + E_y F_y + E_z F_z) dx dy dz = (F, E)_U.$$

Соответствующая *электростатическая энергия* равна  $\frac{1}{2} \|E\|_U^2$ , где

$$\|E\|_U^2 = (E, E)_U = \int_U (E_x^2 + E_y^2 + E_z^2) dx dy dz.$$

Предположим, что

$$E = -du \quad \text{и} \quad F = -dv,$$

где  $u$  и  $v$  — гладкие функции. Тогда

$$\begin{aligned} d(u \star dv) &= du \wedge \star dv + u d\star dv \\ &= du \wedge \star dv + u \Delta v dx \wedge dy \wedge dz \end{aligned}$$

или

$$du \wedge \star dv = d(u \star dv) - (u \Delta v) dx \wedge dy \wedge dz.$$

Соответственно

$$(du, dv)_U = \int_U du \wedge \star dv = \int_U d(u \star dv) - \int_U u \Delta v dx \wedge dy \wedge dz$$

и тогда

$$\boxed{(du, dv)_U = \int_{\partial U} u \star dv - \int_U u \Delta v dx \wedge dy \wedge dz.} \quad (16.6)$$

Это *первая формула Грина*. Поскольку  $(E, F)_U = (F, E)_U$ , в первой формуле Грина  $u$  и  $v$  можно переставить, а потом из одной формулы вычесть другую. Тогда получим

$$\boxed{\int_{\partial U} (u \star dv - v \star du) = \int_U (u \Delta v - v \Delta u) dx \wedge dy \wedge dz.} \quad (16.7)$$

Это *вторая формула Грина*. Пусть точка  $P$  лежит в области  $U$ , и дан маленький шарик  $B_\varepsilon$  с радиусом  $\varepsilon$  и с центром в  $P$ , который целиком находится в области  $U$  (рис. 16.12). Применим вторую формулу Грина к области  $U$ , из которой вынут шарик  $B_\varepsilon$ . Граница этой области будет состоять из  $\partial U$  и  $\partial B_\varepsilon$ , с противоположной ориентацией. Возьмем

$$v = r_P^{-1},$$

тогда

$$dv = \frac{dr}{r_P^2}, \quad \star dv = -\tau_P$$

и

$$\Delta v = 0 \quad \text{для} \quad U - B_\varepsilon.$$

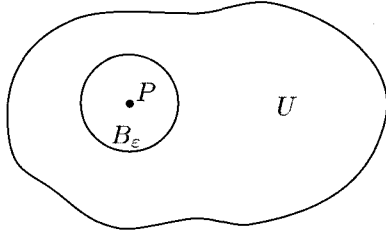


Рис. 16.12

На поверхности  $\partial B_\varepsilon$  функция  $v$  принимает постоянное значение  $\varepsilon^{-1}$ . Тогда мы получаем

$$\begin{aligned} - \int_{\partial U} \left( u \tau_P + \frac{\star du}{r_P} \right) + \int_{\partial B_\varepsilon} u \tau_P + \varepsilon^{-1} \int_{\partial B_\varepsilon} \star du \\ = - \int_{U - B_\varepsilon} \frac{\Delta u}{r_P} dx \wedge dy \wedge dz. \end{aligned}$$

Согласно теореме Стокса,

$$\int_{\partial B_\varepsilon} \star du = \int_{\partial B_\varepsilon} d\star du,$$

так что этот интеграл порядка  $\varepsilon^3$ . Следовательно, третий член слева стремится к нулю как  $O(\varepsilon^2)$ . Предполагается, что функция  $u$  дифференцируема, и поэтому  $u(x) = u(P) + O(\varepsilon)$  на границе  $\partial B_\varepsilon$ . Тогда мы можем заменить

$$\int_{\partial B_\varepsilon} u \tau_P \quad \text{на} \quad u(P) \int_{\partial B_\varepsilon} \tau_P,$$

с погрешностью порядка  $O(\varepsilon)$ . Однако

$$\int_{\partial B_\varepsilon} \tau_P = 4\pi,$$

т. е. равен площади сферы единичного радиуса. Устремив  $\varepsilon \rightarrow 0$ , мы получаем

$$u(P) = \frac{1}{4\pi} \int_{\partial U} \left( u \tau_P + \frac{\star du}{r_P} \right) - \frac{1}{4\pi} \int_U \frac{\Delta u}{r_P} dx dy dz. \quad (16.8)$$

## 16.8. Гармонические функции

Функция  $u$  называется *гармонической*, если  $\Delta u = 0$ . Для гармонических функций последний член в уравнении (16.8) обращается в нуль и мы имеем

$$u(P) = \frac{1}{4\pi} \int_{\partial U} \left( u \tau_P + \frac{\star du}{r_P} \right). \quad (16.9)$$

Это показывает, что гармоническая функция  $u$  внутри области определяется ее значением и значениями ее производных на границе. На самом деле мы докажем, что функция  $u$  определяется только своими значениями на границе  $\partial U$ . Давайте применим полученную формулу к шару  $W$  с центром в точке  $P$  (лежащим внутри  $U$ , см. рис. 16.13):

$$u(P) = \frac{1}{4\pi} \int_{\partial W} u \tau_P + \int_{\partial W} \frac{\star du}{r_P}.$$

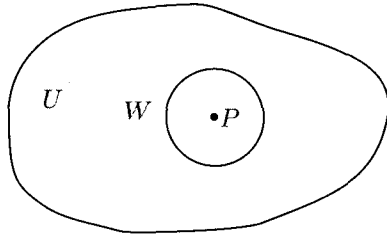


Рис. 16.13

Величина  $r_P$  постоянна на  $\partial W$ , и по теореме Стокса

$$\int_{\partial W} \star du = \int_W d\star du = 0, \quad \text{поскольку} \quad \Delta u = 0.$$

Тогда

$$u(P) = \frac{1}{4\pi} \int_{\partial W} u \tau_P. \quad (16.10)$$

Выражение справа — это просто среднее значение  $u$  по сфере  $\partial W$ . Итак, мы доказали следующее.

Если  $u$  — гармоническая функция в некоторой области, то значение  $u$  в любой точке равно ее среднему значению по любой сфере с центром в этой точке, внутренняя область которой целиком находится внутри области.

Отсюда можно получить несколько интересных следствий. Предположим, что вокруг точки  $\mathbf{x}_0$  имеется окрестность  $Z$ , такая, что

$$u(\mathbf{x}) \leq u(\mathbf{x}_0) \quad \text{во всех точках внутри } Z.$$

Пусть  $S_a$  — сфера с радиусом  $a$ , целиком лежащая в области  $Z$  и имеющая центр в точке  $\mathbf{x}_0$ . Тогда  $u(\mathbf{x}) \leq u(\mathbf{x}_0)$  во всех точках  $S_a$ . А среднее значение  $u$  по сфере  $S_a$  меньше или равно  $u(\mathbf{x}_0)$ . Предположим теперь, что есть некоторая точка  $\mathbf{y}$  из  $S_a$ , где  $u(\mathbf{y}) < u(\mathbf{x}_0)$ . Тогда  $u(\mathbf{x}) < u(\mathbf{x}_0)$  во всех точках  $\mathbf{x}$  вблизи  $\mathbf{y}$ , и поэтому среднее от  $u$  по сфере  $S_a$  будет строго меньше  $u(\mathbf{x}_0)$ . Но это невозможно. Тогда

$$u(\mathbf{y}) = u(\mathbf{x}_0) \quad \text{во всех точках сферы } S_a.$$

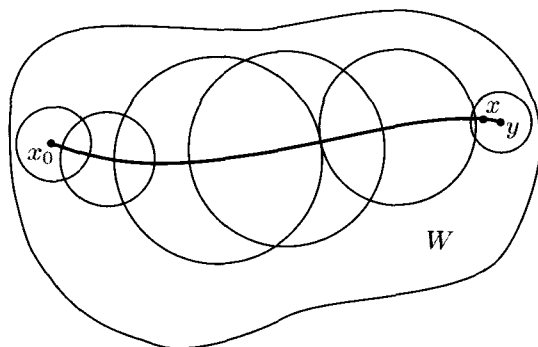


Рис. 16.14

Теперь предположим, что  $W$  — открытое и *связное* множество, т. е. допустим, что любые две точки области  $W$  могут быть соединены непрерывной кривой, целиком лежащей в  $W$ . Допустим также, что функция  $u$  достигает своего максимума в некоторой точке  $\mathbf{x}_0 \in W$ . Пусть  $\mathbf{y}$  — любая другая точка  $W$  и пусть  $C$  — кривая, соединяющая точки  $\mathbf{x}_0$  и  $\mathbf{y}$ . Вокруг каждой точки  $\mathbf{x}$  на кривой мы можем построить достаточно малый шарик с центром в точке  $\mathbf{x}$ , целиком лежащий в  $W$ . В силу компактности  $C$  мы можем выбрать конечное число таких шариков, покрывающих  $C$ . Теперь можно утверждать следующее. Имеется конечное число сфер  $S_{a_1}, \dots, S_{a_k}$  таких, что внутренняя часть всех сфер лежит целиком в  $W$ . Сфера  $S_{a_1}$  имеет центр в точке  $\mathbf{x}_0$ , центр  $\mathbf{x}_i$  сферы  $S_{a_{i+1}}$  лежит в  $S_{a_i}$ , а  $\mathbf{y} \in S_{a_k}$  (см. рис. 16.14). Но это значит, что  $u(\mathbf{x}_0) = u(\mathbf{x}_1) = \dots = u(\mathbf{x}_k) = u(\mathbf{y})$ . Другими словами, мы установили следующее.

Пусть  $u$  — гармоническая функция в открытом связном множестве  $W$ . Предположим, что она достигает максимума в некоторой точке  $\mathbf{x}_0 \in W$ . Тогда  $u$  постоянна на множестве  $W$ .

Отсюда непосредственно следует вывод<sup>3</sup>.

<sup>3</sup>называемый (в том числе далее в этой главе) *принципом максимума* для гармонических функций. Вышедоказанная теорема о среднем и соответственно принцип максимума справедливы для гармонических функций не только в  $\mathbb{R}^3$ , но и в пространстве  $\mathbb{R}^n$  любой размерности  $n \geq 2$ . — *Прим. ред.*

Пусть  $U$  — открытое связное множество и пусть  $\bar{U}$  обозначает его замыкание, так что  $\bar{U} = U \cup \partial U$ , где  $\partial U$  обозначает границу  $U$ . Предположим, что  $\bar{U}$  ограничено. Тогда если  $u$  — функция, гармоническая в  $U$  и непрерывная на  $\bar{U}$ , то либо

$$u(\mathbf{x}) < \max_{\mathbf{y} \in \partial U} u(\mathbf{y}) \quad \text{для любого } \mathbf{x} \in U,$$

либо  $u$  постоянна.

*Доказательство.* Поскольку  $\bar{U}$  замкнуто и ограничено, то оно компактно. Функция  $u$  непрерывна на  $\bar{U}$ . Теорема из вещественного анализа говорит, что  $u$  должна достигать своего максимума в некоторой точке  $\mathbf{x}_0$  множества  $\bar{U}$ . Если одной из таких точек будет  $\mathbf{x}_0 \in U$ , то из наших предыдущих результатов следует, что  $u$  постоянна. Если же  $u$  не постоянна, то  $\mathbf{x}_0 \in \partial U$ , и мы доказали наше предположение.

Отсюда можно получить еще один результат.

Пусть  $U$  — открытое связное множество, а его замыкание  $\bar{U}$  компактно. Пусть  $u$  и  $v$  — функции, непрерывные на  $\bar{U}$  и гармонические в  $U$ . Допустим, что

$$u(\mathbf{y}) = v(\mathbf{y}) \quad \text{для всех } \mathbf{y} \in \partial U.$$

Тогда

$$u(\mathbf{x}) = v(\mathbf{x}) \quad \text{для всех } \mathbf{x} \in U.$$

*Доказательство.* Функция  $u - v$  гармоническая и обращается в нуль на границе  $\partial U$ . В силу предыдущего принципа  $u(\mathbf{x}) - v(\mathbf{x}) \leq 0$  для  $\mathbf{x} \in U$ . Аналогично,  $v(\mathbf{x}) - u(\mathbf{x}) \leq 0$ , что и требовалось доказать.

Другая формулировка последнего предложения: в области  $U$  гармоническая функция полностью определяется своими значениями на границе. В этом состоит *теорема единственности*: существует только одна гармоническая функция, определяемая своими граничными значениями. При этом возникает проблема: верна ли соответствующая теорема существования? Эта проблема известна как *задача Дирихле*.

## 16.9. Метод ортогональной проекции

Давайте рассмотрим другое доказательство *единственности* решения задачи Дирихле для ограниченной области  $U$ . Воспользуемся первой формулой Грина (16.6)

$$(du, dv)_U = \int_{\partial U} u \star dv - \int_U u \Delta v \, dx \wedge dy \wedge dz.$$

Пусть  $C^0$  обозначает пространство всех дифференцируемых функций на  $\bar{U}$ , а  $C^0_{\text{внутр}}$  — его подпространство, состоящее из функций, тождественно равных нулю на границе  $\partial U$ . Пусть  $H$  обозначает пространство гармонических функций  $v$  в области  $U$ , которые непрерывны на  $\bar{U}$  и удовлетворяют уравнению  $\Delta v = 0$  в области  $U$ . Тогда, если  $u \in C^0_{\text{внутр}}$ , то поскольку  $u = 0$  на границе  $\partial U$ , мы имеем

$$\int_{\partial U} u \star dv = 0 \quad \text{для любой гладкой функции } v.$$

Если  $v \in H$ , то  $u \Delta v \equiv 0$  для любой функции  $u$ . Тогда если  $u \in C^0_{\text{внутр}}$  и  $v \in H$ , то  $(du, dv)_U = 0$ . Другими словами, пространства  $dC^0_{\text{внутр}}$  и  $dH$  ортогональны относительно  $(\cdot, \cdot)_U$ . В частности, предположим, что  $u$  — гармоническая функция, обращающаяся в нуль на границе. Тогда

$$(du, du)_U = 0.$$

Однако

$$(du, du)_U = \int_U \left( \left( \frac{\partial u}{\partial x} \right)^2 + \left( \frac{\partial u}{\partial y} \right)^2 + \left( \frac{\partial u}{\partial z} \right)^2 \right) dx \wedge dy \wedge dz.$$

Это выражение обращается в нуль только при условии, что все частные производные  $u$  равны нулю в области  $U$ . Но тогда  $u$  должна быть постоянной, и поскольку  $u = 0$  на границе  $\partial U$ , то  $u \equiv 0$ . Таким образом, гармоническая функция, которая равна нулю на границе  $\partial U$ , должна быть равна нулю везде. Как и раньше, это доказывает *единственность* решения задачи Дирихле.

Пространства  $dC^0$ ,  $dH$  и так далее являются бесконечномерными векторными пространствами. В этой книге мы не занимались такими пространствами или проекционными операторами  $\pi$  в таких пространствах. Эта теория может быть разработана и, следуя Герману Вейлю, может быть доказан следующий принцип.

Пространство  $dH$  является ортогональным дополнением  $dC^0_{\text{внутр}}$  к  $dC^0$ .

Или другими словами, если  $E = dv$  и  $(E, du) = 0$  для всех  $u \in C^0_{\text{внутр}}$ , то  $\Delta v = 0$  в области  $U$ . Поэтому, взяв любую функцию  $\psi$ , мы можем разложить  $d\psi$  в сумму

$$d\psi = du + dv, \quad \text{где } u \in C^0_{\text{внутр}} \text{ и } v \in H.$$

Другими словами,  $du = \pi d\psi$ , где  $\pi$  обозначает проекцию на компоненту  $dC^0_{\text{внутр}}$ , и  $dv = (1 - \pi) d\psi$ . Теперь можно решить задачу Дирихле: для данной гладкой функции  $\phi$ , определенной на  $\partial U$ , найти такую функцию  $v$ , чтобы выполнялось уравнение  $\Delta v = 0$  и  $v = \phi$  на границе  $\partial U$ .

*Решение.* Выберем произвольную функцию  $\psi$ , которая совпадает с  $\phi$  на границе  $\partial U$  и определена во всей области  $U$ . Разложим  $d\psi$  на слагаемые, как это показано выше. Тогда поскольку  $u$  обращается в нуль на границе  $\partial U$ , мы знаем, что  $v$  совпадает с  $\psi$  на границе  $\partial U$  и, следовательно, является искомым решением.

Заметим, что поскольку  $(du, dv)_U = 0$ , то

$$\|d\psi\|_U^2 = \|du\|_U^2 + \|dv\|_U^2,$$

откуда следует, что

$$\|d\psi\|_U^2 \geq \|dv\|_U^2.$$

Таким образом, среди всех функций, значения которых совпадают с  $\psi$  на границе, решение задачи Дирихле  $v$  может характеризоваться как функция с наименьшим значением  $\|dv\|_U^2$ . Это так называемый принцип Дирихле. Если считать, что  $\frac{1}{2}\|E\|_U^2$  — энергия электрического поля, то принцип Дирихле говорит, что решением задачи Дирихле является такая функция  $v$ , у которой  $dv$  дает минимальную энергию, при условии, что на границе  $\partial U$  выполняется равенство  $v = \phi$ .

Мы не доказываем здесь, что решение задачи Дирихле существует для произвольных областей (доказательство можно найти в главе 12 книги Loomis and Sternberg). Тем не менее, мы покажем, как решается эта задача для областей определенного типа. Кроме того, мы выведем некоторые общие принципы электростатики.

## 16.10. Функции Грина

Предположим, что в области  $U$  задача Дирихле может быть решена. Здесь мы получим для этой области  $U$  аналог той функции Грина, которая обсуждалась в главе 13 в дискретном случае.

Для любой точки  $\mathbf{x}$  в пространстве  $\mathbb{R}^3$  положим

$$K(\mathbf{x}, \mathbf{y}) = \frac{1}{4\pi} \frac{1}{\|\mathbf{x} - \mathbf{y}\|}.$$

Рассматривая  $K$  как функцию  $\mathbf{y}$  (при фиксированном  $\mathbf{x}$ ), мы имеем

$$\star dK(\mathbf{x}, \cdot) = \frac{1}{4\pi} \tau_{\mathbf{x}}.$$

Формулу (16.8) из параграфа 16.7 мы можем переписать через  $K$  в виде

$$u(\mathbf{x}) = \int_{\partial U} (u \star dK(\mathbf{x}, \cdot) + K(\mathbf{x}, \cdot) \star du) - \int_U K(\mathbf{x}, \cdot) \Delta u \, dV.$$

Теперь для заданного  $\mathbf{x} \in U$  функция  $K(\mathbf{x}, \mathbf{y})$  является дифференцируемой функцией  $\mathbf{y}$ , когда  $\mathbf{y}$  изменяется вдоль границы  $\partial U$ . По предположению мы можем найти такую функцию  $h(\mathbf{x}, \cdot)$ , гармоническую в области  $U$  и непрерывную на  $\bar{U}$ , чтобы выполнялось условие

$$h(\mathbf{x}, \mathbf{y}) = -K(\mathbf{x}, \mathbf{y}) \quad \text{для всех } \mathbf{y} \in \partial U.$$

Далее, для фиксированного значения  $\mathbf{y}$  функция  $h(\mathbf{x}, \mathbf{y})$  является непрерывной функцией  $\mathbf{x}$ . В соответствии с принципом максимума гармоническая функция  $h(\mathbf{x}_1, \mathbf{y}) - h(\mathbf{x}_2, \mathbf{y})$  достигает своего максимального значения на границе  $\partial U$ . Следовательно,

$$|h(\mathbf{x}_1, \mathbf{y}) - h(\mathbf{x}_2, \mathbf{y})| \leq \max_{\mathbf{z} \in \partial U} |K(\mathbf{x}_1, \mathbf{z}) - K(\mathbf{x}_2, \mathbf{z})|.$$

Также очевидно, что  $K(\mathbf{x}, \mathbf{z})$  равномерно непрерывна по  $\mathbf{x}$  для всех  $\mathbf{z} \in \partial U$ , пока  $\mathbf{x}$  находится на ограниченном снизу расстоянии от  $\partial U$ . Итак, мы построили такую функцию  $h_U$ , для которой выполняются следующие условия:

- (i) функция  $h_U(\mathbf{x}, \mathbf{y})$  непрерывна по  $\mathbf{x}$  и  $\mathbf{y}$  для  $\mathbf{x}, \mathbf{y} \in \bar{U}$  и дифференцируема по  $\mathbf{y}$  в области  $U$ ;
- (ii) для каждого фиксированного значения  $\mathbf{x}$  выполняется уравнение  $\Delta_{\mathbf{y}} h_U = 0$ , т. е.  $\Delta h_U(\mathbf{x}, \cdot) = 0$ ;
- (iii)  $G_U(\mathbf{x}, \mathbf{y}) = K(\mathbf{x}, \mathbf{y}) + h_U(\mathbf{x}, \mathbf{y}) = 0$  для  $\mathbf{y} \in \partial U$ .

Функция  $G_U$  называется *функцией Грина* области  $U$ . Предположим, что  $G_U$  существует, и давайте выясним ее свойства. В дальнейшем мы будем писать  $G$  вместо  $G_U$ , если ясно, о какой области  $U$  идет речь. Сначала покажем, что для  $\mathbf{x} \in U$  и  $\mathbf{y} \in U$  выполняется условие

$$G(\mathbf{x}, \mathbf{y}) = G(\mathbf{y}, \mathbf{x}).$$

Пусть  $B_{\mathbf{x}, \varepsilon}$  и  $B_{\mathbf{y}, \varepsilon}$  обозначают маленькие шарики вокруг точек  $\mathbf{x}$  и  $\mathbf{y}$ . Применим вторую формулу Грина к функциям  $u = G(\mathbf{x}, \cdot)$  и  $v = G(\mathbf{y}, \cdot)$ , заменив область  $U$  на  $U - B_{\mathbf{x}, \varepsilon} - B_{\mathbf{y}, \varepsilon}$ . Поскольку обе функции гармонические в рассматриваемой области и обращаются в нуль на границе  $\partial U$ , мы получаем

$$\begin{aligned} \int_{\partial B_{\mathbf{x}, \varepsilon}} G(\mathbf{x}, \cdot) \star dG(\mathbf{y}, \cdot) + \int_{\partial B_{\mathbf{y}, \varepsilon}} G(\mathbf{x}, \cdot) \star dG(\mathbf{y}, \cdot) \\ = \int_{\partial B_{\mathbf{x}, \varepsilon}} G(\mathbf{y}, \cdot) \star dG(\mathbf{x}, \cdot) + \int_{\partial B_{\mathbf{y}, \varepsilon}} G(\mathbf{y}, \cdot) \star dG(\mathbf{x}, \cdot). \end{aligned}$$

Теперь покажем, что при  $\varepsilon \rightarrow 0$  левая часть равенства стремится к  $G(\mathbf{x}, \mathbf{y})$ , а правая — к  $G(\mathbf{y}, \mathbf{x})$ . В силу симметрии достаточно установить это для левой части.

Запишем

$$\int_{\partial B_{\mathbf{y}, \varepsilon}} G(\mathbf{x}, \cdot) \star dG(\mathbf{y}, \cdot) = \int_{\partial B_{\mathbf{y}, \varepsilon}} G(\mathbf{x}, \cdot) \star dK(\mathbf{y}, \cdot) + \int_{\partial B_{\mathbf{y}, \varepsilon}} G(\mathbf{x}, \cdot) \star dh.$$

Первый член справа стремится к  $G(\mathbf{x}, \mathbf{y})$ , как и в параграфе 16.6, поскольку  $4\pi \star dK$  — это телесный угол из точки  $\mathbf{y}$ . Второй член стремится к нулю, потому что  $G(\mathbf{x}, \cdot)$  и  $h$  — гладкие функции в области  $B_{\mathbf{y}, \varepsilon}$ . С другой стороны,

$$\begin{aligned} \int_{\partial B_{\mathbf{x}, \varepsilon}} G(\mathbf{x}, \cdot) \star dG(\mathbf{y}, \cdot) \\ = \int_{\partial B_{\mathbf{x}, \varepsilon}} K(\mathbf{x}, \cdot) \star dG(\mathbf{y}, \cdot) + \int_{\partial B_{\mathbf{x}, \varepsilon}} h(\mathbf{x}, \cdot) \star dG(\mathbf{y}, \cdot). \end{aligned}$$

Так же, как и выше, второй член стремится к нулю. Первый член можно переписать в виде

$$\frac{1}{4\pi} \int_{\partial B_{\mathbf{x}, \varepsilon}} \star dG(\mathbf{y}, \cdot) = \frac{1}{4\pi} \int_{B_{\mathbf{x}, \varepsilon}} d \star dG(\mathbf{y}, \cdot) = 0,$$

поскольку  $G(\mathbf{y}, \cdot)$  — гармоническая функция в области  $B_{\mathbf{x}, \varepsilon}$ . Итак, мы доказали, что  $G(\mathbf{x}, \mathbf{y}) = G(\mathbf{y}, \mathbf{x})$ .

Пусть  $u$  обозначает любую гладкую функцию в области  $U$ . Применим вторую формулу Грина к  $u$  и  $v = G(\mathbf{x}, \cdot)$  в области  $U - B_{\mathbf{x}, \varepsilon}$ . Поскольку  $G(\mathbf{x}, \cdot) = 0$  на границе  $\partial U$ , мы получим

$$\begin{aligned} \int_{\partial U} u \star dG(\mathbf{x}, \cdot) - \int_{\partial B_{\mathbf{x}, \varepsilon}} u \star dG(\mathbf{x}, \cdot) + \int_{\partial B_{\mathbf{x}, \varepsilon}} G(\mathbf{x}, \cdot) \star du \\ = \int_U G(\mathbf{x}, \cdot) \Delta u \, d\mathbf{x}^1 \wedge \cdots \wedge d\mathbf{x}^n. \end{aligned}$$

Третий интеграл слева можно записать в виде

$$\begin{aligned} \int_{\partial B_{\mathbf{x}, \varepsilon}} K(\mathbf{x}, \cdot) \star du + \int_{\partial B_{\mathbf{x}, \varepsilon}} h(\mathbf{x}, \cdot) \star du \\ = \frac{1}{4\pi\varepsilon} \int_{\partial B_{\mathbf{x}, \varepsilon}} \star du + \int_{\partial B_{\mathbf{x}, \varepsilon}} h(\mathbf{x}, \cdot) \star du = \frac{1}{4\pi\varepsilon} O(\varepsilon^2) + O(\varepsilon^2). \end{aligned}$$

Следовательно, он стремится к нулю. Тогда мы получим

$$u(\mathbf{x}) = \int_U G(\mathbf{x}, \cdot) \Delta u \, d\mathbf{x}^1 \wedge \cdots \wedge d\mathbf{x}^n + \int_{\partial U} u \star dG(\mathbf{x}, \cdot).$$

Это уравнение показывает, что если мы знаем, что существует решение задачи Пуассона  $\Delta F = f$  с условием  $F = 0$  на границе, то оно имеет вид

$$F(\mathbf{x}) = \int G(\mathbf{x}, \mathbf{y}) f(\mathbf{y}) d\mathbf{y}. \quad (16.11)$$

Аналогично, если мы знаем, что существует гладкое решение задачи Дирихле

$$\Delta u = 0, \quad u(\mathbf{x}) = f(\mathbf{x}) \quad \text{для } \mathbf{x} \in \partial U,$$

то оно имеет вид

$$u(\mathbf{x}) = \int_{\partial U} u \star dG(\mathbf{x}, \cdot). \quad (16.12)$$

## 16.11. Интегральная формула Пуассона

Важно отметить, что эти формулы являются следствием существования функции Грина для  $U$ , т.е. они справедливы во всех случаях, когда мы можем найти такую функцию  $h$ , что (ii) и (iii) выполняются. В качестве примера мы можем в явном виде построить функцию Грина для шара  $B_R$  с радиусом  $R$ . Для простоты предположим, что начало координат находится в центре шара.

Пусть для  $\mathbf{x} \neq 0$  точка  $\mathbf{x}'$  обозначает ее образ при инверсии по отношению к сфере радиуса  $R$ , т.е.

$$\mathbf{x}' = \frac{R^2}{\|\mathbf{x}\|^2} \mathbf{x} \quad (\text{см. рис. 16.11}).$$

Определим функцию  $G_R$  формулой

$$4\pi G_R(\mathbf{x}, \mathbf{y}) = \begin{cases} \frac{1}{\|\mathbf{y} - \mathbf{x}\|} - \frac{R}{\|\mathbf{x}\| \|\mathbf{y} - \mathbf{x}'\|}, & \text{если } \mathbf{x} \neq 0, \\ \frac{1}{\|\mathbf{y}\|} - \frac{1}{R}, & \text{если } \mathbf{x} = 0. \end{cases}$$

Если  $\mathbf{x} \in B_R$ , то  $\mathbf{x}' \in B_R$  и поэтому вторые слагаемые в правой части формулы являются непрерывными и гармоническими в  $\overline{B_R}$ . Нам надо просто проверить выполнение свойства (iii). Для  $\|\mathbf{y}\| = R$  из подобных треугольников (или прямым вычислением) мы имеем

$$\frac{R}{\|\mathbf{x}\|} = \frac{\|\mathbf{y} - \mathbf{x}'\|}{\|\mathbf{y} - \mathbf{x}\|},$$

следовательно,

$$G_R(\mathbf{x}, \mathbf{y}) = 0 \quad \text{для} \quad \|\mathbf{y}\| = R.$$

А это и есть свойство (iii). Итак мы проверили, что  $G_R$  является функцией Грина для шара с радиусом  $R$ .

Чтобы применить нашу формулу к решению задачи Дирихле с помощью функции Грина, мы должны вычислить  $\star dG_R$  на сфере радиуса  $R$ . Имеем

$$\begin{aligned} 4\pi \star dG_R(\mathbf{x}, \cdot) &= \tau_{\mathbf{x}} - \frac{R}{\|\mathbf{x}\|} \tau_{\mathbf{x}'} \\ &= \sum \frac{\mathbf{y}^i - \mathbf{x}^i}{\|\mathbf{y} - \mathbf{x}\|^3} - \frac{R}{\|\mathbf{x}\|} \frac{\mathbf{y}^i - \mathbf{x}^i}{\|\mathbf{y} - \mathbf{x}'\|^3} \star d\mathbf{y}^i. \end{aligned}$$

Но

$$\frac{R}{\|\mathbf{x}\| \|\mathbf{y} - \mathbf{x}'\|^3} = \frac{\|\mathbf{x}\|^2}{R^2 \|\mathbf{y} - \mathbf{x}\|^3}, \quad \text{если} \quad \|\mathbf{y}\| = R.$$

Итак, мы видим, что для  $\|\mathbf{y}\| = R$

$$\begin{aligned} 4\pi \star dG_R &= \frac{1}{\|\mathbf{y} - \mathbf{x}\|^3} \sum \left( \mathbf{y}^i - \mathbf{x}^i - \frac{\|\mathbf{x}\|^2}{R^2} \mathbf{y}^i - \frac{R^2}{\|\mathbf{x}\|^2} \mathbf{x}^i \right) \star d\mathbf{y}^i \\ &= \frac{R^2 - \|\mathbf{x}\|^2}{\|\mathbf{y} - \mathbf{x}\|^3 R^2} \sum \mathbf{y}^i \star d\mathbf{y}^i \\ &= \frac{R^2 - \|\mathbf{x}\|^2}{\|\mathbf{y} - \mathbf{x}\|^3 R^2} \star r dr. \end{aligned}$$

Если мы подставим в наше решение задачи Дирихле

$$u(\mathbf{x}) = \int_{\partial U} u(\mathbf{y}) \star dG(\mathbf{x}, \mathbf{y}),$$

то получим

$$4\pi u(\mathbf{x}) = \int_{S_R} \frac{R^2 - \|\mathbf{x}\|^2}{\|\mathbf{y} - \mathbf{x}\|^3 R^2} u(\mathbf{y}) \star r \, dr.$$

Однако на сфере радиуса  $R$  форма  $\star r \, dr$  просто равна  $R \, dS_R$ , где  $dS_R$  — элемент площади сферы радиуса  $R$ . Подставив это в последнюю формулу, получаем *интегральную формулу Пуассона*:

$$u(\mathbf{x}) = \frac{R^2 - \|\mathbf{x}\|^2}{4\pi R} \int_{S_R} \frac{u(\mathbf{y})}{\|\mathbf{y} - \mathbf{x}\|^3} dS_R \quad (16.13)$$

При выводе формулы (16.13) мы воспользовались предположением, что функция  $u$  дифференцируема в некоторой окрестности шара  $B_R$ , и что она гармоническая для  $\|\mathbf{x}\| < R$ . На самом деле нам было достаточно предположить, что  $u$  — дифференцируемая и гармоническая при  $\|\mathbf{x}\| < R$  и что она непрерывна на замкнутом шаре  $\|\mathbf{x}\| \leq R$ . По существу, формула (16.13) для любого  $\|\mathbf{x}\| < R$  останется верной, если заменить  $R$  на  $R_\alpha$ , где  $\mathbf{x} < R_\alpha < R$ . Если затем  $R_\alpha$  устремить к  $R$ , то формула (16.13) восстанавливается в первоначальном виде в силу предположения о непрерывности функции  $u$ .

Формула (16.13) дает решение задачи Дирихле для шара при условии, что мы знаем, что такое решение существует, т. е. если  $u$  — произвольная гармоническая функция на открытом шаре и непрерывная на замкнутом, то она удовлетворяет уравнению (16.13). А сейчас покажем, что (16.13) действительно является решением задачи Дирихле для заданных граничных значений. Предположим, что нам дана непрерывная функция  $u$ , определенная на сфере  $S_R$ . Тогда нам дана  $u(\mathbf{y})$  для всех  $\mathbf{y} \in S_R$ . Пусть уравнение (16.13) определяет  $u(\mathbf{x})$  для  $\|\mathbf{x}\| < R$ . Мы должны показать, что

- (а)  $u$  гармоническая в области  $\|\mathbf{x}\| < R$ ,
- (б)  $u(\mathbf{x}) \rightarrow u(\mathbf{y}_0)$ , если  $\mathbf{x} \rightarrow \mathbf{y}_0$  и  $\|\mathbf{y}_0\| = R$ .

Чтобы доказать (а), заметим, что  $G_R(\mathbf{x}, \mathbf{y})$  — дифференцируемая функция  $\mathbf{x}$  и  $\mathbf{y}$  в области  $\|\mathbf{x}\| < R_1 < R$ ,  $R_1 < \|\mathbf{y}\| < R^2/R_1$ .

Кроме того, по построению она — гармоническая функция от  $\mathbf{y}$ . Для  $\|\mathbf{x}\| < R$  и  $\|\mathbf{y}\| < R$  мы знаем, что

$$G_R(\mathbf{x}, \mathbf{y}) = G_R(\mathbf{y}, \mathbf{x}).$$

Тогда для фиксированного  $\mathbf{y}$ , при условии  $\|\mathbf{y}\| < R$ , функция  $G_R(\cdot, \mathbf{y})$  гармоническая в области  $B_R - \{\mathbf{y}\}$ . Устремив  $\|\mathbf{y}\| \rightarrow R$ , мы видим, что  $G_R(\mathbf{x}, \mathbf{y})$  — гармоническая функция  $\mathbf{x}$  для  $\|\mathbf{x}\| < R_1 < R$  и любого фиксированного  $\mathbf{y} \in S_R$ . Таким образом,  $\partial G_R(\mathbf{x}, \mathbf{y}) / \partial y^i$  — гармоническая функция  $\mathbf{x}$  для любого  $\mathbf{y} \in S_R$ . Другими словами, все коэффициенты  $\star dG_R(\cdot, \mathbf{y})$  являются гармоническими функциями  $\mathbf{x}$  для каждого  $\mathbf{y} \in S_R$ . Это же можно сказать и о всех коэффициентах  $u(\mathbf{y}) \star dG_R(\cdot, \mathbf{y})$ . Следовательно, функция  $u(\mathbf{x}) = \int_{S_R} u \star dG_R(\mathbf{x}, \cdot)$  является гармонической функцией  $\mathbf{x}$ , поскольку интеграл сходится равномерно при  $\|\mathbf{x}\| < R_1 < R$  (так же, как и все интегралы от различных производных по  $\mathbf{x}$ ). Это доказывает (а).

Для доказательства (b) сначала отметим, что тождественная единица является гармонической функцией везде, поэтому к ней применимо уравнение (16.13). Тогда получаем

$$\frac{R^2 - \|\mathbf{x}\|^2}{4\pi R} \int_{S_R} \frac{dS_R}{\|\mathbf{y} - \mathbf{x}\|^3} = 1 \quad \text{для любого } \|\mathbf{x}\| < R. \quad (16.14)$$

Пусть  $\mathbf{y}_0$  обозначает точку на сфере  $S_R$  и пусть  $u$  — непрерывная функция на этой сфере. Для любого  $\varepsilon > 0$  мы можем найти такое  $\delta > 0$ , что

$$|u(\mathbf{y}) - u(\mathbf{y}_0)| < \varepsilon \quad \text{при } \|\mathbf{y} - \mathbf{y}_0\| \leq 2\delta, \quad \mathbf{y} \in S_R.$$

Пусть

$$Z_1 = \{\mathbf{y} \in S_R : \|\mathbf{y} - \mathbf{y}_0\| > 2\delta\} \quad \text{и} \quad Z_2 = \{\mathbf{y} \in S_R : \|\mathbf{y} - \mathbf{y}_0\| \leq 2\delta\}.$$

Тогда в силу (16.13) и (16.14) для  $\|\mathbf{x}\| < R$  мы имеем

$$u(\mathbf{x}) - u(\mathbf{y}_0) = \frac{R^2 - \|\mathbf{x}\|^2}{4\pi R} \int_{S_R} \frac{u(\mathbf{y}) - u(\mathbf{y}_0)}{\|\mathbf{y} - \mathbf{x}\|^3} dS_R.$$

Поэтому

$$|u(\mathbf{x}) - u(\mathbf{y}_0)| \leq I_1 + I_2,$$

где

$$I_1 = \frac{R^2 - \|\mathbf{x}\|^2}{4\pi R} \int_{Z_1} \frac{|u(\mathbf{y}) - u(\mathbf{y}_0)|}{\|\mathbf{y} - \mathbf{x}\|^3} dS_R$$

и

$$I_2 = \frac{R^2 - \|\mathbf{x}\|^2}{4\pi R} \int_{Z_2} \frac{|u(\mathbf{y}) - u(\mathbf{y}_0)|}{\|\mathbf{y} - \mathbf{x}\|^3} dS_R.$$

Если  $\|\mathbf{y}_0 - \mathbf{x}\| < \delta$ , то для всех  $\mathbf{y} \in Z_1$  мы имеем  $\|\mathbf{y} - \mathbf{x}\| > \|\mathbf{y} - \mathbf{y}_0\| - \|\mathbf{x} - \mathbf{y}_0\|$ , так что  $\|\mathbf{y} - \mathbf{x}\| > \delta$ . Итак, для всех  $\mathbf{x}$ , удовлетворяющих условию  $\|\mathbf{x} - \mathbf{y}_0\| < \delta$ , интеграл в  $I_1$  равномерно ограничен. Поскольку  $\|\mathbf{x}\| \rightarrow R$  при  $\mathbf{x} \rightarrow \mathbf{y}_0$ , мы приходим к выводу, что  $I_1 \rightarrow 0$  при  $\mathbf{x} \rightarrow \mathbf{y}_0$ .

Займемся теперь  $I_2$ . Мы знаем, что  $|u(\mathbf{y}) - u(\mathbf{y}_0)| < \varepsilon$  для всех  $\mathbf{y} \in Z_2$ , так что

$$I_2 < \frac{R^2 - \|\mathbf{x}\|^2}{4\pi R} \int_{Z_2} \frac{\varepsilon dS_R}{\|\mathbf{y} - \mathbf{x}\|^3} = \varepsilon \frac{R^2 - \|\mathbf{x}\|^2}{4\pi R} \int_{S_R} \frac{dS_R}{\|\mathbf{y} - \mathbf{x}\|^3} = \varepsilon$$

в силу (16.14). Таким образом, мы доказали (b).

Итак, доказана

**Теорема.** Пусть  $u$  — непрерывная функция, определенная на сфере  $S_R$ . Существует единственная непрерывная на ограниченном ею шаре  $\|\mathbf{x}\| \leq R$  функция, которая совпадает с данной функцией на сфере  $S_R$  и является гармонической в области  $\|\mathbf{x}\| < R$  и. Эта функция определяется для всех  $\|\mathbf{x}\| < R$  уравнением (16.13).

Мы затратили изрядное количество усилий на подробное доказательство существования решения задачи Дирихле в простом случае сферы. Детали доказательств для более сложных областей мы приводить не будем. Дадим только идею одного метода, подробности которого при желании можно восстановить.

Давайте аппроксимируем евклидово пространство сетью. В конце каждого вектора  $\mathbf{v}$ , координаты которого являются двоичными дробями со знаменателем  $2^N$ , поместим узел. Вектора  $\mathbf{v}$  имеют вид

$$\mathbf{v} = \frac{1}{2^N} \begin{pmatrix} k \\ l \\ m \end{pmatrix}, \quad \text{где } k, l, m — \text{целые числа.}$$

Здесь  $N$  — фиксированное большое целое число. Каждый узел соединим с шестью ближайшими соседями (рис. 16.15); тогда все пространство разделится на кубические ячейки. Каждой ветви присвоим емкость  $2^N$ , а через  $\Delta_N$  обозначим оператор Лапласа для этой сети. Легко видеть, что для любой дважды дифференцируемой функции  $u$

$$\Delta_N u \rightarrow \Delta u \quad \text{при } N \rightarrow \infty.$$

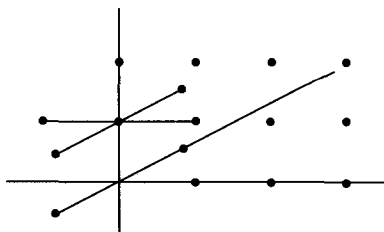


Рис. 16.15

Если нам дана ограниченная область  $U$  с границей  $\partial U$ , то можно рассмотреть конечную подсеть, состоящую из всех узлов в области  $U$  и всех ветвей, соединяющих эти узлы. Узлы, расположенные ближе всего к границе, назовем граничными узлами. Используя наши результаты для конечных цепей, мы можем решить задачу Дирихле для такого конечного приближения. Зададим непрерывную функцию  $\phi$  на границе  $\partial U$ . Можно приписать значения  $\phi_N$  на граничных узлах нашего конечного приближения, взяв  $\phi_N(\mathbf{n}) = \phi(\mathbf{x})$ , где  $\mathbf{x}$  — это ближайшая к узлу  $\mathbf{n}$  точка на границе. (Если таких ближайших точек несколько, то выбираем одну из них.) Таким образом, непрерывную задачу Дирихле мы аппроксимировали дискретной конечной задачей, которую умеем решать. При  $N \rightarrow \infty$  мы ожидаем, что последовательность решений  $U_N$  будет сходиться к функции  $u$  (определенной для всех двоично-рациональных векторов и продолженной по непрерывности), которая дает решение исходной задачи Дирихле. Этот метод доказательства работает, если форма границы не слишком «дикая». Само доказательство требует аккуратных оценок, что находится за рамками этой книги.

## Резюме

### А. Электростатика и дифференциальные формы

Вы должны уметь формулировать законы электростатики в терминах дифференциальных форм  $D$  и  $E$ .

Вы должны знать формулы Грина, уметь их доказывать и применять в задачах электростатики в пространстве  $\mathbb{R}^3$ .

### В. Звездный оператор в пространстве $\mathbb{R}^3$

Для данной дифференциальной формы  $\omega$  в пространстве  $\mathbb{R}^3$ , записанной в координатах, дифференциалы которых взаимно ортогональны, надо уметь строить ортогональный базис и с его помощью строить  $\star\omega$ .

Вы должны уметь понимать связь между электростатикой в пространстве  $\mathbb{R}^3$  и теорией цепей с емкостями, а также уметь объяснить, как звездный оператор играет ту же роль, что и матрица емкостей.

## Задачи

16.1. Пусть  $D = \tau_0 = (x dy \wedge dz + y dz \wedge dx + z dx \wedge dy)/r^3$ , где  $r^2 = x^2 + y^2 + z^2$ . Эта формула дает электрическое смещение в случае единичного положительного заряда, находящегося в начале координат.

(а) Вычислите  $\int_{R_1} D$ , где  $R_1$  — диск  $z = z_0$ ,  $x^2 + y^2 \leq a^2$ .

(б) Вычислите  $\int_{R_2}$ , где  $R_2$  — боковая поверхность цилиндра

$$-z_0 < z < z_0, \quad x^2 + y^2 = a^2.$$

(с) Используя результаты (а) и (б), проверьте, что  $\int_{R_3} D = 4\pi$ , где  $R_3$  — цилиндр с боковой поверхностью  $R_2$  и двумя дисками  $R_1$  снизу и сверху.

16.2. Дана 1-форма  $A = A_x dx + A_y dy + A_z dz$ . Величины  $A_x$ ,  $A_y$  и  $A_z$  являются функциями  $x$ ,  $y$ ,  $z$ . Используя линейный оператор  $\star$ , удовлетворяющий условиям

$$\begin{aligned} \star dx &= dy \wedge dz, & \star dy &= dz \wedge dx, & \star dz &= dx \wedge dy, \\ \star dy \wedge dz &= dx, & \star dz \wedge dx &= dy, & \star dx \wedge dy &= dz, \\ \star 1 &= dx \wedge dy \wedge dz, & \star dx \wedge dy \wedge dz &= 1, \end{aligned}$$

вычислите следующее:

- (a)  $\star dA$ ;
- (b)  $\star d\star A$ ;
- (c)  $\star d\star dA$ ;
- (d)  $\star d\star df$ , где  $f(x, y, z)$  — скалярное поле;
- (e)  $\star(A \wedge B)$ , где  $A$  и  $B$  — 1-формы;
- (f)  $\star(A \wedge \star B)$ , где  $A$  и  $B$  — 1-формы.

16.3. Сферические координаты  $r$ ,  $\theta$  и  $\phi$  определены формулами

$$x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi, \quad z = r \cos \theta.$$

- (a) Выразите  $dx$ ,  $dy$ ,  $dz$ ,  $dy \wedge dz$ ,  $dz \wedge dx$ ,  $dx \wedge dy$ ,  $dx \wedge dy \wedge dz$  в сферических координатах (т.е. используя формулы замены координат, сделайте перенос этих дифференциальных форм).
  - (b) Звездный оператор, определенный в упражнении 16.2, является линейным преобразованием, которое задано его действием в базисе  $dx$ ,  $dy$ ,  $dz$ ,  $dy \wedge dz$ ,  $\dots$ . Вычислите этот оператор в сферических координатах, т.е. выразите  $\star dr$ ,  $\star d\theta$  и  $\star d\phi$  через  $d\theta \wedge d\phi$ ,  $d\phi \wedge dr$  и  $dr \wedge d\theta$ . Выразите  $\star(d\theta \wedge d\phi)$ ,  $\star(d\phi \wedge dr)$  и  $\star(dr \wedge d\theta)$  через  $dr$ ,  $d\theta$ ,  $d\phi$ .
  - (c) Выразите через сферические координаты 2-форму  $\tau_0$  из упражнения 16.1 и покажите, что она равна  $\star d(1/r)$ .
- 16.4. Если  $u$  обозначает электрический потенциал в пространстве  $\mathbb{R}^3$ , то его лапласиан  $\Delta u$  можно записать в виде

$$\Delta u \, dx \wedge dy \wedge dz = d\star du.$$

- (a) Считая, что  $u$  является функцией  $x$ ,  $y$  и  $z$ , проверьте, что

$$\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}.$$

- (b) Считая, что  $u$  является функцией  $r$ ,  $\theta$  и  $\phi$ , получите формулу для  $\Delta u$  через частные производные  $u$  по  $r$ ,  $\theta$  и  $\phi$ .
- 16.5. Рассмотрите 2-форму  $D$ , которая в сферических координатах имеет вид  $D = r^3(1-r) \sin \theta \, d\theta \wedge d\phi$  ( $r \leq 1$ ),  $D = 0$  ( $r > 1$ ).
- (a) Проверьте, что если  $E = r(1-r)dr$  для  $r \leq 1$ , то  $D = \star E$ . (Заметим, что  $dr \wedge \star dr = dx \wedge dy \wedge dz$ .)

- (b) Найдите такую потенциальную функцию  $u$ , для которой  $E = -du$ . Вычислите  $\partial D$  (которая равна  $-dD$ ). Проверьте, что  $\int_D E = -\int_{\partial D} u$ , как это утверждается теоремой Стокса. (Вспомните, что  $\int_D E \equiv \int_{R^3} E \wedge D$ .)

16.6. Дифференциальная форма, представляющая  $D$  для единичного положительного заряда, может быть выражена в сферических координатах в виде  $\tau = \sin \theta \, d\theta \wedge d\phi$ .

- (a) Вычислите  $\int_S \tau$ , где  $S$  — сфера  $x^2 + y^2 + z^2 = 1$ .  
 (b) Покажите, что  $d\tau = 0$ .  
 (c) Используя теорему Стокса, объясните, почему *не может* существовать 1-форма  $\omega$ , определенная во всем пространстве  $\mathbb{R}^3$ , за исключением, может быть, начала координат, для которой  $d\omega = \tau$ .  
 (d) Найдите 1-форму  $\omega$ , определенную первом октанте пространства  $\mathbb{R}^3$  (где  $x > 0$ ,  $y > 0$ ,  $z > 0$ ), для которой  $\tau = d\omega$ . При этом для нахождения соответствующей  $\omega$  можно пользоваться сферическими координатами, но потом выразите  $\omega$  в декартовых координатах  $x, y, z$ .

16.7. Пусть  $u(r, \theta)$  — гладкая функция в пространстве  $\mathbb{R}^2$  с компактным носителем. Вычислите

$$\int_{\mathbb{R}^2} du \wedge \omega, \quad \text{где } \omega = d\theta = \frac{x \, dy - y \, dx}{x^2 + y^2}.$$

(Поскольку  $\omega$  не определена в начале координат, сначала надо интегрировать по области вне диска с радиусом  $\varepsilon$  и центром в начале координат, а потом взять предел при  $\varepsilon \rightarrow 0$ .)

16.8. Пусть  $\phi$  обозначает поворот в пространстве  $\mathbb{R}^3$ , так что  $\phi$  представляется матрицей  $A$ , для которой  $AA^t = \mathbb{I}$  и  $\text{Det } A = +1$ . Дана 1-форма  $\omega = F_x \, dx + F_y \, dy + F_z \, dz$ . Докажите, что в этом случае

$$\star(\phi^* \omega) = \phi^*(\star \omega).$$

16.9. Рассмотрите пять функций

$$v_1 = r^2, \quad v_2 = r^2 \cos^2 \theta (= z^2), \quad v_3 = r^4, \quad v_4 = r^4 \cos^2 \theta (= r^2 z^2),$$

$$v_5 = r^4 \cos^4 \theta (= z^4).$$

Они определяют пятимерное векторное пространство, которое обозначим  $C^0$ .

- (a) Для каждой из этих пяти функций вычислите  $dv$ ,  $\star dv$  и  $\Delta v$ .
- (b) Вычислите все скалярные произведения типа  $(dv_1, dv_2)_U$ , считая, что  $U$  — единичный шар  $r \leq 1$ .
- (c) Найдите базис двумерного пространства  $H$  подпространства  $C^0$ , которое состоит из функций, удовлетворяющих условию  $\Delta v = 0$ .
- (d) Постройте матрицу  $\mathbb{I} - \pi$ , которая является ортогональной проекцией  $dC^0$  на  $dH$ . Выпишите также матрицу  $\pi$ .
- (e) Используя матрицу  $\mathbb{I} - \pi$ , найдите функцию  $u$ , удовлетворяющую условию  $\Delta u = 0$  и равную  $\cos^4 \theta + \cos^2 \theta$  на сфере  $r = 1$ . (Задача Дирихле.)
- (f) С помощью  $\pi$  найдите функцию  $\phi$ , которая удовлетворяет уравнению  $\Delta \phi = 5r^2 - 4r^2 \cos^2 \theta + 1$  и обращается на единичной сфере в нуль. (Решение уравнения Пуассона.)

## Глава 17

---

# Токи, потоки и магнитостатика

---

В этой главе мы продолжаем изучение внешнего дифференциального исчисления. Здесь рассматриваются потоки и векторные поля, внутренние произведения и производные Ли. Все это прилагается к задачам магнитостатики.

### 17.1. Токи

На первый взгляд кажется, что можно сразу же дать непрерывный вариант описания цепей с сопротивлениями. Все, что надо сделать, — это лишь немного изменить то, что уже обсуждалось в предыдущей главе. Мы сохраняем форму  $E$  и уравнение  $E = -du$ , но энергию заменяем мощностью. Электростатическую 2-форму  $D$  заменяем на 2-форму  $J$ , которая должна представлять размазанный вариант токов, текущих по ветвям. У нас есть спаривание между током  $J$  и электрическим полем  $E$ :

$$\int E \wedge J,$$

что измеряется в единицах мощности. Согласно первому закону Кирхгофа,  $\partial I = 0$ . Поэтому соответствующий размазанный вариант имеет вид

$$\partial J = 0$$

или, поскольку  $\partial = -d$ ,

$$dJ = 0.$$

Вместо соотношения  $D = \varepsilon \star E$  для изотропной электростатической среды мы имеем

$$J = \sigma \star E,$$

где  $\sigma$  называется *проводимостью*. Если мы теперь введем величину  $r = 1/\sigma$ , называемую *удельным электрическим сопротивлением*, то предыдущее уравнение принимает вид

$$E = r \star^{-1} J,$$

что является размазанным вариантом закона Ома

$$V = RI.$$

Теперь надо остановиться и понять, что же представляет собой форма  $J$ . Рассмотрим некоторую поверхность  $S$ . Что означает  $\int_S J$ ? Пусть мы имеем конечную цепь, и поверхность пересекает ветви  $\alpha$ ,  $\beta$  и  $\gamma$ , ориентация которых показана на рис. 17.1. Тогда, принимая во внимание эти ориентации, можно считать, что  $S$  определяет коцепь, приписывающую каждому току  $I$  число  $I_\alpha - I_\beta + I_\gamma$ . В общем случае  $S$  определяет коцепь, приписывающую каждому току  $I$  *полный ток, протекающий через поверхность  $S$*  с учетом соответствующей ориентации. Таким образом, в размазанном варианте можно считать, что  $\int_S J$  представляет собой ток, протекающий через поверхность  $S$ .

Если мы представляем ток как движущиеся заряды, то  $\int_S J$  оказывается скоростью переноса зарядов через поверхность  $S$ .

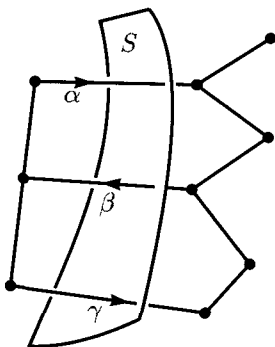


Рис. 17.1

Если  $S$  — замкнутая поверхность, т. е.  $S = \partial U$ , то мы имеем

$$\int_S J = \int_U dJ = 0.$$

Это значит, что полная скорость переноса заряда через поверхность  $S$  равна нулю, и заряд внутри нее не накапливается, что полностью согласуется с первым законом Кирхгофа, который мы обсуждали раньше. Значит, все прекрасно работает.

Существует, однако, геометрическая интерпретация, которая требует дополнительного объяснения. Мы знаем, что на самом деле ток создается движущимися электронами. Каждый электрон, будучи точечной частицей, движется вдоль некоторой кривой. Тогда форма  $J$ , проинтегрированная<sup>1</sup> по поверхности  $S$ , на самом деле считает число электронов, пересекающих поверхность  $S$  (в любом направлении) в единицу времени. У нас должен быть геометрический объект, непосредственно связанный с семейством траекторий отдельных электронов, который описывает их движение и из которого мы можем восстановить 2-форму  $J$ . В дальнейшем мы увидим, что таким объектом является *векторное поле*.

## 17.2. Потоки и векторные поля

В этом параграфе, чтобы избежать нагромождения индексов, мы в основном будем работать в пространстве  $\mathbb{R}^3$ . Однако, тот факт, что это пространство имеет три измерения, нам нигде не потребуются. Поэтому все выводы и формулы будут справедливы для пространств любого числа измерений.

Пусть  $U$  обозначает открытую область в пространстве  $\mathbb{R}^k$  и пусть  $I$  — некоторый интервал в  $\mathbb{R}$ , содержащий начало координат. Кроме того, пусть  $\phi : I \times U \rightarrow \mathbb{R}^k$  — дифференцируемое отображение. Момент времени будем обозначать  $t \in I$ . Тогда мы можем рассматривать  $\phi$  двумя способами. Выберем точку  $\mathbf{p} \in U$ . Тогда отображение

$$\phi(\cdot, \mathbf{p}) : t \rightarrow \phi(t, \mathbf{p})$$

<sup>1</sup>Здесь имеется в виду абсолютный, или неориентированный, интеграл (первого рода). Его значение неотрицательно. — *Прим. ред.*

определяет кривую в  $\mathbb{R}^k$ . Конечно, каждой точке  $\mathbf{p}$  соответствует своя кривая. Рассмотрим касательный вектор  $\phi'(t, \mathbf{p}) = (d/dt)\phi(t, \mathbf{p})$ . С другой стороны, для каждого момента времени  $t$  мы можем рассмотреть отображение  $\phi_t : U \rightarrow \mathbb{R}^k$ , заданное формулой

$$\phi_t(\mathbf{p}) = \phi(t, \mathbf{p}).$$

Предположим, что

$$\phi_0 = \text{id},$$

т.е. что

$$\phi(0, \mathbf{p}) = \mathbf{p}$$

для всех точек  $\mathbf{p}$ . Тогда для каждой точки  $\mathbf{p}$  касательная  $\phi'(0, \mathbf{p})$  к кривой  $\phi(\cdot, \mathbf{p})$  является касательным вектором в этой точке  $\mathbf{p}$ . Итак, мы получили правило, согласно которому каждой точке  $\mathbf{p}$  в области  $U$  ставится в соответствие касательный вектор в этой точке. Это правило называется *векторным полем* в области  $U$  (рис. 17.2). Векторное поле будем обозначать  $\xi$ . Итак,  $\xi(\mathbf{p})$  является касательным вектором в точке  $\mathbf{p}$ .

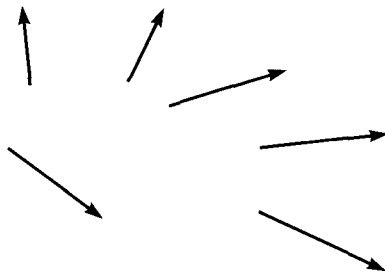


Рис. 17.2

Для произвольной точки  $\mathbf{p}$  предположение  $\phi_0 = \text{id}$  означает, что  $d\phi_0 = \text{id}$ , т.е.  $d(\phi_t)_{\mathbf{p}}$  несингулярна для достаточно малых значений  $t$ . По теореме о неявной функции  $\phi_t$  оказывается взаимно-однозначной функцией с дифференцируемой обратной в некоторой окрестности точки  $\mathbf{p}$ . Чтобы избежать нагромождения обозначений, предположим, что  $\phi_t : U \rightarrow \mathbb{R}^k$  имеет дифференцируемое обратное во всей области  $U$ . Выберем момент времени

$s \in I$ . Тогда определим

$$\psi_t = \phi_t \circ \phi_s^{-1}.$$

Тогда  $\psi_t$  обладает всеми свойствами  $\phi_t$  и при этом

$$\psi_s = \text{id}.$$

Итак, для каждого момента времени  $s \in I$  мы получаем векторное поле  $\xi_s$ . Значит, мы показали, что при данном  $\phi : I \times U \rightarrow \mathbb{R}^k$ , получается однопараметрическое семейство векторных полей  $\xi_s$ . Для каждой точки  $\mathbf{p} \in U$  существует некоторая окрестность  $W$  вокруг  $\mathbf{p}$  и некоторая окрестность  $K$  вокруг 0, так что  $\xi_s$  является векторным полем, определенным в  $W$ .

**Пример.** (а) Зададим вектор  $\mathbf{w}$ . Пусть  $\phi(t, \mathbf{p}) = \mathbf{p} + t\mathbf{w}$ . Тогда  $\phi'(0, \mathbf{p}) = \mathbf{w}$  для всех  $\mathbf{p}$ . Кроме того,

$$\begin{aligned} \phi_t \circ \phi_s^{-1}(\mathbf{p}) &= \phi_t(\mathbf{p} - s\mathbf{w}) \\ &= \mathbf{p} + (t - s)\mathbf{w} \\ &= \phi_{t-s}(\mathbf{p}), \end{aligned}$$

следовательно,

$$\xi_s(\mathbf{p}) \equiv \mathbf{w} \quad \text{для всех } \mathbf{p} \text{ и } s.$$

(b) Пусть задана матрица  $A$ . Определим  $\phi(t, \mathbf{p}) = e^{tA}\mathbf{p}$ . Тогда

$$\phi'(0, \mathbf{p}) = A\mathbf{p}.$$

Кроме того,

$$\begin{aligned} \phi_t \circ \phi_s^{-1}(\mathbf{p}) &= e^{tA}(e^{-sA}\mathbf{p}) \\ &= e^{(t-s)A}\mathbf{p} \\ &= \phi_{t-s}(\mathbf{p}) \end{aligned}$$

и, следовательно,

$$\xi_s(\mathbf{p}) = A\mathbf{p} \quad \text{для всех } \mathbf{p} \text{ и } s.$$

В этом случае векторное поле каждой точке  $\mathbf{p}$  ставит в соответствие вектор  $A\mathbf{p}$ .

Заметим, что в обоих примерах векторное поле  $\xi_s$  не зависит от  $s$ . Это является следствием тождества

$$\phi_t \circ \phi_s^{-1} = \phi_{t-s}$$

или, заменив  $s$  на  $-s$ , тождества

$$\phi_{t+s} = \phi_t \circ \phi_s.$$

Отображение  $\phi$ , удовлетворяющее этому тождеству всюду, где определены обе части, называется (стационарным) *поток*ом. В случае  $\phi_t(U) = U$  и  $I = \mathbb{R}$  мы говорим об *однопараметрической группе преобразований области  $U$* .

Итак, для потоков мы получили фиксированное векторное поле  $\xi = \xi_s$  для всех  $s$ .

Давайте сначала рассмотрим *линейное* векторное поле

$$\xi(\mathbf{p}) = A\mathbf{p},$$

где  $A$  — заданная матрица. Тогда, опираясь на результаты главы 3, мы знаем, что можно построить  $e^{tA}$  и потом определить  $\phi$  по формуле

$$\phi(t, \mathbf{p}) = e^{tA}\mathbf{p}.$$

Итак, мы получили теорему существования для обычных дифференциальных уравнений. Начав с  $\xi$ , мы получили  $\phi$ . Соответствующая теорема справедлива и в общем случае.

Пусть векторное поле  $\xi_s$  зависит гладко от  $s \in I$  и  $\mathbf{p} \in U$ . Тогда для каждой точки  $\mathbf{p}$  существует окрестность  $W \ni \mathbf{p}$  в пространстве  $\mathbb{R}^k$  и окрестность  $K$  нуля в  $\mathbb{R}$  такие, что отображение  $\phi : K \times W \rightarrow \mathbb{R}^k$ ,  $\phi_0 = \text{id}$  гладкое, а зависящие от  $s$  векторные поля, определяемые  $\phi$ , точно совпадают с  $\xi_s$ . Если же  $\xi_s \equiv \xi$  (т. е. не зависит от  $s$ ), то  $\phi$  удовлетворяет тождеству

$$\phi_{s+k} = \phi_s \circ \phi_t$$

всюду, где определены обе части.

Доказать эту теорему несложно. Это делается методом последовательных приближений. Здесь мы не будем приводить это доказательство. Оно стандартно, и его можно найти в любом учебнике по теории обыкновенных дифференциальных уравнений, в частности, в главе 6 книги Loomis and Sternberg.

Предположим, что имеются векторное поле  $\xi$  и дифференцируемая функция  $f$ . В каждой точке  $\mathbf{p}$  мы можем построить производную по направлению:  $D_{\xi(\mathbf{p})}f$ . Тогда получается новая функция, которая каждой точке  $\mathbf{p}$  приписывает значение  $D_{\xi(\mathbf{p})}f$ . Например, предположим, что в пространстве  $\mathbb{R}^3$  имеется векторное поле

$$\xi(\mathbf{p}) = \begin{pmatrix} a(\mathbf{p}) \\ b(\mathbf{p}) \\ c(\mathbf{p}) \end{pmatrix},$$

где  $a$ ,  $b$  и  $c$  — функции. Тогда

$$D_{\xi(\mathbf{p})}f = a(\mathbf{p})\frac{\partial f}{\partial x}(\mathbf{p}) + b(\mathbf{p})\frac{\partial f}{\partial y}(\mathbf{p}) + c(\mathbf{p})\frac{\partial f}{\partial z}(\mathbf{p}).$$

Или, в сокращенном виде,

$$D_{\xi}f = a\frac{\partial f}{\partial x} + b\frac{\partial f}{\partial y} + c\frac{\partial f}{\partial z}.$$

Поэтому в дальнейшем мы будем писать

$$\xi = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial z}.$$

В этих обозначениях, например, символ  $\partial/\partial x$  обозначает постоянное векторное поле

$$\mathbf{p} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Однако,

$$D_{\partial/\partial x}f = \frac{\partial f}{\partial x},$$

так что данное векторное поле действует как частная производная. Аналогично, векторное поле

$$\xi = \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

можно записать в виде

$$\xi = a \frac{\partial}{\partial x} + b \frac{\partial}{\partial y} + c \frac{\partial}{\partial z}.$$

Оно действует на функции по формуле

$$D_{\xi}f = a \frac{\partial f}{\partial x} + b \frac{\partial f}{\partial y} + c \frac{\partial f}{\partial z}.$$

Можно изложить это по-другому: дифференциал

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz$$

можно рассматривать как функцию, которая каждой точке приписывает вектор-строку

$$\left( \frac{\partial f}{\partial x}(\mathbf{p}), \frac{\partial f}{\partial y}(\mathbf{p}), \frac{\partial f}{\partial z}(\mathbf{p}) \right).$$

Тогда в каждой точке мы можем вычислить произведение этого

вектора-строки на вектор-столбец  $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$  и получить выражение

$$\langle \xi, df \rangle = \left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) \begin{pmatrix} a \\ b \\ c \end{pmatrix} = a \frac{\partial f}{\partial x} + b \frac{\partial f}{\partial y} + c \frac{\partial f}{\partial z}.$$

Итак,

$$D_{\xi}f = \langle \xi, df \rangle.$$

Очень полезно интерпретировать величину  $D_{\xi}f$  еще одним способом. Пусть имеется такое же отображение  $\phi : I \times U \rightarrow \mathbb{R}^k$  и

соответствующие векторные поля  $\xi_s$ . В каждый момент времени  $t$  мы можем рассмотреть функцию  $\phi_t^* f$ . Вспомним, что

$$\phi_t^* f(\mathbf{p}) = f(\phi_t(\mathbf{p})) = f(\phi(t, \mathbf{p})).$$

Так что

$$\frac{1}{t}(\phi_t^* f - f)$$

— по-прежнему хорошая функция. Вычислив ее в некоторой точке  $\mathbf{v}$  и устремив  $t \rightarrow 0$ , мы увидим, что

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{1}{t}(\phi_t^* f - f)(\mathbf{p}) &= \lim_{t \rightarrow 0} \frac{1}{t}[f(\phi(t, \mathbf{p})) - f(\mathbf{p})] \\ &= D_{\xi(\mathbf{p})} f. \end{aligned}$$

Откуда получаем

$$D_{\xi} f = \lim_{t \rightarrow 0} \frac{1}{t}(\phi_t^* f - f).$$

Теперь мы видим, что естественным геометрическим объектом, связанным с потоком, является векторное поле. В предыдущем параграфе мы увидели, что электрический ток (который по существу является потоком электронов) определяется 2-формой. Поэтому нам требуется математическая операция, которая позволит перейти от векторного поля, описывающего поток электронов, к 2-форме, определяющей ток. Эта 2-форма будет зависеть от плотности электронов. Очевидно, что более высокая плотность электронов,двигающихся вдоль тех же линий потока, будет создавать больший ток. Поэтому нам нужна такая математическая операция, которая даст возможность переходить от векторных полей  $\xi$  и 3-форм  $\rho dx \wedge dy \wedge dz$  к 2-формам. Эту операцию мы опишем в следующем параграфе.

## 17.3. Внутреннее произведение

Пусть  $V$  — векторное пространство, а  $\omega$  обозначает элемент  $\Lambda^k(V^*)$ . Вспомним, что  $\omega$  является функцией (полилинейной и

кососимметричной)  $k$  векторов в пространстве  $V$ . Для каждого набора  $\mathbf{v}_1, \dots, \mathbf{v}_k$  мы получаем число

$$\omega(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k).$$

Пусть  $\mathbf{v}$  обозначает вектор в  $V$ . Определим функцию  $k-1$  векторов в пространстве  $V$ , подставляя  $\mathbf{v}$  на первую позицию в  $\omega$ . Таким образом, мы рассматриваем функцию  $\mathbf{w}_1, \dots, \mathbf{w}_{k-1}$ , задаваемую

$$\omega(\mathbf{v}, \mathbf{w}_1, \dots, \mathbf{w}_{k-1}).$$

Эта функция векторов  $\mathbf{w}$  также оказывается полилинейной и кососимметричной. Следовательно, она является элементом  $\Lambda^{k-1}(V^*)$ . Назовем эту функцию *внутренним произведением*  $\mathbf{v}$  и  $\mathbf{w}$  и будем обозначать ее

$$i(\mathbf{v})\omega.$$

Итак,

$$i(\mathbf{v})\omega(\mathbf{w}_1, \dots, \mathbf{w}_{k-1}) = \omega(\mathbf{v}, \mathbf{w}_1, \dots, \mathbf{w}_{k-1}).$$

Давайте посмотрим, как выглядит внутреннее произведение в различных случаях. Допустим, что в пространстве  $\mathbb{R}^3$  у нас есть базис

$$\begin{aligned} \frac{\partial}{\partial x} &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{и элемент дуального базиса} \quad dx = (1, 0, 0), \\ \frac{\partial}{\partial y} &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{и элемент дуального базиса} \quad dy = (0, 1, 0), \\ \frac{\partial}{\partial z} &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{и элемент дуального базиса} \quad dz = (0, 0, 1). \end{aligned}$$

Тогда если  $k = 1$ , т.е.  $\omega$  является 1-формой, то  $i(\mathbf{v})\omega$  равняется  $\omega(\mathbf{v})$ , т.е. есть число. Таким образом,

$$i\left(\frac{\partial}{\partial x}\right)dx = 1, \quad i\left(\frac{\partial}{\partial x}\right)dy = 0, \quad i\left(\frac{\partial}{\partial x}\right)dz = 0$$

и так далее.

Если  $k = 2$ , то

$$\begin{aligned} i\left(\frac{\partial}{\partial x}\right)(dx \wedge dy)(\mathbf{w}) &= dx \wedge dy\left(\frac{\partial}{\partial x}, \mathbf{w}\right) \\ &= dx\left(\frac{\partial}{\partial x}\right)dy(\mathbf{w}) - dy\left(\frac{\partial}{\partial x}\right)dx(\mathbf{w}) \\ &= dy(\mathbf{w}), \end{aligned}$$

поскольку  $dy(\partial/\partial x) = 0$ . Значит,

$$i\left(\frac{\partial}{\partial x}\right)dx \wedge dy = dy.$$

Аналогично

$$i\left(\frac{\partial}{\partial x}\right)dx \wedge dz = dz$$

и

$$i\left(\frac{\partial}{\partial x}\right)dy \wedge dz = 0.$$

Кроме того,

$$\begin{aligned} i\left(\frac{\partial}{\partial y}\right)(dx \wedge dy)(\mathbf{w}) &= dx\left(\frac{\partial}{\partial y}\right)dy(\mathbf{w}) - dx(\mathbf{w})dy\left(\frac{\partial}{\partial y}\right) \\ &= -dx(\mathbf{w}). \end{aligned}$$

Отсюда

$$i\left(\frac{\partial}{\partial y}\right)dx \wedge dy = -dx, \quad i\left(\frac{\partial}{\partial y}\right)dx \wedge dz = 0, \quad i\left(\frac{\partial}{\partial y}\right)dy \wedge dz = dz.$$

Аналогично

$$i\left(\frac{\partial}{\partial z}\right)dx \wedge dy = 0, \quad i\left(\frac{\partial}{\partial z}\right)dx \wedge dz = -dx, \quad i\left(\frac{\partial}{\partial z}\right)dy \wedge dz = -dy.$$

Кроме того,

$$\begin{aligned}
 i\left(\frac{\partial}{\partial x}\right)(dx \wedge dy \wedge dz)(\mathbf{w}_1, \mathbf{w}_2) &= (dx \wedge dy \wedge dz)\left(\frac{\partial}{\partial x}, \mathbf{w}_1, \mathbf{w}_2\right) \\
 &= \text{Det} \begin{pmatrix} dx\left(\frac{\partial}{\partial x}\right) & dy\left(\frac{\partial}{\partial x}\right) & dz\left(\frac{\partial}{\partial x}\right) \\ dx(\mathbf{w}_1) & dy(\mathbf{w}_1) & dz(\mathbf{w}_1) \\ dx(\mathbf{w}_2) & dy(\mathbf{w}_2) & dz(\mathbf{w}_2) \end{pmatrix} \\
 &= \text{Det} \begin{pmatrix} 1 & 0 & 0 \\ dx(\mathbf{w}_1) & dy(\mathbf{w}_1) & dz(\mathbf{w}_1) \\ dx(\mathbf{w}_2) & dy(\mathbf{w}_2) & dz(\mathbf{w}_2) \end{pmatrix} \\
 &= \text{Det} \begin{pmatrix} dy(\mathbf{w}_1) & dz(\mathbf{w}_1) \\ dy(\mathbf{w}_2) & dz(\mathbf{w}_2) \end{pmatrix}.
 \end{aligned}$$

Следовательно,

$$i\left(\frac{\partial}{\partial x}\right)(dx \wedge dy \wedge dz) = dy \wedge dz.$$

Аналогично

$$i\left(\frac{\partial}{\partial y}\right)(dx \wedge dy \wedge dz) = -dx \wedge dz$$

и

$$i\left(\frac{\partial}{\partial z}\right)(dx \wedge dy \wedge dz) = dx \wedge dy.$$

Теперь ясно, что  $i(\mathbf{v})\omega$  билинейна по  $\mathbf{v}$  и  $\omega$ . Поэтому если

$$\xi = a\frac{\partial}{\partial x} + b\frac{\partial}{\partial y} + c\frac{\partial}{\partial z}$$

и

$$\omega = \rho dx \wedge dy \wedge dz,$$

то

$$i(\xi)\omega = \rho a dy \wedge dz - \rho b dx \wedge dz + \rho c dx \wedge dy.$$

Итак, мы определили внутреннее произведение  $i(\xi)\omega$  вектора  $\xi$  и элемента  $\omega$  пространства  $\Lambda^k(V^*)$ . Однако, используя это

определение в каждой точке, мы можем также определить внутреннее произведение  $i(\xi)\omega$ , где  $\xi$  — векторное поле, а  $\omega$  — дифференциальная форма степени  $k$ . И тогда все полученные формулы выполняются в пространстве  $\mathbb{R}^3$ .

В частности, если мы рассмотрим векторное поле  $\xi$ , представляющее стационарный поток электронов, и 3-форму  $\omega = \rho dx \wedge dy \wedge dz$ , дающую плотность электронов (точнее, размазанное приближение к ней), то получим 2-форму

$$J = i(\xi)\omega,$$

определяющую ток. Явный вид формулы в декартовых координатах дается предыдущим выражением. Чтобы понять, почему  $J$  представляет ток, можно воспользоваться исходным определением внутреннего произведения. Возьмем точку  $\mathbf{p}$  и построим небольшой параллелограмм на векторах  $\mathbf{w}_1$  и  $\mathbf{w}_2$ , выходящих из точки  $\mathbf{p}$ . Вектор  $\mathbf{v} = h\xi(\mathbf{p})$  представляет линейное приближение движения частицы, помещенной в точку  $\mathbf{p}$ , на протяжении небольшого временного интервала  $h$ . Другими словами, параллелепипед, построенный на векторах  $\mathbf{v}$ ,  $\mathbf{w}_1$  и  $\mathbf{w}_2$  (рис. 17.3), дает (приблизительно) область, проходимую параллелограммом за время  $h$ . Полный заряд этого параллелограмма дается выражением

$$\begin{aligned}\omega(\mathbf{v}, \mathbf{w}_1, \mathbf{w}_2) &= [i(\mathbf{v})\omega](\mathbf{w}_1, \mathbf{w}_2) \\ &= h[i(\xi(\mathbf{p}))\omega](\mathbf{w}_1, \mathbf{w}_2) \\ &= hJ(\mathbf{p})(\mathbf{w}_1, \mathbf{w}_2).\end{aligned}$$

Итак,  $J(\mathbf{p})(\mathbf{w}_1, \mathbf{w}_2)$  действительно представляет скорость потока заряда через параллелограмм, построенный на векторах  $\mathbf{w}_1$  и  $\mathbf{w}_2$ .

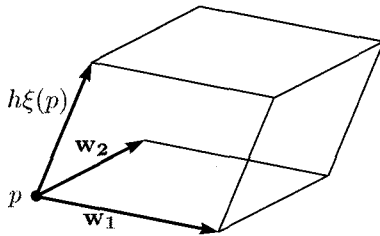


Рис. 17.3

Поэтому, если  $J$  проинтегрировать по поверхности, то получится скорость потока заряда через эту поверхность.

В заключение этого параграфа давайте докажем очень важную формулу для внутреннего произведения:

$$i(\xi)[\omega_1 \wedge \omega_2] = (i(\xi)\omega_1) \wedge \omega_2 + (-1)^{\deg \omega_1} \omega_1 \wedge i(\xi)\omega_2. \quad (17.1)$$

(Проверка показывает, что в пространстве  $\mathbb{R}^3$  формула (17.1) справедлива для всех случаев, которые мы рассмотрели.) В формуле (17.1)  $\xi$  обозначает векторное поле, а  $\omega_1$  и  $\omega_2$  дифференциальные формы. Поскольку (17.1) — чисто алгебраическое выражение, то достаточно доказать его для каждой точки, т.е. в случае, когда  $\xi = \mathbf{v}$  является вектором в пространстве  $V$ ,  $\omega_1$  — элементом  $\Lambda^k(V^*)$ , а  $\omega_2$  — элементом  $\Lambda^l(V^*)$ . Обе части выражения линейны по  $\xi$ ,  $\omega_1$  и  $\omega_2$ , когда два других аргумента фиксированы. Мы можем написать  $\xi = a_1 \mathbf{v}_1 + \dots + a_n \mathbf{v}_n$ , где  $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  является базисом. В силу линейности по  $\xi$ , достаточно проверить формулу для  $\xi = \mathbf{v}_i$ , т.е. для базисного элемента. Без потери общности предположим

$$\xi = \mathbf{v}_1.$$

Сначала докажем формулу в случае, когда  $\omega_1$  является 1-формой. В силу линейности можно считать, что

$$\omega_1 = \mathbf{v}_i^*, \quad i = 1, \dots, n,$$

т.е. является элементом дуального базиса. Кроме того, предположим, что  $\omega_2$  является базисным элементом вида

$$\omega_2 = \mathbf{v}_{j_1}^* \wedge \dots \wedge \mathbf{v}_{j_l}^*.$$

Рассмотрим теперь разные возможности.

(а) Пусть  $i = 1, j_1 > 1$ . Тогда  $\mathbf{v}_j^*(\mathbf{v}_1) = 0$  для всех  $j = j_1, \dots, j_l$ , и в определителе выражения для

$$(\omega_1 \wedge \omega_2)(\mathbf{v}_1, \mathbf{w}_1, \dots, \mathbf{w}_l)$$

все элементы, кроме левого верхнего, обращаются в нуль. Например, если  $\omega_2 = \mathbf{v}_2^* \wedge \mathbf{v}_3^* \wedge \mathbf{v}_4^*$ , то

$$\begin{aligned}
 (\omega_1 \wedge \omega_2)(\mathbf{v}, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3) &= \text{Det} \begin{pmatrix} \mathbf{v}_1^*(\mathbf{v}_1) & \mathbf{v}_2^*(\mathbf{v}_1) & \mathbf{v}_3^*(\mathbf{v}_1) & \mathbf{v}_4^*(\mathbf{v}_1) \\ \mathbf{v}_1^*(\mathbf{w}_1) & \mathbf{v}_2^*(\mathbf{w}_1) & \mathbf{v}_3^*(\mathbf{w}_1) & \mathbf{v}_4^*(\mathbf{w}_1) \\ \mathbf{v}_1^*(\mathbf{w}_2) & \mathbf{v}_2^*(\mathbf{w}_2) & \mathbf{v}_3^*(\mathbf{w}_2) & \mathbf{v}_4^*(\mathbf{w}_2) \\ \mathbf{v}_1^*(\mathbf{w}_3) & \mathbf{v}_2^*(\mathbf{w}_3) & \mathbf{v}_3^*(\mathbf{w}_3) & \mathbf{v}_4^*(\mathbf{w}_3) \end{pmatrix} \\
 &= \text{Det} \begin{pmatrix} 1 & 0 & 0 & 0 \\ \mathbf{v}_1^*(\mathbf{w}_1) & \mathbf{v}_2^*(\mathbf{w}_1) & \mathbf{v}_3^*(\mathbf{w}_1) & \mathbf{v}_4^*(\mathbf{w}_1) \\ \mathbf{v}_1^*(\mathbf{w}_2) & \mathbf{v}_2^*(\mathbf{w}_2) & \mathbf{v}_3^*(\mathbf{w}_2) & \mathbf{v}_4^*(\mathbf{w}_2) \\ \mathbf{v}_1^*(\mathbf{w}_3) & \mathbf{v}_2^*(\mathbf{w}_3) & \mathbf{v}_3^*(\mathbf{w}_3) & \mathbf{v}_4^*(\mathbf{w}_3) \end{pmatrix} \\
 &= \text{Det} \begin{pmatrix} \mathbf{v}_2^*(\mathbf{w}_1) & \mathbf{v}_3^*(\mathbf{w}_1) & \mathbf{v}_4^*(\mathbf{w}_1) \\ \mathbf{v}_2^*(\mathbf{w}_2) & \mathbf{v}_3^*(\mathbf{w}_2) & \mathbf{v}_4^*(\mathbf{w}_2) \\ \mathbf{v}_2^*(\mathbf{w}_3) & \mathbf{v}_3^*(\mathbf{w}_3) & \mathbf{v}_4^*(\mathbf{w}_3) \end{pmatrix} \\
 &= (\mathbf{v}_2^* \wedge \mathbf{v}_3^* \wedge \mathbf{v}_4^*)(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3).
 \end{aligned}$$

Итак, выражения для  $(\omega_1 \wedge \omega_2)(\mathbf{v}, \mathbf{w}_1, \dots, \mathbf{w}_l)$  и  $\omega_2(\mathbf{w}_1, \dots, \mathbf{w}_l)$  чрез определители совпадают. Следовательно,

$$i(\mathbf{v})(\omega_1 \wedge \omega_2) = \omega_2 = [i(\mathbf{v})\omega_1]\omega_2 - \omega_1 \wedge i(\mathbf{v})\omega_2,$$

поскольку в этом случае  $i(\mathbf{v})\omega_1 = 1$  и  $i(\mathbf{v})\omega_2 = 0$ .

(b) Пусть  $i = 1$ ,  $j_1 = 1$ . Тогда  $\omega_1 \wedge \omega_2 = 0$ . В этом случае  $i(\mathbf{v}_1)\omega_1 = 1$  и тогда

$$(i(\mathbf{v}_1)\omega_1) \wedge \omega_2 = \omega_2.$$

В то же время

$$i(\mathbf{v}_1)\omega_2 = \mathbf{v}_{j_2}^* \wedge \dots \wedge \mathbf{v}_{j_l}^*.$$

Тогда

$$\omega_1 \wedge i(\mathbf{v}_1)\omega_2 = \mathbf{v}_1^* \wedge \mathbf{v}_{j_2}^* \wedge \dots \wedge \mathbf{v}_{j_l}^* = \omega_2.$$

Таким образом,

$$i(\mathbf{v}_1)\omega_1 \wedge \omega_2 - \omega_1 \wedge i(\mathbf{v}_1)\omega_2 = \omega_2 - \omega_2 = 0.$$

(c) Пусть  $i > 1$ ,  $j_1 = 1$ . В этом случае можно написать  $\omega_2 = \mathbf{v}_1^* \wedge \omega_3$ . Тогда левая часть формулы (17.1) переписывается в виде

$$\begin{aligned}
 i(\mathbf{v})[\omega_1 \wedge \mathbf{v}_1^* \wedge \omega_3] &= -i(\mathbf{v})[\mathbf{v}_1^* \wedge (\omega_1 \wedge \omega_3)] \quad (\text{переставляя } \mathbf{v}_1 \text{ с } \omega_1) \\
 &= -\omega_1 \wedge \omega_3,
 \end{aligned}$$

согласно случаю (а). Первый член в правой части формулы (17.1) обращается в нуль, а второй член равен  $-\omega_1 \wedge \omega_3$  (еще раз применяем случая (а)).

(d) Пусть  $i > 1, j_1 > 1$ . Тогда

$$\begin{aligned} i(\mathbf{v}_1)(\omega_1 \wedge \omega_2) &= 0, \\ i(\mathbf{v}_1)\omega_1 &= 0 \end{aligned}$$

и

$$i(\mathbf{v}_1)\omega_2 = 0.$$

Поэтому обе части формулы (17.1) обращаются в нуль. Итак, мы доказали формулу (17.1) для случая, когда  $\deg \omega_1 = 1$ . Но тогда ассоциативный закон внешнего умножения дает нам возможность доказать эту формулу в общем случае. Например, предположим, что  $\omega_1 = \sigma_1 \wedge \sigma_2$ , где  $\deg \sigma_1 = \deg \sigma_2 = 1$ . Тогда

$$\begin{aligned} i(\mathbf{v})(\sigma_1 \wedge \sigma_2 \wedge \omega_2) &= i(\mathbf{v})\sigma_1 \wedge (\sigma_2 \wedge \omega_2) - \sigma_1 \wedge i(\mathbf{v})(\sigma_2 \wedge \omega_2) \\ &= i(\mathbf{v})\sigma_1 \wedge (\sigma_2 \wedge \omega_2) - \sigma_1 \wedge i(\mathbf{v})\sigma_2 \wedge \omega_2 + \sigma_1 \wedge \sigma_2 \wedge i(\mathbf{v})\omega_2 \\ &= [i(\mathbf{v})\sigma_1 \wedge \sigma_2 - \sigma_1 \wedge i(\mathbf{v})\sigma_2] \wedge \omega_2 + \sigma_1 \wedge \sigma_2 \wedge i(\mathbf{v})\omega_2 \\ &= i(\mathbf{v})(\sigma_1 \wedge \sigma_2) \wedge \omega_2 + (\sigma_1 \wedge \sigma_2) \wedge i(\mathbf{v})\omega_2. \end{aligned}$$

Итак, мы доказали формулу, когда  $\omega_1 = \sigma_1 \wedge \sigma_2$ , и тем самым, в силу линейности, для всех случаев, когда  $\deg \omega_1 = 2$ . Дальше доказательство продолжаем по индукции. Например, если  $\deg \sigma_1 = 1$  и  $\deg \sigma_2 = 2$ , то

$$\begin{aligned} i(\mathbf{v})(\sigma_1 \wedge \sigma_2 \wedge \omega_2) &= i(\mathbf{v})\sigma_1 \wedge (\sigma_2 \wedge \omega_2) - \sigma_1 \wedge i(\mathbf{v})(\sigma_2 \wedge \omega_2) \\ &= i(\mathbf{v})\sigma_1 \wedge \sigma_2 \wedge \omega_2 - \sigma_1 \wedge i(\mathbf{v})\sigma_2 \wedge \omega_2 - \sigma_1 \wedge \sigma_2 \wedge i(\mathbf{v})\omega_2 \\ &= i(\mathbf{v})(\sigma_1 \wedge \sigma_2) \wedge \omega_2 - \sigma_1 \wedge \sigma_2 \wedge i(\mathbf{v})\omega_2. \end{aligned}$$

Поэтому

$$i(\mathbf{v})\omega_1 \wedge \omega_2 = i(\mathbf{v})\omega_1 \wedge \omega_2 - \omega_1 \wedge i(\mathbf{v})\omega_2,$$

если  $\deg \omega_1 = 3$ , и так далее.

## 17.4. Производные Ли

Так же, как в параграфе 17.2, пусть  $\phi : I \times U \rightarrow \mathbb{R}^k$ . Тогда для каждого  $t \in I$  мы имеем отображение

$$\phi_t : U \rightarrow \mathbb{R}^k, \quad \phi_t(\mathbf{p}) = \phi(t, \mathbf{p}),$$

причем предполагается, что

$$\phi_0 = \text{id}.$$

Пусть  $\xi = \xi_0$  обозначает векторное поле, связанное с  $\phi_t$  в момент  $t = 0$ .

Пусть  $\omega$  — дифференциальная форма степени  $k$ . Рассмотрим форму  $\phi_t^* \omega$  и соответственно форму

$$\frac{1}{t}(\phi_t^* \omega - \omega)$$

при  $t \neq 0$ . Утверждается, что это выражение имеет предел при  $t = 0$ . Действительно, если  $\omega = f$  — функция, то такой предел при  $t = 0$  существует и, как мы уже видели, он записывается в виде

$$D_\xi f = \lim_{t \rightarrow 0} \frac{1}{t}(\phi_t^* f - f).$$

Если  $\omega = df$ , то

$$\phi_t^* df = d\phi_t^* f$$

и поэтому

$$\frac{1}{t}(\phi_t^* df - df) = d \left[ \frac{1}{t}(\phi_t^* f - f) \right].$$

Мы имеем право переставить пределы в частных производных, определяющих  $d$  и  $D_\xi$ . Тогда получим

$$\lim_{t \rightarrow 0} \frac{1}{t}(\phi_t^* df - df) = dD_\xi f.$$

Наиболее общий вид линейной дифференциальной формы является суммой выражений вида  $f dg$ . Тогда

$$\begin{aligned} \phi_t^*(f dg) - f dg &= (\phi_t^* f) \phi_t^* dg - f dg \\ &= \phi_t^* f \phi_t^* dg - f \phi_t^* dg + f(\phi_t^* dg) - f dg. \end{aligned}$$

Следовательно,

$$\lim_{t \rightarrow 0} \frac{1}{t} (\phi_t^* f dg - f dg) = (D_\xi f) dg + f d(D_\xi g).$$

Итак,  $\lim_{t \rightarrow 0} (1/t)(\phi_t^* \omega - \omega)$  существует для любой 1-формы. Более того, этот предел можно выразить через  $\xi$  и  $\omega$ . Этот предел будем обозначать  $D_\xi \omega$ . Далее, если

$$\omega = \omega_1 \wedge \omega_2,$$

где  $\omega_1$  и  $\omega_2$  — 1-формы, то мы имеем

$$\begin{aligned} \phi_t^*(\omega_1 \wedge \omega_2) - \omega_1 \wedge \omega_2 &= \phi_t^* \omega_1 \wedge \phi_t^* \omega_2 - \omega_1 \wedge \omega_2 \\ &= (\phi_t^* \omega_1 - \omega_1) \wedge \phi_t^* \omega_2 + \omega_1 \wedge (\phi_t^* \omega_2 - \omega_2). \end{aligned}$$

Разделив это выражение на  $t$  и вычисляя предел, получим, что

$$\lim_{t \rightarrow 0} \frac{1}{t} (\phi_t^*(\omega_1 \wedge \omega_2) - \omega_1 \wedge \omega_2)$$

существует и равен

$$D_\xi \omega_1 \wedge \omega_2 + \omega_1 \wedge D_\xi \omega_2.$$

Так как любая 2-форма может быть представлена как сумма членов вида  $\omega_1 \wedge \omega_2$ , то

$$\lim_{t \rightarrow 0} \frac{1}{t} (\phi_t^* \omega - \omega)$$

существует для любой 2-формы  $\omega$ . Этот предел мы будем обозначать  $D_\xi \omega$ , потому что он зависит только от  $\xi$  и  $\omega$ . Действуя таким образом, мы убедимся, что этот предел существует для форм  $\omega$  любой степени и зависит только от  $\xi$  и  $\omega$ . Мы его обозначим  $D_\xi \omega$  и назовем *производной Ли* от  $\omega$  по  $\xi$ . Для любых форм  $\omega_1$  и  $\omega_2$ , как и выше, можно доказать формулу

$$D_\xi(\omega_1 \wedge \omega_2) = D_\xi \omega_1 \wedge \omega_2 + \omega_1 \wedge D_\xi \omega_2.$$

Определение и равенство  $d\phi^* = \phi^*d$ , показывают, что

$$dD_\xi \omega = D_\xi d\omega.$$

Существует формула, связывающая внутреннее произведение  $i(\xi)$ , внешнюю производную  $d$  и производную Ли  $D_\xi$ . Прежде чем получить эту формулу, вспомним, что  $d$  отображает  $l$ -формы в  $(l+1)$ -формы и удовлетворяет равенству

$$d(\omega_1 \wedge \omega_2) = d\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge d\omega_2,$$

если степень  $\omega_1$  равна  $k$ . Кроме того,  $i(\xi)$  отображает  $l$ -формы в  $(l-1)$ -формы и удовлетворяет равенству

$$i(\xi)(\omega_1 \wedge \omega_2) = i(\xi)\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge i(\xi)\omega_2.$$

Поэтому  $d \circ i(\xi)$  отображает  $l$ -формы в  $l$ -формы и удовлетворяет равенству

$$\begin{aligned} [d \circ i(\xi)](\omega_1 \wedge \omega_2) &= d(i(\xi)\omega_1 \wedge \omega_2 + (-1)^k \omega_1 \wedge i(\xi)\omega_2) \\ &= di(\xi)\omega_1 \wedge \omega_2 + (-1)^{k-1} i(\xi)\omega_1 \wedge d\omega_2 \\ &\quad + (-1)^k d\omega_1 \wedge i(\xi)\omega_2 + (-1)^{2k} \omega_1 \wedge di(\xi)\omega_2. \end{aligned}$$

Заметим, что  $(-1)^{2k} = 1$ . Кроме того, имеется аналогичная формула

$$\begin{aligned} [i(\xi) \circ d](\omega_1 \wedge \omega_2) &= i(\xi)d\omega_1 \wedge \omega_2 + (-1)^{k+1} d\omega_1 \wedge i(\xi)\omega_2 \\ &\quad + (-1)^k i(\xi)\omega_1 \wedge d\omega_2 + \omega_1 \wedge i(\xi)d\omega_2. \end{aligned}$$

Если мы сложим эти два выражения, то средние члены сократятся, и мы получим *правило Лейбница*

$$\begin{aligned} [i(\xi) \circ d + d \circ i(\xi)](\omega_1 \wedge \omega_2) \\ = [i(\xi)d + di(\xi)]\omega_1 \wedge \omega_2 + \omega_1 \wedge [i(\xi)d + di(\xi)]\omega_2. \end{aligned}$$

А сейчас докажем фундаментальную формулу дифференциального исчисления

$$\boxed{D_\xi = i(\xi)d + di(\xi)}. \quad (17.2)$$

*Доказательство.* Обе части этого уравнения являются операторами, которые, действуя на произведения, удовлетворяют тождеству Лейбница. Каждая форма является суммой произведений

функций  $f$  и дифференциалов  $dg$ . Поэтому нам достаточно проверить тождество

$$D_{\xi}\omega = i(\xi)d\omega + di(\xi)\omega$$

для  $\omega = f$  и для  $\omega = dg$ . В случае  $\omega = f$  по условию мы имеем  $i(\xi)f = 0$  (не существует  $(-1)$ -форм). Тогда формула упрощается

$$D_{\xi}f = i(\xi)df = df(\xi),$$

что мы уже получили в параграфе 17.2. В случае  $\omega = dg$  мы имеем

$$\begin{aligned} D_{\xi}dg &= dD_{\xi}g \\ &= d(i(\xi)dg) \\ &= (i(\xi)d + di(\xi))dg, \end{aligned}$$

поскольку  $d^2 = 0$ . Итак, формула верна в общем случае. В приложении к этой главе мы дадим другое доказательство этой формулы в более общем виде.

Например, предположим, что

$$\omega = \rho dx \wedge dy \wedge dz$$

в пространстве  $\mathbb{R}^3$ . Тогда  $d\omega = 0$ , поскольку в трехмерном пространстве не существует ненулевых 4-форм. Тогда

$$D_{\xi}\omega = di(\xi)\omega.$$

Если мы положим

$$J = i(\xi)\omega,$$

то условие

$$dJ = 0$$

эквивалентно

$$D_{\xi}\omega = 0.$$

Это — инфинитезимальный способ выражения того, что  $\phi_t^*\omega = \omega$  для всех  $t$ . Тем самым, первый закон Кирхгофа можно переформулировать так: поток  $\phi_t$  сохраняет плотность заряда  $\omega$ , так что  $D_{\xi}\omega = 0$ .

## 17.5. Магнетизм

Электрические и магнитные явления в очень примитивной форме были известны еще с древних времен. Исследования этих загадочных явлений учеными семнадцатого и восемнадцатого веков привели в девятнадцатом веке к пониманию фундаментальной роли электромагнитных сил в природе и к революционным изменениям в обществе, возникшим в результате открытия электрических технологий.

В древности было известно, что если потереть янтарь (окаменевшую древесную смолу) тканью, то небольшие кусочки легкого материала поднимутся в воздух и приклеются к нему. Янтарь по-гречески называется «электрон». В книге «О магнетизме», опубликованной в Лондоне в 1600 году, У. Гилберт ввел термин «электричество» для обозначения притяжения, свойственного янтарю. Он посвятил янтарю целую большую главу, «чтобы показать природе притяжения тел к нему и подчеркнуть огромную разницу между этим свойством и магнитными действиями» и отличить «простую симпатию, связывающую железо и магниты, от неразборчивого поведения янтаря». Древние рассматривали притяжения янтаря и магнитов как одинаковые явления, давая им одинаковые объяснения. Кардано (1550) осознал ряд различий между электрическими и магнитными притяжениями. Однако, только Гилберт обнаружил, что электростатическое свойство притяжения янтаря может быть и у многих других твердых веществ. Конечно, электростатическое притяжение заряженного тела к легкому нейтральному телу за счет индуцированной поляризации и более сильное притяжение к телу с зарядом противоположного знака — очень сложные явления как для точного теоретического расчета, так и для эксперимента. Легкие кусочки бумаги летят к янтарю хаотически и приклеиваются к нему. Затем через некоторое время они падают. Поэтому, конечно, поведение янтаря по сравнению с постоянным притяжением железа к магниту или магнита к магниту можно считать «неразборчивым». По существу прошло еще почти два века, пока люди осознали, что фундаментальное притяжение и отталкивание в рамках электростатики действует между заряженными телами, и пока был открыт правильный закон взаимодействия между ними. Это было сделано

Кулоном в 1789 году в тщательно поставленном эксперименте с помощью изобретенных им крутильных весов.

В древнегреческих источниках имеются сведения о камнях, которые могут притягивать железо. Сообщалось, что один такой большой камень в окрестности города Магнезия в Малой Азии вытягивает кусочки железа из амуниции пастухов и гвозди из их ботинок. По имени этого города мы и получили термин «магнит». Было обнаружено, что этот камень может передавать свойство притягивать железной игле, причем у последней оно будет зависеть от направления. В том же эксперименте с крутильными весами Кулон обнаружил и измерил силы притяжения и отталкивания между магнитными полюсами. (Он еще раз открыл закон обратного квадрата.) Однако, современное понимание природы магнитных сил пришло после работ Ампера. Оказалось, что истинным объектом изучения было не притяжение железа к магниту или притяжение между магнитами, а сила, которую магнит оказывает на движущийся заряд (или, еще точнее, сила, с которой один движущийся заряд действует на другой).

Весной 1820 года, читая лекцию об электрических токах, Эрстед заметил, что стрелка компаса, случайно оказавшегося рядом, отклоняется током. Объявление об этом результате — что электрический ток действует на магнит — всколыхнуло научный мир. Электричество и магнетизм, разделенные в течение почти двух веков, вновь оказались связанными между собой. Ампер осознал, что если ток отклоняет магнитную стрелку, то магнит должен действовать на ток и два тока должны действовать друг на друга. В течение нескольких недель с 18 сентября по 9 октября 1820 года, Ампер без усталости провел серию блестящих экспериментов и геометрических исследований. В некотором смысле понятие дифференциальных форм возникло из математического анализа, проделанного Ампером, который открыл основные законы магнитостатики.

Магнитное поле создает силу, действующую на ток, а ток представляется нам как система движущихся зарядов. Значит, магнитное поле действует на заряды. В то же время сила зависит от скорости движения заряда. *Магнитная индукция* является

2-формой  $B$  в пространстве  $\mathbb{R}^3$ . Предположим, что заряд  $e$  движется через точку  $\mathbf{p}$  со скоростью  $\mathbf{v}$ . Тогда возникает сила

$$ei(\mathbf{v})B_{\mathbf{p}}.$$

В этом выражении  $B$  является 2-формой, поэтому  $B_{\mathbf{p}}$  равно значению  $B$  в точке  $\mathbf{p}$ , т. е. является элементом  $\Lambda^2(\mathbb{R}^3)$ . Таким образом  $i(\mathbf{v})B_{\mathbf{p}}$  — линейная функция касательных векторов в точке  $\mathbf{p}$ , т. е. то, что мы называли силовым полем (ср. том 1, с. 348).

Итак, если частица с зарядом  $e$  движется вдоль кривой  $\gamma$ , то в момент  $t$  на нее действует сила

$$ei(\gamma'(t))B_{\gamma(t)}.$$

Тогда из определения внутреннего произведения немедленно следует, что

$$i(\mathbf{v})i(\mathbf{w})\omega = -i(\mathbf{w})i(\mathbf{v})\omega$$

для любых  $\mathbf{v}$ ,  $\mathbf{w}$  и  $\omega$ . В частности,

$$i(\mathbf{v})i(\mathbf{v})\omega = 0,$$

или, если  $\omega \in \Lambda^2(V^*)$ , то

$$[i(\mathbf{v})\omega](\mathbf{v}) = 0.$$

Это значит, что составляющая силы вдоль направления движения отсутствует. Тогда сила, действующая на линейный проводник с током, будет перпендикулярна этому проводнику. (Это было проверено Ампером в остроумном эксперименте, когда ток протекал по подвижному проводу, соединяющему два сосуда со ртутью.)

В каждой точке форма  $B \neq 0$  будет определять направление в пространстве — прямую, задаваемую уравнением

$$i(\mathbf{w})B = 0.$$

(Читателю следует убедиться, что это два линейно независимых уравнения для  $\mathbf{w}$ .) Железные опилки, помещенные в магнитное поле (способные вращаться, но не двигаться поступательно) повернутся вдоль этого направления, создавая *магнитные силовые*

линии. (На токи вдоль именно этих направлений не действует сила.)

Ампер не только открыл силу, с которой магнитное поле действует на ток. Он также вывел количественную формулу закона (впервые обнаруженного Эрстедом), согласно которому ток создает магнитное поле. Для этого полезно ввести *1-форму*  $H$ , которую вслед за Зоммерфельдом будем называть *магнитным возбуждением* или *магнитным полем*. Второй закон Ампера говорит, что если произвольная поверхность  $S$  ограничена кривой  $\gamma = \partial S$ , то

$$\int_{\gamma} H = 4\pi \int_S J.$$

Другими словами, интеграл от магнитного поля по замкнутой кривой, ограничивающей некоторую поверхность, равен потоку тока через эту поверхность.

По теореме Стокса это то же самое, что

$$dH = 4\pi J.$$

В отличие от электростатики, мы можем сформулировать еще один закон магнитостатики в виде

$$\int_S B = 0 \quad \text{для любой замкнутой поверхности } S.$$

(Согласно Герцу, это значит, что «не существует магнитных зарядов».) И, наконец, нам надо получить закон, связывающий  $B$  и  $H$ . В неферромагнитных материалах он имеет вид

$$B = \mu \star H,$$

где  $\mu$  называется *магнитной восприимчивостью*. А сейчас давайте выпишем все законы магнитостатики.

- (i) Сила, действующая на элемент тока  $\mathbf{I} = e\mathbf{v}$  в точке  $\mathbf{p}$ , равна  $i(\mathbf{I})B_{\mathbf{p}}$ .
- (ii)  $\int_S B = 0$  для любой замкнутой поверхности  $S$ . Следовательно,

$$dB = 0.$$

$$(iii) \int_{\partial S} H = 4\pi \int_S J, \text{ или}$$

$$dH = 4\pi J \quad (\text{закон Ампера}).$$

$$(iv) B = \mu \star H.$$

Предположим, что мы находимся в области  $U$ , где  $H^2(U) = 0$ . Тогда условие  $dB = 0$  означает, что

$$B = dA$$

для некоторой 1-формы  $A$ . Конечно,

$$A = A_x dx + A_y dy + A_z dz$$

не определяется однозначно уравнением  $B = dA$ , поскольку функция  $A + d\psi$ , где  $\psi$  — произвольная функция, удовлетворяет тому же уравнению. Давайте возьмем такую функцию  $\psi$ , чтобы

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} = - \left( \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} \right).$$

Как только мы выбрали  $A$ , задача сводится к решению уравнения Пуассона.<sup>2</sup> Тогда, заменив  $A$  на  $A + d\psi$ , мы получим

$$\frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} = 0.$$

Теперь

$$H = \frac{1}{\mu} \star^{-1} B = \frac{1}{\mu} \star^{-1} dA$$

и

$$dH = J.$$

Предположим, что в рассматриваемой области  $\mu$  — постоянная. Тогда

$$\boxed{d\star^{-1}dA = \mu J.}$$

<sup>2</sup>Последующие вычисления станут яснее после того, как мы введем общую форму оператора  $\star$ . См. следующую главу.

В этом случае

$$d(A_x dx) = -\frac{\partial A_x}{\partial y} dx \wedge dy + \frac{\partial A_x}{\partial z} dz \wedge dx,$$

поэтому

$$\star^{-1}d(A_x dx) = -\frac{\partial A_x}{\partial y} dz + \frac{\partial A_x}{\partial z} dy$$

и

$$\begin{aligned} (d\star^{-1}d)(A_x dx) &= -\frac{\partial^2 A_x}{\partial x \partial y} dx \wedge dz - \frac{\partial^2 A_x}{\partial y^2} dy \wedge dz \\ &\quad + \frac{\partial^2 A_x}{\partial z \partial x} dx \wedge dy - \frac{\partial^2 A_x}{\partial z^2} dy \wedge dz. \end{aligned}$$

Проделав те же вычисления для  $A_y dy$  и  $A_z dz$  и сложив все вместе, получаем

$$(d\star^{-1}d)A = -(\Delta A_x dy \wedge dz + \Delta A_y dz \wedge dx + \Delta A_z dx \wedge dy),$$

где

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

— обычный оператор Лапласа. Тогда если написать

$$J = J_x dy \wedge dz + J_y dz \wedge dx + J_z dx \wedge dy,$$

то уравнения

$\begin{aligned} dA &= B, \\ B &= \mu \star H, \\ dH &= 4\pi J \end{aligned}$
---

сводятся к решению трех уравнений Пуассона:

$$\Delta A_x = -4\pi J_x,$$

$$\Delta A_y = -4\pi J_y,$$

$$\Delta A_z = -4\pi J_z.$$

Например, мы хотим получить  $B$  и  $H$  вне длинного тонкого прямолинейного провода, по которому течет постоянный ток  $I$ .

Предположим, что радиус провода равен  $a$ , и расположен он вдоль оси  $z$ . В этом случае  $J_x = J_y = 0$  и

$$J_z = \begin{cases} 0 & \text{вне провода,} \\ \frac{I}{\pi a^2} & \text{внутри провода.} \end{cases}$$

Тогда  $A_x = A_y = 0$ . Решая уравнение для  $A_z$ , получаем

$$A_z = -2\mu I \log \sqrt{x^2 + y^2}$$

для точек вне провода. Тогда

$$\begin{aligned} B = dA &= -\frac{\partial A_z}{\partial x} dx \wedge dz - \frac{\partial A_z}{\partial y} dy \wedge dz \\ &= 2\mu I \left( \frac{x}{r^2} dx \wedge dz + \frac{y}{r^2} dy \wedge dz \right) \\ &= \frac{2\mu I}{r} dr \wedge dz \end{aligned}$$

в цилиндрических координатах

$$\begin{aligned} x &= r \cos \theta, & y &= r \sin \theta, & z &= z, \\ r &= \sqrt{x^2 + y^2}. \end{aligned}$$

Силовыми линиями являются круги с центром в середине провода, и  $H = 1/\mu \star^{-1} B$  пропорционально  $d\theta$ . Если бы мы заранее знали, что

$$H = f d\theta,$$

то из соображений симметрии мы могли бы заключить, что  $f = f(r)$ . Тогда из соотношения

$$\int_{\gamma} H = 4\pi I$$

мы получили бы

$$f(r) = \frac{2I}{r}.$$

Поскольку  $d\theta/r = dr \wedge dz$ , мы бы тогда имели

$$B = \mu \star H = 2\mu T dr \wedge dz.$$

В общем случае решение уравнения Пуассона показывает, что  $A$  определяется объемным интегралом

$$A_x(\mathbf{v}) = \int \frac{1}{\|\mathbf{v} - \mathbf{w}\|} j_x(\mathbf{w}) d^3\mathbf{w}.$$

Аналогичные формулы получаются для  $A_y$  и  $A_z$ .

## Приложение: альтернативное доказательство фундаментальной формулы дифференциального исчисления

Здесь мы дадим альтернативное доказательство основной формулы (17.2). По существу, мы докажем ее в более общем случае. Пусть  $W \subset \mathbb{R}^m$  и  $Z \subset \mathbb{R}^n$  — открытые множества. Предположим, что нам дано дифференцируемое отображение  $\phi : W \times I \rightarrow Z$ , где  $I$  — некоторый интервал в  $\mathbb{R}$ , содержащий начало координат. Для каждой фиксированной точки  $\mathbf{w} \in W$  мы получим кривую  $\phi(\mathbf{w}, t)$ , зависящую от  $t$ . Пусть  $\xi_t(\mathbf{w})$  обозначает касательный вектор к кривой  $\phi(\mathbf{w}, t)$ . Это касательный вектор в пространстве образов  $\mathbb{R}^n$ , но зависящий от  $t$  и  $\mathbf{w} \in W$ . Кроме того, пусть отображение  $\phi_t : W \rightarrow Z$  определено как

$$\phi_t(\mathbf{w}) = \phi(\mathbf{w}, t).$$

Мы считаем, что  $\phi_t$  образуют однопараметрическое семейство отображений из  $W$  в  $Z$ .

Пусть  $\sigma$  — дифференциальная  $(k+1)$ -форма в пространстве  $Z$ . Для каждого  $t$  определим  $k$ -форму  $\phi_t^* i(\xi_t) \sigma$  в  $W$  согласно формуле

$$[\phi_t^* i(\xi_t) \sigma](\eta_1, \dots, \eta_k) = \sigma(\vec{\xi}_t(\mathbf{w}), d\phi_t \eta_1, \dots, d\phi_t \eta_k),$$

где  $\eta_1, \dots, \eta_k$  — касательные вектора в точке  $\mathbf{w}$ . Все  $k+1$  векторов, возникающие внутри  $\sigma$  в правой части этого соотношения, являются касательными векторами к кривым  $\phi_t(\mathbf{w})$ . Поэтому правая часть имеет смысл. Примем это как определение левой части данной формулы. Для обобщения формулы (17.2) мы хотим доказать следующее.

Пусть  $\sigma_t$  обозначает гладкое однопараметрическое семейство форм в  $Z$ . Тогда  $\phi_t^* \sigma_t$  является гладким семейством форм в  $W$ , и основная формула дифференциального исчисления форм принимает вид:

$$\frac{d}{dt} \phi_t^* \sigma_t = \phi_t^* \frac{d\sigma_t}{dt} + \phi_t^* (i(\xi_t) d\sigma_t) + d\phi_t^* i(\xi_t) \sigma_t. \quad (17.3)$$

Сначала докажем формулу для частного случая, когда  $W = Z = M \times I$ , где  $M$  является открытым подмножеством пространства  $\mathbb{R}^m$ , а  $\phi_t$  является отображением  $\psi_t : M \times I \rightarrow M \times I$ , определенным формулой

$$\psi_t(\mathbf{x}, s) = (\mathbf{x}, s + t).$$

Наиболее общий вид дифференциальной формы в  $M \times I$  можно записать в виде

$$ds \wedge a + b,$$

где  $a$  и  $b$  являются формами в  $M$ , которые могут зависеть от  $t$  и  $s$ . (В терминах локальных координат  $s, x^1, \dots, x^n$  эти формы оказываются суммами членов вида  $c dx^{i_1} \wedge \dots \wedge dx^{i_k}$ , где  $c$  является функцией, зависящей от  $t, s$  и  $\mathbf{x}$ .) Чтобы показать зависимость от  $\mathbf{x}$  и  $s$ , мы перепишем наше выражение в виде

$$\sigma_t = ds \wedge a(\mathbf{x}, s, t) dx + b(\mathbf{x}, s, t) dx.$$

В этих обозначениях очевидно, что

$$\psi_t^* \sigma_t = ds \wedge a(\mathbf{x}, s + t, t) dx + b(\mathbf{x}, s + t, t) dx$$

и поэтому

$$\begin{aligned} \frac{d\psi_t^* \sigma_t}{dt} &= ds \wedge \frac{\partial a}{\partial s}(\mathbf{x}, s + t, t) dx + \frac{\partial b}{\partial s}(\mathbf{x}, s + t, t) dx \\ &\quad + ds \wedge \frac{\partial a}{\partial t}(\mathbf{x}, s + t, t) dx + \frac{\partial b}{\partial t}(\mathbf{x}, s + t, t) dx. \end{aligned}$$

Следовательно,

$$\frac{d\psi_t^* \sigma_t}{dt} - \psi_t^* \left( \frac{d\sigma_t}{dt} \right) = ds \wedge \frac{\partial a}{\partial s}(\mathbf{x}, s + t, t) dx + \frac{\partial b}{\partial s}(\mathbf{x}, s + t, t) dx. \quad (17.4)$$

В этом случае также очевидно, что вектор, касательный к кривой  $\psi_t(\mathbf{x}, s)$ , равен  $\partial/\partial s$ , вычисленной в точке  $(\mathbf{x}, s+t)$ .

Тогда  $\partial/\partial s$  оказывается векторным полем и

$$i\left(\frac{\partial}{\partial s}\right)\sigma_t = a\,dx.$$

Следовательно,

$$\psi_t^*\left(i\left(\frac{\partial}{\partial s}\right)\sigma_t\right) = a(\mathbf{x}, s+t, t)\,dx$$

и поэтому

$$d\psi_t^*\left(i\left(\frac{\partial}{\partial s}\right)\sigma_t\right) = \frac{\partial a}{\partial s}(\mathbf{x}, s+t, t)\,ds \wedge dx + d_x a(\mathbf{x}, s+t, t)\,dx. \quad (17.5)$$

Здесь  $d_x$  обозначает внешнюю производную формы  $a(\mathbf{x}, s+t, t)\,dx$  в  $M$  при фиксированном значении  $s$ . Аналогично

$$d\sigma_t = -ds \wedge d_x a\,dx + \frac{\partial b}{\partial s}\,ds \wedge dx + d_x b\,dx.$$

Тогда

$$i\left(\frac{\partial}{\partial s}\right)d\sigma_t = -d_x a\,dx + \frac{\partial b}{\partial s}\,dx$$

и

$$\psi_t^* i\left(\frac{\partial}{\partial s}\right)d\sigma_t = -d_x a(\mathbf{x}, s+t, t)\,dx + \frac{\partial b}{\partial s}(\mathbf{x}, s+t, t)\,dx. \quad (17.6)$$

В результате сложения соотношений (17.4)–(17.6) мы получаем доказательство формулы (17.3) для  $\psi_t$ .

Пусть теперь  $\phi : W \times I \rightarrow Z$  имеет вид  $\phi(\mathbf{w}, s) = \phi_s(\mathbf{w})$ . В результате действия  $\phi$  прямые, параллельные  $I$  и проходящие через  $\mathbf{w}$  в  $W \times I$ , переходят в кривые  $\phi_s(\mathbf{w})$  в  $Z$ . Другими словами,

$$d\phi\left(\frac{\partial}{\partial s}\right)_{(\mathbf{w}, t)} = \xi_t(\mathbf{w}).$$

Если мы рассмотрим отображение  $l : W \rightarrow W \times I$ , заданное правилом  $l(\mathbf{w}) = (\mathbf{w}, 0)$ , то отображение  $\phi_t$  можно записать как

$\phi \circ \psi_t \circ l$ . Тогда  $\phi_t^* \sigma_t = l^* \psi_t^* \phi^* \sigma_t$ . Поскольку  $l$  и  $\phi$  не зависят от  $t$ , мы имеем

$$\frac{d}{dt} \phi_t^* \sigma_t = l^* \frac{d}{dt} \psi_t^* (\phi^* \sigma_t).$$

В точке  $w, t$  пространства  $W \times I$  можно написать

$$i \left( \frac{\partial}{\partial s} \right) \phi^* \sigma_t = (d\phi)^* \left\{ i \left( d\phi \frac{\partial}{\partial s} \right) \sigma_t \right\} = (d\phi)^* i(\xi_t) \sigma_t,$$

и тогда

$$l^* \psi_t^* \left( i \left( \frac{\partial}{\partial s} \right) \phi^* \sigma_t \right) = l^* \psi_t^* \phi^* (i(\xi_t) \sigma_t) = \phi_t^* (i(\xi_t) \sigma_t).$$

Подстановка в формулу для  $(d/dt)[\psi_t^* \phi^* \sigma_t]$  дает (17.3).

## Резюме

### А. Дифференциальные формы, векторные поля и магнетизм

Для данного векторного поля  $\xi$  и дифференциальной формы  $\omega$  надо знать, как определить и вычислить производную Ли  $D_\xi \omega$ .

Вы должны знать определение внутреннего произведения вектора и дифференциальной формы, уметь вычислять его и использовать для формулировки закона о магнитной силе.

Вы должны знать формулировки законов магнитостатики и технику вычисления  $B$  для данного тока  $J$  с помощью дифференциальных форм.

## Задачи

17.1. Пусть вектора  $\{e_1, e_2, e_3\}$  образуют базис в пространстве  $\mathbb{R}^3$ , так что

$$dx[e_1] = dy[e_2] = dz[e_3] = 1.$$

Пусть  $\xi = x e_2 - y e_1$ . Это векторное поле  $\xi$  описывает скорость, связанную с равномерным вращением вокруг оси  $z$ . Мы можем переписать это выражение в виде  $\xi = x(\partial/\partial y) - y(\partial/\partial x)$ .

- (а) Имеется постоянная плотность заряда, описываемая формой  $\omega = \rho_0 dx \wedge dy \wedge dz$ . Используя соотношение  $J = i(\xi)\omega$ , определите ток, связанный с вращением этого однородного распределения заряда.
- (б) Пусть  $\omega = z dx \wedge dy$ . В явном виде проверьте выполнение тождеств  $D_\xi \omega = i(\xi)d\omega + di(\xi)\omega$  и  $dD_\xi \omega = D_\xi d\omega$ , где  $\xi = x(\partial/\partial y) - y(\partial/\partial x)$ .

- 17.2. Пусть  $I$  обозначает ток на единицу длины, который распространяется аксиально симметрично относительно оси  $z$  таким образом, что, проходя через поверхность цилиндра радиуса  $r$ , ось которого является осью  $z$ , не зависит от  $r$ . В этом случае скорость носителей заряда определяется формулой

$$\xi = \frac{x\mathbf{e}_1 + y\mathbf{e}_2}{x^2 + y^2} \quad \text{или} \quad \xi = \frac{1}{x^2 + y^2} \left( x \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right).$$

Вычислите 2-форму  $J$ , используя равенство  $J = i(\xi)dx \wedge dy \wedge dz$  и убедитесь, что  $dJ = 0$ .

- 17.3. Предположим, что  $B = B_0 dz$ , где  $B_0$  — константа. Найдите функцию  $A$ , удовлетворяющую условиям

$$B = dA \quad \text{и} \quad \frac{\partial A_x}{\partial x} + \frac{\partial A_y}{\partial y} + \frac{\partial A_z}{\partial z} = 0.$$

- 17.4. Предположим, что по проводу с радиусом  $a$  движется аксиально симметричный ток  $J$ , величина которого пропорциональна  $r^2$ . В этом случае очевидно, что  $J_x = J_y = 0$  и

$$j_z = \begin{cases} 0 & \text{вне провода,} \\ \frac{2I(x^2 + y^2)}{\pi a^4} & \text{внутри провода.} \end{cases}$$

- (а) Найдите функцию  $A_z$ , зависящую только от  $r = \sqrt{x^2 + y^2}$ , для которой

$$\Delta A_z = -4\pi J_z.$$

- (б) Вычислите  $B$  и  $H$  в цилиндрических и декартовых координатах. Проверьте, что  $dH = 4\pi J$ .

- 17.5. Очень длинный цилиндр с радиусом  $a$  и осью вдоль  $z$  имеет однородную плотность заряда  $\rho_0$ . Вне цилиндра зарядов нет. Цилиндр вращается вокруг своей оси с угловой скоростью  $\omega$ , так что  $J = \omega r \rho_0$  для  $r < a$ .

- 
- (a) Выразите  $J$  в цилиндрических и декартовых координатах (см. задачу 17.1).
- (b) В силу симметрии форма  $B$  имеет вид  $B = f(r) dx \wedge dy$  и  $H = 1/\mu f(r) dz$ . Получите такое  $H$ , чтобы выполнялось условие  $dH = 4\pi J$ .
- (c) Найдите функцию  $A$  внутри и вне вращающегося цилиндра. В цилиндрических координатах  $A$  будет иметь вид  $g(r) d\theta$ .

# Глава 18

---

## Звездный оператор

---

В этой главе мы даем общее обсуждение звездного оператора. Это позволит нам сформулировать в следующей главе уравнения Максвелла и понять связь между ними и геометрией пространства-времени. Кроме того, мы объясним различные операторы и тождества векторного анализа.

### 18.1. Скалярные произведения и внешняя алгебра

Пусть  $V$  обозначает конечномерное векторное пространство с невырожденным скалярным произведением. Мы не предполагаем, что это скалярное произведение  $(\cdot, \cdot)$  положительно определено. Однако, полезно иметь в виду положительно определенный случай, чтобы интуитивно чувствовать, что происходит. В этом параграфе мы хотим объяснить, как  $(\cdot, \cdot)$  индуцирует скалярное произведение в каждом из пространств  $\Lambda^k(V^*)$ . Идея состоит в том, что за понятием длины следуют площадь параллелограмма, объем параллелепипеда и так далее.

Сначала отметим, что скалярное произведение в пространстве  $V$  индуцирует скалярное произведение в пространстве  $V^*$ . Действительно, любое скалярное произведение в  $V$  индуцирует отображение  $L : V \rightarrow V^*$ , задаваемое формулой

$$L(\mathbf{v})(\mathbf{w}) = (\mathbf{v}, \mathbf{w}), \quad \mathbf{v}, \mathbf{w} \in V.$$

В этом уравнении  $L(\mathbf{v})$  является элементом  $V^*$ , т. е. функцией в пространстве  $V$ , и поэтому ее можно вычислить для любого  $\mathbf{w} \in V$ . Отображение  $L$  определено так, что вычисление приводит к  $(\mathbf{v}, \mathbf{w})$ . Если  $(\cdot, \cdot)$  невырождено, то из того, что  $(\mathbf{v}, \mathbf{w}) = 0$  для всех  $\mathbf{w}$ , следует, что  $\mathbf{v} = 0$ . Следовательно, соотношение  $L(\mathbf{v})(\mathbf{w}) = 0$  для всех  $\mathbf{w}$  означает  $\mathbf{v} = 0$ , т. е. из  $L(\mathbf{v}) = 0$  следует, что  $\mathbf{v} = 0$ . Поскольку  $\text{Dim } V^* = \text{Dim } V$ , можно сделать вывод, что  $(\cdot, \cdot)$  не сингулярно тогда и только тогда, когда  $L$  является изоморфизмом. Мы предположили, что  $(\cdot, \cdot)$  невырождено, но тогда на самом деле  $L$  оказывается изоморфизмом. Теперь можно определить скалярное произведение в пространстве  $V^*$

$$(\alpha, \beta)_{V^*} = (L^{-1}\alpha, L^{-1}\beta)_V, \quad \alpha, \beta \in V^*.$$

Например, предположим, что  $V = \mathbb{R}^n$  состоит из векторов-столбцов. Тогда  $V^* = \mathbb{R}^{n*}$  состоит из векторов-строк. В общем случае скалярное произведение в пространстве  $\mathbb{R}^n$  дается симметричной матрицей  $S$ . Допустим, что  $S$  диагональна. Тогда

$$(\mathbf{v}, \mathbf{w})_V = \sum s_i x_i y_i,$$

если

$$\mathbf{v} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{и} \quad \mathbf{w} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

В этом случае

$$L\mathbf{v} = (s_1 x_1, \dots, s_n x_n).$$

Итак, если

$$\begin{aligned} \alpha &= (\alpha_1, \dots, \alpha_n), \\ \beta &= (\beta_1, \dots, \beta_n), \end{aligned}$$

то

$$L^{-1}\alpha = \begin{pmatrix} \alpha_1/s_1 \\ \vdots \\ \alpha_n/s_n \end{pmatrix}, \quad L^{-1}\beta = \begin{pmatrix} \beta_1/s_1 \\ \vdots \\ \beta_n/s_n \end{pmatrix}.$$

Следовательно,

$$(\alpha, \beta)_{V^*} = \sum \frac{1}{s_i} \alpha_i \beta_i.$$

Заметим, что  $s_i$  стоит в знаменателе. Поэтому мы должны предположить, что все  $s_i \neq 0$ . А это и есть предположение о невырожденности  $(\cdot, \cdot)$ .

Если  $(\cdot, \cdot)$  положительно определено и  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  является ортонормированным базисом в пространстве  $V$ , то  $\{\mathbf{e}_1^*, \dots, \mathbf{e}_n^*\}$  будет дуальным базисом. Исследование определения  $L$  показывает, что

$$L(\mathbf{e}_i) = \mathbf{e}_i^*$$

и, следовательно,  $\{\mathbf{e}_1^*, \dots, \mathbf{e}_n^*\}$  образует ортонормированный базис в пространстве  $V^*$ . В более общем случае, если не предполагать положительной определенности  $(\cdot, \cdot)$  (но при этом предполагать невырожденность), то можно взять в качестве  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  ортогональный базис с условием  $(\mathbf{e}_i, \mathbf{e}_i) = \pm 1$ . Тогда  $\{\mathbf{e}_1^*, \dots, \mathbf{e}_n^*\}$  оказывается ортогональным базисом в  $V^*$  и

$$(\mathbf{e}_i^*, \mathbf{e}_i^*)_{V^*} = (\mathbf{e}_i, \mathbf{e}_i)_V.$$

Теперь можно дать определение  $(\cdot, \cdot)$  для двух любых линейных дифференциальных форм. Например, в евклидовом пространстве  $\mathbb{R}^n$  базис 1-форм  $dx^1, \dots, dx^n$  ортонормирован и

$$(dx^i, dx^i) = +1.$$

Пусть в пространстве  $\mathbb{R}^3$  есть две формы  $\omega = A_x dx + A_y dy + A_z dz$  и  $\tau = B_x dx + B_y dy + B_z dz$ . Тогда  $(\omega, \tau) = A_x B_x + A_y B_y + A_z B_z$ . Заметим, что  $(\omega, \tau)$  является функцией.

Давайте рассмотрим пространство  $\Lambda^k(V^*)$ . Каждый элемент этого пространства является суммой вида

$$\alpha = \alpha^1 \wedge \dots \wedge \alpha^k + \beta^1 \wedge \dots \wedge \beta^k + \dots$$

Утверждается, что существует единственное скалярное произведение, определенное в пространстве  $\Lambda^k(V^*)$  и удовлетворяющее условию

$$(\alpha^1 \wedge \dots \wedge \alpha^k, \gamma^1 \wedge \dots \wedge \gamma^k) = \text{Det}((\alpha^i, \gamma^j)).$$

Очевидно, что если такое скалярное произведение существует, то этим уравнением оно определено единственным образом, поскольку в силу билинейности  $(\cdot, \cdot)$  можно продолжить на все  $\alpha$  и  $\gamma$ , которые представляют собой суммы разложимых элементов, т. е. выражений вида  $\alpha^1 \wedge \dots \wedge \alpha^k$  и  $\gamma^1 \wedge \dots \wedge \gamma^k$ . Задача состоит в том, чтобы доказать согласованность такого определения, т. е. показать, что если  $\alpha$  и/или  $\gamma$  записываются в виде сумм разложимых элементов двумя различными способами, то мы получим одно и то же значение  $(\alpha, \gamma)$ . Такое доказательство становится простым, если мы воспользуемся более абстрактным определением пространства  $\Lambda^k(V^*)$ , чем то, чем пользовались до сих пор. Это определение и доказательство приводятся в приложении к этой главе, поэтому здесь мы не будем отвлекаться.

Сейчас получим скалярное произведение (значения которого суть функции) в пространстве  $k$ -форм. Например, для 2-форм

$$(\omega^1 \wedge \omega^2, \tau^1 \wedge \tau^2) = (\omega^1, \tau^1)(\omega^2, \tau^2) - (\omega^1, \tau^2)(\omega^2, \tau^1).$$

В пространстве  $\mathbb{R}^3$  для евклидова скалярного произведения

$$(dx \wedge dy, dx \wedge dy) = +1.$$

В четырехмерном пространстве-времени для скалярного произведения Лоренца мы имеем

$$(dt \wedge dx, dt \wedge dx) = -1,$$

но

$$(dx \wedge dy, dx \wedge dy) = +1.$$

В качестве иллюстрации, как делать такие вычисления в более общей системе координат, рассмотрим сферические координаты в пространстве  $\mathbb{R}^3$ , которые определяются формулами

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \theta = \arctan(\sqrt{(x^2 + y^2)}/z), \quad \phi = \arctan(y/x).$$

Если мы вычислим дифференциалы этих функций и выразим результаты через базис  $\{dx, dy, dz\}$  с коэффициентами, записанными для удобства в координатах  $r, \theta$  и  $\phi$ , то получим

$$dr = \sin \theta \cos \phi dx + \sin \theta \sin \phi dy + \cos \theta dz,$$

$$d\theta = \frac{1}{r}(\cos \theta \cos \phi dx + \cos \theta \sin \phi dy - \sin \theta dz),$$

$$d\phi = \frac{1}{r \sin \theta}(-\sin \phi dx + \cos \phi dy).$$

Используя ортонормированность  $dx, dy$  и  $dz$ , прямое вычисление показывает, что  $dr, d\theta$  и  $d\phi$  являются ортогональными элементами пространства  $V^*$  в каждой точке  $(r, \theta, \phi)$ . При этом

$$(dr, dr) = 1, \quad (d\theta, d\theta) = \frac{1}{r^2}, \quad (d\phi, d\phi) = \frac{1}{r^2 \sin^2 \theta}.$$

Отсюда немедленно следует, что  $dr \wedge d\theta, dr \wedge d\phi$  и  $d\theta \wedge d\phi$  образуют ортогональный базис в пространстве  $\Lambda^2(\mathbb{R}^3)$ , а скалярные произведения

$$(dr \wedge d\theta, dr \wedge d\theta) = \frac{1}{r^2} \quad \text{и т. д.}$$

Наконец, мы можем вычислить скалярное произведение базисного элемента  $dr \wedge d\theta \wedge d\phi$  на себя:

$$(dr \wedge d\theta \wedge d\phi, dr \wedge d\theta \wedge d\phi) = \frac{1}{r^4 \sin^2 \theta}.$$

Чтобы подвести итог полученным результатам, лучше всего отметить, что

$$dr, \quad r d\theta, \quad r \sin \theta d\phi$$

образуют *ортонормированный* базис  $\Lambda^1(\mathbb{R}^{3*})$  в каждой точке. Поэтому мы можем работать с этими тремя дифференциалами так же, как с  $dx, dy$  и  $dz$ .

## 18.2. Звездный оператор

Рассмотрим конечномерное векторное пространство  $V$ . В главе 8 мы обсуждали понятие *ориентации* векторного пространства  $V$ . Вспомним это определение. Любые два базиса  $\{e_1, \dots, e_n\}$

и  $\{f_1, \dots, f_n\}$  связаны между собой матрицей замены базиса  $B$ . Два базиса называются *эквивалентными* или *определяют одинаковую ориентацию*, если  $\text{Det } B > 0$ . В противном случае они определяют противоположные ориентации. Множество всех базисов разбивается на два класса эквивалентности, и каждый класс эквивалентности называется *ориентацией*. С помощью пространства  $\Lambda^n(V)$  мы можем сформулировать это иначе. Базис  $\{e_1, \dots, e_n\}$  определяет ненулевой вектор в одномерном пространстве  $\Lambda^n(V)$ , а именно

$$e_1 \wedge \dots \wedge e_n.$$

Второй базис  $\{f_1, \dots, f_n\}$  определяет элемент  $f_1 \wedge \dots \wedge f_n$ . Эти два элемента отличаются скалярным множителем  $\text{Det } B$ . Таким образом, выбор ориентации пространства  $V$  равносильно выбору одного из двух классов эквивалентности ненулевых элементов  $\Lambda^n(V)$ . Эти два элемента считаются эквивалентными, если они отличаются друг от друга *положительным* множителем.

Предположим, что в пространстве  $V$  есть невырожденное скалярное произведение. Пусть  $\{e_1, \dots, e_n\}$  будет ортогональным базисом в  $V$ , с условием  $(e_i, e_i) = \pm 1$ . Два таких базиса отличаются друг от друга матрицей замены базиса  $B$ , для которой  $\text{Det } B = \pm 1$ . Таким образом, выбор ориентации  $V$  вместе со скалярным произведением определяет единственный элемент  $e_1 \wedge \dots \wedge e_n$  пространства  $\Lambda^n(V)$ . Поэтому у нас есть также единственный дуальный элемент из пространства  $\Lambda^n(V^*)$ , который будем обозначать  $\sigma$ .

Предположим, что мы рассматриваем дифференциальные формы. Тогда в каждой точке пространства  $V$  мы получаем единственный элемент  $\Lambda^n(V^*)$ , другими словами, мы получаем  $n$ -форму. Чтобы избежать громоздких обозначений (но рискуя возможной путаницей), эту  $n$ -форму тоже будем обозначать  $\sigma$ . Итак, если  $V = \mathbb{R}^n$  —  $n$ -мерное евклидово пространство, то

$$\sigma = dx^1 \wedge dx^2 \wedge \dots \wedge dx^n.$$

Рассматривая пространство-время, договоримся время считать *первой* координатой. Тогда для четырехмерного пространства

$$\sigma = dt \wedge dx \wedge dy \wedge dz.$$

(Заметим, что в этом случае  $(\sigma, \sigma) = -1$ .)

Сейчас мы уже можем определить линейное отображение из  $\Lambda^k(V^*)$  в  $\Lambda^{n-k}(V^*)$ , называемое *звездным оператором*. Для этого будем идентифицировать  $\Lambda^k(V^*)$  и  $\Lambda^{n-k}(V^*)$  с пространством линейных функций на  $\Lambda^{n-k}(V^*)$ . Потом отождествим элементы  $\Lambda^k(V^*)$  и  $\Lambda^{n-k}(V^*)$ , которые соответствуют одной и той же линейной функции. Такая идентификация  $\Lambda^k(V^*)$  и  $\Lambda^{n-k}(V^*)$  будет называться звездным оператором.

Сначала покажем, как *внешнее* произведение вместе с выбором  $\sigma \in \Lambda^n(V^*)$  приписывает каждому элементу  $\lambda \in \Lambda^k(V^*)$  линейную функцию на пространстве  $\Lambda^{n-k}(V^*)$ . Действительно, если  $\omega \in \Lambda^{n-k}(V^*)$ , то  $\lambda \wedge \omega$  является элементом  $\Lambda^n(V^*)$ . Следовательно, эта функция должна быть пропорциональна некоторому множителю  $\sigma$ . Другими словами, мы можем написать

$$\lambda \wedge \omega = f(\omega)\sigma.$$

Итак, каждый элемент  $\lambda$  определяет линейную функцию

$$\omega \mapsto f(\omega).$$

При этом существует единственный элемент пространства  $\Lambda^{n-k}(V^*)$ , обозначаемый в дальнейшем  $\star\lambda$ , который определяет ту же самую функцию  $f(\omega)$  из  $\Lambda^{n-k}(V^*)$  в  $\mathbb{R}$  через *скалярное* произведение

$$(\star\lambda, \omega) = f(\omega).$$

Таким образом, задав любой элемент  $\lambda \in \Lambda^k(V^*)$ , мы можем определить  $\star\lambda \in \Lambda^{n-k}(V^*)$  условием

$$\lambda \wedge \omega = (\star\lambda, \omega)\sigma$$

для всех  $\omega \in \Lambda^{n-k}(V^*)$ . Заметим, что это определение звездного оператора зависит от выбора ориентации. Изменение ориентации меняет знак  $\sigma$  и, следовательно, знак  $\star\lambda$ .

Чтобы вычислить  $\star\lambda$ , в общем случае мы должны применить это определение, используя в качестве  $\omega$  по очереди каждый базисный элемент  $\Lambda^{n-k}(V^*)$ . Однако, если  $\lambda$  является базисным элементом  $\Lambda^k(V^*)$  вида  $dx^{i_1} \wedge \cdots \wedge dx^{i_k}$ , то нам достаточно рассмотреть только один базисный элемент  $\Lambda^{n-k}(V^*)$ , а именно тот  $\omega$ ,

который содержит  $n - k$  множителей, не входящих в  $\lambda$ . В этом случае  $\lambda \wedge \omega$  оказывается внешним произведением всех дифференциалов от  $dx^1$  до  $dx^n$  в каком-то порядке, т. е. оно равно  $\pm\sigma$ . Тогда  $\star\lambda = \pm\omega$ . Осталась единственная трудность — правильно определить знак.

Общее определение

$$\lambda \wedge \omega = (\star\lambda, \omega)\sigma$$

следует дополнить соглашениями для случаев  $k = n$  и  $k = 0$ . Базисный элемент пространства  $\Lambda^0(\mathbb{R}^n)$  обозначим 1. При этом скалярное произведение  $(1, 1) = 1$  и тривиальное внешнее произведение  $1 \wedge \omega = \omega \wedge 1 = \omega$ . Тогда если  $\lambda = \sigma$ , то нам надо рассмотреть только базисный элемент  $\omega = 1$  и

$$\sigma \wedge 1 = (\star\lambda, 1)\sigma,$$

так что  $\star\sigma = 1$ . С другой стороны, если  $\lambda = 1$ , то рассматриваем  $\omega = \sigma$  и получаем

$$1 \wedge \sigma = (\star\lambda, \sigma)\sigma.$$

В этом случае  $\star 1 = (\sigma, \sigma)\sigma$ . Это значит, что если мы работаем с евклидовым скалярным произведением, для которого  $(\sigma, \sigma) = 1$ , то мы имеем  $\star\sigma = 1$  и  $\star 1 = \sigma$ . Однако, в важном случае двумерного или четырехмерного пространства-времени, со скалярным произведением Лоренца,  $(\sigma, \sigma) = -1$ , следовательно,

$$\star\sigma = +1, \quad \text{но} \quad \star 1 = -\sigma.$$

Вычисление звездного оператора лучше всего иллюстрируется на примерах.

**Пример 1.** В пространстве  $\mathbb{R}^2$  имеется евклидово скалярное произведение. Базис пространства  $\Lambda^1(V^*)$  образован  $dx$  и  $dy$  с условиями  $(dx, dx) = 1$ ,  $(dy, dy) = 1$  и  $\sigma = dx \wedge dy$ .

$$\star(dx \wedge dy) = 1 \quad (\text{верно всегда});$$

$$dx \wedge dy = (\star dx, dy) dx \wedge dy, \quad \text{поэтому} \quad \star dx = dy;$$

$$dy \wedge dx = (\star dy, dx) dx \wedge dy, \quad \text{поэтому} \quad \star dy = -dx;$$

$$\star 1 = dx \wedge dy, \quad \text{потому что} \quad (\sigma, \sigma) = +1.$$

**Пример 2.** Дано двумерное пространство-время, где определено скалярное произведение Лоренца. В пространстве  $\Lambda^1(V)$  имеется базис  $\{dt, dx\}$ , удовлетворяющий условиям  $(dt, dt) = +1$ ,  $(dx, dx) = -1$  и  $\sigma = dt \wedge dx$ .

$$\star dt \wedge dx = 1;$$

$$dt \wedge dx = (\star dt, dx) dt \wedge dx, \quad \text{поэтому} \quad \star dt = -dx;$$

$$dx \wedge dt = (\star dx, dt) dt \wedge dx, \quad \text{поэтому} \quad \star dx = -dt;$$

$$\star 1 = -dt \wedge dx, \quad \text{потому что} \quad (\sigma, \sigma) = -1.$$

**Пример 3.** В пространстве  $\mathbb{R}^3$  определено евклидово скалярное произведение. Базисом пространства  $\Lambda^1(V)$  является  $\{dx, dy, dz\}$ . При этом выполняются условия  $(dx, dx) = (dy, dy) = (dz, dz) = 1$  и  $\sigma = dx \wedge dy \wedge dz$ .

$$\star(dx \wedge dy \wedge dz) = 1$$

$$(dx \wedge dy) \wedge dz = (\star(dx \wedge dy), dz)\sigma, \quad \text{тогда} \quad \star(dx \wedge dy) = dz;$$

$$(dx \wedge dz) \wedge dy = (\star(dx \wedge dz), dy)\sigma, \quad \text{тогда} \quad \star(dx \wedge dz) = -dy;$$

$$(dy \wedge dz) \wedge dx = (\star(dy \wedge dz), dx)\sigma, \quad \text{тогда} \quad \star(dy \wedge dz) = dx;$$

$$dx \wedge (dy \wedge dz) = (\star dx, dy \wedge dz)\sigma, \quad \text{тогда} \quad \star dx = dy \wedge dz;$$

$$dy \wedge (dx \wedge dz) = (\star dy, dx \wedge dz)\sigma, \quad \text{тогда} \quad \star dy = -dx \wedge dz;$$

$$dz \wedge (dx \wedge dy) = (\star dz, dx \wedge dy)\sigma, \quad \text{тогда} \quad \star dz = dx \wedge dy;$$

$$\star 1 = dx \wedge dy \wedge dz.$$

**Пример 4.** Рассмотрим четырехмерное пространство-время, где определено скалярное произведение Лоренца. Базисом пространства  $\Lambda^1(V)$  является  $\{dt, dx, dy, dz\}$  с условиями  $(dt, dt) = 1$ ,  $(dx, dx) = (dy, dy) = (dz, dz) = -1$  и  $\sigma = dt \wedge dx \wedge dy \wedge dz$ . Тогда

$$\star(dt \wedge dx \wedge dy \wedge dz) = 1.$$

Далее,

$$(dt \wedge dx \wedge dy) \wedge dz = (\star(dt \wedge dx \wedge dy), dz)\sigma,$$

поэтому

$$\star(dt \wedge dx \wedge dy) = -dz.$$

Аналогично

$$\star(dt \wedge dx \wedge dz) = +dy;$$

$$\star(dt \wedge dy \wedge dz) = -dx.$$

Кроме того,

$$(dx \wedge dy \wedge dz) \wedge dt = (\star(dx \wedge dy \wedge dz), dt)\sigma,$$

поэтому

$$\star(dx \wedge dy \wedge dz) = -dt.$$

Рассмотрим 2-формы:

$$(dt \wedge dx) \wedge (dy \wedge dz) = (\star(dt \wedge dx), (dy \wedge dz))\sigma,$$

тогда

$$\star(dt \wedge dx) = dy \wedge dz.$$

Аналогично

$$\star(dt \wedge dy) = -dx \wedge dz,$$

$$\star(dt \wedge dz) = dx \wedge dy.$$

Далее,

$$(dx \wedge dy) \wedge (dt \wedge dz) = (\star(dx \wedge dy), dt \wedge dz)\sigma.$$

Тогда

$$\star(dx \wedge dy) = -dt \wedge dz.$$

Аналогично

$$\star(dx \wedge dz) = dt \wedge dy,$$

$$\star(dy \wedge dz) = -dt \wedge dx.$$

Теперь рассмотрим 1-формы:

$$dt \wedge (dx \wedge dy \wedge dz) = (\star dt, dx \wedge dy \wedge dz)\sigma.$$

Таким образом,

$$\star dt = -dx \wedge dy \wedge dz.$$

Кроме того,

$$dx \wedge (dt \wedge dy \wedge dz) = (\star dx, dt \wedge dy \wedge dz)\sigma,$$

так что

$$\star dx = -dt \wedge dy \wedge dz.$$

Аналогично

$$\star dy = dt \wedge dx \wedge dz,$$

$$\star dz = -dt \wedge dx \wedge dy.$$

И наконец,

$$\star 1 = -dt \wedge dx \wedge dy \wedge dz.$$

Очевидно, что в приведенных примерах  $\star(\star\lambda) = \pm\lambda$ . Чтобы сформулировать общее правило, в пространстве  $V$  выберем ортогональный базис  $\mathbf{e}_1, \dots, \mathbf{e}_n$ , при условии, что  $(\mathbf{e}_i, \mathbf{e}_i) = \pm 1$ . Рассмотрим случай, когда  $\lambda \in \Lambda^k(V^*)$  является произведением  $k$  векторов  $\mathbf{e}_i^*$ , и  $\omega$  является произведением оставшихся  $(n - k)$  векторов  $\mathbf{e}_j^*$  с некоторым знаком  $\pm$ . Допустим, что  $\lambda$  и  $\omega$  таковы, что  $\lambda \wedge \omega = \sigma$  (например,  $\lambda = dx$ ,  $\omega = dy \wedge dz$  в пространстве  $\mathbb{R}^3$ ). Тогда  $\star\lambda = \pm\omega$  и  $\star\omega = \pm\lambda$ . Чтобы определить знаки, положим  $\star\lambda = c_1\omega$  и  $\star\omega = c_2\lambda$ . Тогда

$$\sigma = \lambda \wedge \omega = (\star\lambda, \omega)\sigma.$$

Откуда

$$(\star\lambda, \omega) = c_1(\omega, \omega) = 1.$$

Кроме того, согласно правилу перестановки множителей во внешнем произведении,

$$(-1)^{k(n-k)}\sigma = \omega \wedge \lambda = (\star\omega, \lambda)\sigma,$$

откуда

$$(\star\omega, \lambda) = c_2(\lambda, \lambda) = (-1)^{k(n-k)}.$$

Поскольку  $\lambda$  и  $\omega$  являются базисными элементами, мы имеем

$$(\lambda, \lambda)(\omega, \omega) = (\lambda \wedge \omega, \lambda \wedge \omega) = (\sigma, \sigma).$$

Откуда

$$c_1 c_2(\sigma, \sigma) = (-1)^{k(n-k)}.$$

Поскольку  $(\sigma, \sigma) = \pm 1$ , мы можем перенести этот множитель в другую часть уравнения. Тогда получим

$$\star(\star\lambda) = c_1 \star\omega = c_1 c_2 \lambda,$$

где

$$c_1 c_2 = (-1)^{k(n-k)}(\sigma, \sigma).$$

В силу линейности оператора  $\star$ , если все выписанные уравнения выполняются для каждого базисного элемента  $\Lambda^k(V^*)$ , то они должны выполняться и для всех элементов  $\Lambda^k(V^*)$ . Итак, мы доказали, что

$$\boxed{\star\star = (-1)^{k(n-k)}(\sigma, \sigma) \text{ на } \Lambda^k(V^*)}.$$

Если  $n$  — нечетное число и  $(\sigma, \sigma) = +1$ , что справедливо для пространства  $\mathbb{R}^3$  с евклидовым скалярным произведением, то  $\star(\star\lambda) = \lambda$ . Однако, в общем случае это не выполняется.

Используя только что полученный результат, т. е. что

$$\star\star\omega = (-1)^{k(n-k)}(\sigma, \sigma)\omega,$$

мы можем вывести полезную формулу для скалярного произведения двух любых  $k$ -форм. Сначала заметим, что поскольку  $\star\star\omega = \pm\omega$ , то *четыре* последовательных применения звездного оператора приводят к тождеству, т. е.

$$\star\star\star\star\omega = \omega.$$

По определению звездного оператора

$$(\star\star\star\star\lambda, \omega)\sigma = \star\star(\star\lambda \wedge \omega) = (-1)^{k(n-k)}(\sigma, \sigma)\star\lambda \wedge \omega$$

выполняется для двух любых  $k$ -форм  $\lambda$  и  $\omega$ . Следовательно,

$$\begin{aligned}(\lambda, \omega)\sigma &= (-1)^{k(n-k)}(\sigma, \sigma)\star\lambda \wedge \omega \\ &= \omega \wedge \star\lambda(\sigma, \sigma).\end{aligned}$$

Однако,

$$\star\star\star\sigma = \star\star(1) = (\sigma, \sigma)1.$$

Действуя оператором  $\star\star\star$  на обе части уравнения, получаем

$$(\lambda, \omega)(\sigma, \sigma)1 = \star\star\star(\omega \wedge \star\lambda)(\sigma, \sigma)1$$

или

$$(\lambda, \omega) = \star\star\star(\omega \wedge \star\lambda) = \star\star\star(\lambda \wedge \star\omega).$$

Еще один полезный результат получается из

$$(\star\lambda, \star\omega)\sigma = \lambda \wedge \star\omega.$$

Действуем оператором  $\star$  на это уравнение и получаем

$$(\star\lambda, \star\omega) = \star(\lambda \wedge \star\omega) = \star(\omega \wedge \star\lambda).$$

С помощью полученных здесь формул скалярные произведения могут быть записаны через звездный оператор.

### 18.3. Интеграл Дирихле и лапласиан

Пусть  $\omega$  обозначает  $k$ -форму, а  $\lambda$  —  $(k-1)$ -форму. Тогда

$$d(\lambda \wedge \star\omega) = d\lambda \wedge \star\omega + (-1)^{k-1}\lambda \wedge d(\star\omega).$$

Поэтому для любой (ограниченной) области  $U$  мы имеем

$$\int_U d\lambda \wedge \star\omega + (-1)^{k-1} \int_U \lambda \wedge d(\star\omega) = \int_U d(\lambda \wedge \star\omega).$$

Предположим, что  $\lambda$  или  $\omega$  обращаются в нуль на границе  $\partial U$  (в том числе если  $U$  неограничена и если  $\lambda$  или  $\omega$  имеет компактный

носитель внутри  $U$  или если они стремятся к нулю на бесконечности достаточно быстро). Тогда по теореме Стокса имеем

$$\int_U d\lambda \wedge \star\omega + \int_U (-1)^{k-1} \lambda \wedge d(\star\omega) = 0.$$

Действуя  $\star$  на формулу  $(\rho, \tau) = \star\star\star(\rho \wedge \star\tau)$ , мы видим, что

$$d\lambda \wedge \star\omega = (d\lambda, \omega) \star 1.$$

С другой стороны,  $\omega$  является  $k$ -формой, поэтому  $\star\omega$  является  $(n-k)$ -формой и, следовательно,  $d\star\omega$  является  $(n-k+1)$ -формой. Тогда

$$\begin{aligned} \lambda \wedge d\star\omega &= (-1)^{(k-1)(n-k+1)} d\star\omega \wedge \lambda \\ &= (-1)^{(k-1)(n-k+1)} (\star^{-1} d\star\omega, \lambda) \sigma \\ &= (-1)^{(k-1)(n-k+1)} (\lambda, \star^{-1} d\star\omega) \sigma. \end{aligned}$$

Далее,  $\star d\star\omega$  является формой степени  $n - (n - k + 1) = k - 1$ . Тогда поскольку  $\star\star\star\star = \text{id}$ , получаем

$$\star^{-1} d\star\omega = \star\star(\star d\star\omega) = (-1)^{(k-1)(n-k+1)} (\sigma, \sigma) \star d\star\omega$$

и

$$\sigma = (\sigma, \sigma) \star 1.$$

Таким образом,

$$\lambda \wedge d\star\omega = (\lambda, \star\star\star d\star\omega) \star 1.$$

Итак, мы можем написать

$$d\lambda \wedge \star\omega + (-1)^{k-1} \lambda \wedge d\star\omega = [(d\lambda, \omega) + (-1)^{(k-1)} (\lambda, \star^{-1} d\star\omega)] \star 1.$$

Значит, если  $\lambda$  или  $\omega$  обращаются в нуль на границе  $\partial U$  (если область  $U$  ограничена) или имеют компактный носитель (или убывают достаточно быстро на бесконечности), то

$$\int_U (d\lambda, \omega) = \int_U (\lambda, (-1)^k \star\star\star d\star\omega) \sigma.$$

Определим оператор  $d^*$  согласно формуле

$$d^*\omega = (-1)^k \star \star d \star \omega = (-1)^k \star^{-1} d \star \omega.$$

Таким образом,  $d^*$  отображает  $k$ -формы в  $(k-1)$ -формы. Для произвольных  $k$ -форм  $\lambda$  и  $\tau$  давайте напишем

$$(\lambda, \tau)_U = D_U(\lambda, \tau)\sigma = \int_U (\lambda, \tau)\sigma.$$

В этом случае получим

$$D_U(d\lambda, \omega) = D_U(\lambda, d^*\omega)$$

для  $\lambda$  или  $\omega$ , равных нулю на границе  $\partial U$  (или достаточно быстро убывающих на бесконечности).

Заметим, что

$$d^*d^* = 0,$$

поскольку

$$dd = 0.$$

Давайте определим оператор Лапласа  $\square$ , отображающий  $k$ -формы в  $k$ -формы, формулой<sup>1</sup>

$$\square = dd^* + d^*d.$$

Утверждается, что

$$\square d = d\square,$$

$$\square d^* = d^*\square,$$

$$\square \star = \star \square.$$

Чтобы установить равенство  $\square d = d\square$ , заметим, что

$$\square d = (dd^* + d^*d)d = dd^*d$$

и

$$d\square = d(dd^* + d^*d) = dd^*d,$$

<sup>1</sup>Этот оператор  $\square$  отличается от лапласиана, определенного в предыдущих главах, знаком минус.

поскольку  $dd = 0$ . Аналогичное вычисление показывает, что

$$\square d^* = d^* \square.$$

Доказательство равенства  $\square \star = \star \square$  чуть сложнее, потому что необходимо следить за знаками. Начнем с равенства

$$d^* \omega = (-1)^k \star \star \star d \star \omega \quad \text{для } k\text{-формы } \omega.$$

Заменив  $k$ -форму  $\omega$  на  $(n - k)$ -форму  $\star \omega$ , мы имеем

$$d^* \omega = (-1)^{n-k} \star \star \star d \star \star \omega.$$

Поскольку  $d(\star \star \omega)$  является  $(k + 1)$ -формой,

$$\star \star d(\star \star \omega) = (-1)^{(k+1)(n-k-1)} (\sigma, \sigma) d(\star \star \omega).$$

А поскольку  $\omega$  является  $k$ -формой,

$$\star \star \omega = (-1)^{k(n-k)} (\sigma, \sigma) \omega.$$

Собирая вместе эти результаты и учитывая, что  $(\sigma, \sigma)^2 = 1$ , мы получаем

$$d^* \star \omega = (-1)^{n-k} (-1)^{(k+1)(n-k-1)} (-1)^{k(n-k)} \star d \omega$$

или

$$d^* \star \omega = (-1)^{(k+1)(n-k-1)} (-1)^{(k+1)(n-k)} \star d \omega = (-1)^{k-1} \star d \omega.$$

С этим результатом и формулой  $\star d^* \omega = (-1)^k d \star \omega$  доказательство становится совсем простым:

$$\begin{aligned} \star \square \omega &= \star d(d^* \omega) + \star d^*(d \omega), \\ \star \square \omega &= (-1)^k d^* d \omega + (-1)^{k+1} d \star d \omega. \end{aligned}$$

В то же время

$$\begin{aligned} \square \star \omega &= dd^* \star \omega = d^* d \star \omega, \\ \square \star \omega &= (-1)^{k+1} d \star d \omega + (-1)^k d \star d^* \omega. \end{aligned}$$

Следовательно,

$$\star \square = \square \star.$$

**Оператор  $\square$  в пространстве  $\mathbb{R}^3$ .** Вычислим явный вид лапласиана в пространстве  $\mathbb{R}^3$ . Для функции  $f$  имеем

$$\square f = dd^*f + d^*df.$$

Однако,  $\star f$  является 3-формой и  $d\Omega = 0$  для любой 3-формы в  $\mathbb{R}^3$ . Поэтому  $\square f = d^*df$ . Далее,

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz$$

и

$$\star \frac{\partial f}{\partial x} dx = \frac{\partial f}{\partial x} dy \wedge dz.$$

Отсюда

$$d\star \frac{\partial f}{\partial x} dx = \frac{\partial^2 f}{\partial x^2} dx \wedge dy \wedge dz$$

и

$$d^* \left( \frac{\partial f}{\partial x} dx \right) = -\star^{-1} d\star \frac{\partial f}{\partial x} dx = -\frac{\partial^2 f}{\partial x^2}.$$

Проделав аналогичные вычисления для остальных двух членов, получаем

$$\square f = - \left( \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right).$$

Давайте вычислим лапласиан для 1-форм. В этом случае достаточно вычислить  $\square \omega$  для  $\omega = a dx$ . Оператор  $\square$  определяется только скалярным произведением и ориентацией. Поэтому он коммутирует со всеми вращениями, и оси можно переставлять. Тогда, если мы получим формулу для  $\square(a dx)$ , то аналогичная формула получится для  $\square(b dy)$  и  $\square(c dz)$ . В силу линейности мы будем иметь формулу и для  $a dx + b dy + c dz$ . Итак,

$$d(a dx) = da \wedge dx = \frac{\partial a}{\partial y} dy \wedge dx + \frac{\partial a}{\partial z} dz \wedge dx.$$

Поэтому

$$\star(d(a dx)) = -\frac{\partial a}{\partial y} dz + \frac{\partial a}{\partial z} dy.$$

Итак,

$$d\star d(ax) = -\frac{\partial^2 a}{\partial x \partial y} dx \wedge dz - \frac{\partial^2 a}{\partial y^2} dy \wedge dz + \frac{\partial^2 a}{\partial x \partial z} dx \wedge dy + \frac{\partial^2 a}{\partial z^2} dz \wedge dy.$$

Следовательно, используя свойство  $\star = \star^{-1}$  для 2-форм в пространстве  $\mathbb{R}^3$ , получаем

$$d^* d(ax) = \star d\star d(ax) = \frac{\partial^2 a}{\partial x \partial y} dy - \frac{\partial^2 a}{\partial y^2} dx + \frac{\partial^2 a}{\partial x \partial z} dz - \frac{\partial^2 a}{\partial z^2} dx.$$

Далее,

$$\begin{aligned} d^*(ax) &= -\star d\star(ax) = -\star d(ax \wedge dy \wedge dz) \\ &= -\star \left( \frac{\partial a}{\partial x} dx \wedge dy \wedge dz \right) \\ &= -\frac{\partial a}{\partial x}. \end{aligned}$$

Итак,

$$dd^*(ax) = -\frac{\partial^2 a}{\partial x^2} dx - \frac{\partial^2 a}{\partial x \partial y} dy - \frac{\partial^2 a}{\partial x \partial z} dz.$$

Отсюда следует, что

$$\begin{aligned} \square(ax) &= (\square a)dx \\ &= -\left( \frac{\partial^2 a}{\partial x^2} + \frac{\partial^2 a}{\partial y^2} + \frac{\partial^2 a}{\partial z^2} \right) dx. \end{aligned}$$

Таким образом, если  $\omega = a dx + b dy + c dz$ , то мы можем вычислить  $\square \omega$ , подействовав оператором  $\square$  на каждый член.

Произвольная 2-форма в пространстве  $\mathbb{R}^3$  может быть записана как  $\star \omega$ , где  $\omega$  является 1-формой. Поскольку  $\star \square = \square \star$ , мы видим, что для 2-формы

$$\Omega = A dx \wedge dy + B dz \wedge dx + C dy \wedge dz$$

получается тот же результат

$$\square \Omega = (\square A) dx \wedge dy + (\square B) dz \wedge dx + (\square C) dy \wedge dz.$$

Аналогично для 3-формы

$$\square(\rho dx \wedge dy \wedge dz) = (\square\rho) dx \wedge dy \wedge dz.$$

Итак, в обычных прямоугольных координатах в пространстве  $\mathbb{R}^3$  мы можем вычислить  $\square$  для любой формы, действуя оператором  $\square$  на каждый коэффициент.

А теперь давайте повторим рассуждения из последнего параграфа главы 17 — сведение уравнений магнитостатики к уравнению Пуассона. Предположим, что выбраны такие единицы, для которых  $\mu = 1$ . Тогда положим

$$j = -\star J.$$

Теперь уравнения можно записать в виде

$$\begin{aligned} dB &= 0, \\ d^*B &= 4\pi j. \end{aligned}$$

Выберем такое  $A'$ , чтобы

$$dA' = B.$$

(Это возможно, так как мы предположили, что  $H^1 = 0$ .) Решаем уравнение Пуассона

$$\square\psi = d^*A'.$$

Заменяем  $A'$  на

$$A = A' - d\psi$$

и получаем

$$dA = B.$$

Поскольку  $\psi$  является функцией, то

$$\square\psi = d^*d\psi = d^*A'.$$

Следовательно,

$$d^*A = d^*A' - d^*d\psi = 0.$$

Итак наше второе уравнение принимает вид

$$d^*dA = 4\pi j,$$

или, поскольку  $d^*A = 0$ ,

$$\square A = 4\pi j.$$

Кажется, будто в этом доказательстве мы использовали тот факт, что  $B$  является 2-формой,  $A$  — 1-формой, а  $\psi$  — функцией. На самом деле это верно в общем случае. Пусть  $B$  является  $k$ -формой в  $n$ -мерном пространстве. Мы хотим решить уравнения

$$\begin{aligned} dB &= 0, \\ d^*B &= 4\pi j. \end{aligned}$$

Предположим, что рассматриваемая область такова, что  $B = dA'$  для некоторой  $A'$ . Тогда решаем уравнение Пуассона

$$\square \psi = d^*A'.$$

Напишем

$$A = A' - d\psi.$$

Тогда

$$d^*A = dd^*\psi.$$

Следовательно,

$$dd^*A = 0$$

и

$$\square A = (dd^* + d^*d)A = 4\pi j.$$

Таким образом, мы опять свели задачу к решению уравнения Пуассона.

## 18.4. Оператор $\square$ в четырехмерном пространстве-времени

Рассмотрим четырехмерное пространство-время с координатами  $t, x, y, z$  и скалярным произведением, для которого выполняются

условия  $(dt, dt) = 1$ ,  $(dx, dx) = (dy, dy) = (dz, dz) = -1$ . Подействуем оператором  $\square$  на функцию

$$\begin{aligned}\square f &= d^* df \\ &= -\star^{-1} d\star \left( \frac{\partial f}{\partial t} dt + \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz \right) \\ &= -\star^{-1} d \left( -\frac{\partial f}{\partial t} dx \wedge dy \wedge dz - \frac{\partial f}{\partial x} dt \wedge dy \wedge dz \right. \\ &\quad \left. + \frac{\partial f}{\partial y} dt \wedge dx \wedge dz - \frac{\partial f}{\partial z} dt \wedge dx \wedge dy \right) \\ &= -\star^{-1} \left( -\frac{\partial^2 f}{\partial t^2} + \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \right) dt \wedge dx \wedge dy \wedge dz.\end{aligned}$$

Откуда

$$\square f = -\frac{\partial^2 f}{\partial t^2} + \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

В качестве 1-формы сначала рассмотрим  $a dt$ . Получим

$$d(a dt) = \frac{\partial a}{\partial x} dx \wedge dt + \frac{\partial a}{\partial y} dy \wedge dt + \frac{\partial a}{\partial z} dz \wedge dt,$$

тогда

$$\star d(a dt) = -\frac{\partial a}{\partial x} dy \wedge dz + \frac{\partial a}{\partial y} dx \wedge dz - \frac{\partial a}{\partial z} dx \wedge dy$$

и

$$\begin{aligned}d\star d(a dt) &= \frac{\partial^2 a}{\partial t \partial x} dt \wedge dy \wedge dz \\ &\quad + \frac{\partial^2 a}{\partial y \partial t} dt \wedge dx \wedge dz - \frac{\partial^2 a}{\partial y^2} dx \wedge dy \wedge dz \\ &\quad - \frac{\partial^2 a}{\partial z \partial t} dt \wedge dx \wedge dy - \frac{\partial^2 a}{\partial z^2} dx \wedge dy \wedge dz.\end{aligned}$$

Следовательно,

$$\begin{aligned}d^* d(a dt) &= \star^{-1} d\star d(a dt) \\ &= \left( \frac{\partial^2 a}{\partial x^2} + \frac{\partial^2 a}{\partial y^2} + \frac{\partial^2 a}{\partial z^2} \right) dt + \frac{\partial^2 a}{\partial x \partial t} dx + \frac{\partial^2 a}{\partial y \partial t} dy + \frac{\partial^2 a}{\partial z \partial t} dz.\end{aligned}$$

С другой стороны,

$$\begin{aligned}
 d^*(a dt) &= -\star^{-1} d\star(a dt) \\
 &= \star^{-1} d(a dx \wedge dy \wedge dz) \\
 &= \star^{-1} \frac{\partial a}{\partial t} dt \wedge dx \wedge dy \wedge dz \\
 &= -\frac{\partial a}{\partial t}.
 \end{aligned}$$

Откуда

$$dd^*(a dt) = -\frac{\partial^2 a}{\partial t^2} dt - \frac{\partial^2 a}{\partial t \partial x} dx - \frac{\partial^2 a}{\partial t \partial y} dy - \frac{\partial^2 a}{\partial t \partial z} dz.$$

Итак, получаем

$$\square(a dt) = (\square a) dt = \left( -\frac{\partial^2 a}{\partial t^2} + \frac{\partial^2 a}{\partial x^2} + \frac{\partial^2 a}{\partial y^2} + \frac{\partial^2 a}{\partial z^2} \right) dt.$$

Аналогичное вычисление (или использование вращательной инвариантности<sup>2</sup>) показывает, что

$$\begin{aligned}
 \square(b dx) &= (\square b) dx, \\
 \square(c dy) &= (\square c) dy,
 \end{aligned}$$

и так далее.

То же самое выполняется и для 2-форм. Например, рассмотрим  $a dt \wedge dx$ . Получим

$$\begin{aligned}
 d^*(a dt \wedge dx) &= \star^{-1} d\star(a dt \wedge dx) \\
 &= \star^{-1} d(a dy \wedge dz) \\
 &= \star^{-1} \left( \frac{\partial a}{\partial t} dt \wedge dy \wedge dz \right) + \star^{-1} \left( \frac{\partial a}{\partial x} dx \wedge dy \wedge dz \right) \\
 &= -\frac{\partial a}{\partial t} dx - \frac{\partial a}{\partial x} dt.
 \end{aligned}$$

---

<sup>2</sup>Здесь под вращениями следует понимать собственные преобразования Лоренца, сохраняющие лоренцево скалярное произведение и ориентацию пространства-времени. — *Прим. ред.*

Откуда

$$dd^*(a dt) = \left( -\frac{\partial^2 a}{\partial t^2} + \frac{\partial^2 a}{\partial x^2} \right) dt \wedge dx - \frac{\partial^2 a}{\partial t \partial y} dy \wedge dx - \frac{\partial^2 a}{\partial t \partial z} dz \wedge dx \\ - \frac{\partial^2 a}{\partial x \partial y} dt \wedge dy - \frac{\partial^2 a}{\partial x \partial z} dt \wedge dz.$$

С другой стороны,

$$d(a dt \wedge dx) = \frac{\partial a}{\partial y} dy \wedge dt \wedge dx + \frac{\partial a}{\partial z} dz \wedge dt \wedge dx,$$

поэтому

$$\star d(a dt \wedge dx) = -\frac{\partial a}{\partial y} dz + \frac{\partial a}{\partial z} dy$$

и

$$d\star d(a dt \wedge dx) = -\left( \frac{\partial^2 a}{\partial y^2} + \frac{\partial^2 a}{\partial z^2} \right) dy \wedge dz - \frac{\partial^2 a}{\partial y \partial x} dx \wedge dz \\ - \frac{\partial^2 a}{\partial y \partial t} dt \wedge dz + \frac{\partial^2 a}{\partial z \partial t} dt \wedge dy + \frac{\partial^2 a}{\partial x \partial z} dx \wedge dy.$$

Тогда

$$d^* d(a dt \wedge dx) = -\star^{-1} d\star d(a dt \wedge dx)$$

и

$$\square(a dt \wedge dx) = (\square a) dt \wedge dx = \left( -\frac{\partial^2 a}{\partial t^2} + \frac{\partial^2 a}{\partial x^2} + \frac{\partial^2 a}{\partial y^2} + \frac{\partial^2 a}{\partial z^2} \right) dt \wedge dx.$$

Мы предоставляем читателю самостоятельно проверить, что

$$\square(b dx \wedge dz) = (\square b) dx \wedge dz.$$

Итак, мы получили, что для любой 2-формы, записанной через  $dt$ ,  $dx$ ,  $dy$ ,  $dz$ , действие оператора  $\square$  равносильно действию этого оператора на коэффициенты. Справедливость аналогичного утверждения для 3- и 4-форм вытекает теперь из рассмотренных случаев 0-, 1- и 2-форм и равенства  $\star \square = \square \star$ .

## 18.5. Алгебра Клиффорда

Теперь читатель уже, должно быть, догадался, что в пространстве  $\mathbb{R}^n$  мы можем вычислить действие оператора  $\square$ , записанного в *евклидовых координатах*, на произвольную  $k$ -форму как действие этого оператора на каждый коэффициент. Кроме того, уже понятно, что если мы попробуем доказать этот результат теми же методами, которые использовались в последних двух параграфах, то просто запутаемся. Цель этого параграфа — ввести алгебраический формализм, который позволит нам доказать все очень просто. В этой книге нам больше не понадобятся результаты данного параграфа — ведь мы уже доказали результаты в важных случаях трех и четырех измерений. Однако, излагаемые здесь идеи оказались очень важными в современной физике.

Наше доказательство будет основано на следующем свойстве

$$D_U(d\omega, \tau) = D_U(\omega, d^*\tau) + \text{граничные члены}.$$

Это свойство можно перефразировать, говоря, что  $d^*$  является *формальным сопряжением*  $d$ . Слово «формальный» означает «по модулю граничных членов». У нас есть равенство

$$D_U(d\omega, \tau) = D_U(\omega, d^*\tau),$$

только когда делаются различные предположения, которые гарантируют отсутствие вклада от границы. Давайте рассмотрим эту ситуацию в более общем случае.

Пусть  $E$  и  $F$  обозначают векторные пространства, в каждом из которых имеется свое скалярное произведение. Тогда мы можем идентифицировать  $E$  с  $E^*$  и  $F$  с  $F^*$ . В частности, если  $A : E \rightarrow F$  — линейное отображение, то мы можем рассматривать  $A^* : F^* \rightarrow E^*$  как линейное отображение  $A^* : F \rightarrow E$ , определенное условием

$$(e, A^*f)_E = (Ae, f)_F,$$

где  $(\cdot, \cdot)_E$  и  $(\cdot, \cdot)_F$  обозначают скалярные произведения в пространствах  $E$  и  $F$  соответственно. Теперь предположим, что

в векторном пространстве  $V$  мы выбрали систему координат  $x_1, \dots, x_n$ . Тогда соответствующая форма объема имеет вид

$$\sigma = dx_1 \wedge \dots \wedge dx_n.$$

Рассмотрим пространство гладких функций из  $V$  в  $E$ , обозначаемое  $\mathcal{F}(V, E)$ . Если  $u_1$  и  $u_2$  — такие функции, а  $U$  — ограниченная область, то можно определить

$$D_U(u_1, u_2) = \int_U (u_1, u_2)_E \sigma.$$

(Если область  $U$  не ограничена, то все равно мы можем определить  $D_U$  для тех пар функций, которые достаточно быстро убывают на бесконечности.) Можно рассмотреть операторы  $L$  из  $\mathcal{F}(V, E)$  в  $\mathcal{F}(V, F)$ . Тогда каждой функции  $u \in \mathcal{F}(V, E)$  ставится в соответствие элемент  $Lu$  пространства  $\mathcal{F}(V, F)$ . Оператор  $L$  называется *линейным дифференциальным оператором первого порядка*, если существуют такие линейные отображения  $A_1 : E \rightarrow F$ ,  $A_2 : E \rightarrow F$ , ...,  $A_n : E \rightarrow F$  и  $B : E \rightarrow F$ , что

$$Lu = A_1 \frac{\partial u}{\partial x_1} + \dots + A_n \frac{\partial u}{\partial x_n} + Bu.$$

В общем случае отображения  $A_1, \dots, A_n$  и  $B$  могут зависеть от точки в пространстве  $V$ , т. е.  $A_i$  и  $B$  — это функции в пространстве  $V$ , значения которых являются линейными отображениями. Тогда предыдущее уравнение можно записать в форме

$$(Lu)(x) = A_1(x) \frac{\partial u}{\partial x_1}(x) + \dots + A_n(x) \frac{\partial u}{\partial x_n}(x) + B(x)u(x).$$

Если  $A_i$  и  $B$  постоянны, то мы говорим, что  $L$  является линейным дифференциальным оператором первого порядка с постоянными коэффициентами.

Теперь пусть  $u \in \mathcal{F}(V, E)$  и  $v \in \mathcal{F}(V, E)$ . Тогда

$$\begin{aligned} \frac{\partial}{\partial x_i} (A_i u, v)_F &= \left( \frac{\partial A_i}{\partial x_i} u, v \right)_F + \left( A_i \frac{\partial u}{\partial x_i}, v \right)_F + \left( A_i u, \frac{\partial v}{\partial x_i} \right)_F \\ &= \left( A_i \frac{\partial u}{\partial x_i}, v \right)_F + \left( u, \left( \frac{\partial A_i}{\partial x_i} \right)^* v \right)_E + \left( u, A_i^* \frac{\partial v}{\partial x_i} \right)_E. \end{aligned}$$

Поэтому

$$\begin{aligned} \left( A_i \frac{\partial u}{\partial x_i}, v \right)_F \\ = \frac{\partial}{\partial x_i} (A_i u, v)_F - \left( u, A_i^* \frac{\partial v}{\partial x_i} \right)_E - \left( u, \left( \frac{\partial A_i}{\partial x_i} \right)^*, v \right)_E. \end{aligned}$$

Итак, если мы определяем

$$L^* v = -A_1^* \frac{\partial v}{\partial x_1} - A_2^* \frac{\partial v}{\partial x_2} - \dots - A_n^* \frac{\partial v}{\partial x_n} + \left( B^* - \sum \left( \frac{\partial A_i}{\partial x_i} \right)^* \right) v,$$

то

$$(Lu, v)_F = (u, L^* v)_E + \sum \frac{\partial}{\partial x_i} (A_i u, v)_F.$$

Последний член справа сводится к умножению граничного члена на  $\sigma$  и интегрированию. Действительно, если мы напишем

$$\begin{aligned} \tau = \frac{\partial}{\partial x_1} (A_1 u, v)_F dx_2 \wedge \dots \wedge dx_n \\ - \frac{\partial}{\partial x_2} (A_2 u, v)_F dx_1 \wedge dx_3 \wedge \dots \wedge dx_n + \dots, \end{aligned}$$

то

$$d\tau = \sum \frac{\partial}{\partial x_i} (A_i u, v)_F dx_1 \wedge \dots \wedge dx_n.$$

Тогда

$$\int_U (Lu, v)_F \sigma = \int_U (u, L^* v)_E \sigma + \int_{\partial U} \tau.$$

Таким образом, оператор  $L^*$  является формально сопряженным оператора  $L$ .

Строго говоря,  $k$ -форма  $\omega$  не является функцией из пространства  $V$  в фиксированное векторное пространство. (Потому что  $\omega(x)$  является элементом пространства кососимметричных полилинейных  $k$ -функций на касательном пространстве в точке  $x$ , а касательное пространство изменяется от точки к точке.) Однако, как только мы выбрали систему координат,  $k$ -форму  $\omega$  можно рассматривать как функцию из  $V$  в  $\Lambda^k(V^*)$ . Таким образом, мы

можем считать, что  $E = \Lambda^k(V^*)$ . Если мы хотим в качестве дифференциального оператора  $L$  взять  $d$ , то полагаем  $F = \Lambda^{k+1}(V^*)$ . Тогда можно написать

$$d = \varepsilon_1 \frac{\partial}{\partial x_1} + \cdots + \varepsilon_n \frac{\partial}{\partial x_n},$$

где  $\varepsilon_k$  обозначает линейное отображение  $\Lambda^k(V^*) \rightarrow \Lambda^{k+1}(V^*)$ , состоящее из внешнего умножения на  $dx_k$ , т. е.

$$\varepsilon_k \omega = dx_k \wedge \omega.$$

Теперь для вычисления  $d^*$  нам нужна формула для  $\varepsilon_k^*$ . Утверждается, что

$$\varepsilon_k^* = \pm i \left( \frac{\partial}{\partial x_k} \right), \quad \text{причем знак соответствует } (dx_k, dx_k) = \pm 1.$$

Чтобы это доказать, предположим, что  $k = 1$  (это просто вопрос обозначений) и исследуем, что происходит с базисными элементами

$$\omega = dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

пространства  $\Lambda^k(V)$  и базисом  $\Lambda^{k+1}(V^*)$ . Существует две возможности.

(а)  $i_1 = 1$ . Тогда  $\varepsilon_1 \omega = 0$ . С другой стороны, для

$$\tau = dx_{j_1} \wedge \cdots \wedge dx_{j_n}$$

мы имеем

$$i \left( \frac{\partial}{\partial x_1} \right) \tau = 0 \quad (\text{если } j_1 > 1)$$

и

$$i \left( \frac{\partial}{\partial x_1} \right) \tau = dx_{j_2} \wedge \cdots \wedge dx_{j_n} \quad (\text{если } j_1 = 1).$$

В любом случае мы имеем

$$(\varepsilon_1 \omega, \tau) = \left( \omega, i \left( \frac{\partial}{\partial x_1} \right) \tau \right) = 0.$$

(b)  $i_1 > 1$ . Тогда

$$\varepsilon_1 \omega = dx_1 \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

и

$$i\left(\frac{\partial}{\partial x_1}\right)\tau = \begin{cases} 0, & \text{если } j_1 > 1, \\ dx_{j_2} \wedge \cdots \wedge dx_{j_{k+1}}, & \text{если } j_1 = 1. \end{cases}$$

В этом случае

$$(\varepsilon_1 \omega, \tau) = 0 = \left(\omega, i\left(\frac{\partial}{\partial x_1}\right)\tau\right), \quad \text{если } j_1 > 1.$$

Если же  $j_1 = 1$ , то мы по-прежнему имеем

$$(\varepsilon_1 \omega, \tau) = 0 = \left(\omega, i\left(\frac{\partial}{\partial x_1}\right)\tau\right),$$

кроме случаев, когда  $j_2 = i_1$ ,  $j_3 = i_2$  и так далее. Если  $j_2 = i_1$ ,  $j_3 = i_2$  и т. д., то

$$(\varepsilon_1 \omega, \tau) = (dx_1, dx_1)(\omega, \omega),$$

а

$$\left(\omega, i\left(\frac{\partial}{\partial x_1}\right)\tau\right) = (\omega, \omega).$$

Это доказывает, что

$$\varepsilon_1^* = (dx_1, dx_1)i\left(\frac{\partial}{\partial x_1}\right).$$

В общем случае доказательство показывает следующее. Пусть  $L: V \rightarrow V^*$  обозначает линейное отображение, определенное скалярным произведением в пространстве  $V$ . Пусть  $\theta$  — произвольная 1-форма и  $\varepsilon(\theta)$  — внешнее умножение на  $\theta$ . Тогда

$$(\varepsilon(\theta))^* = i(L^{-1}\theta).$$

Определим

$$\iota_p = \pm i\left(\frac{\partial}{\partial x_p}\right) = \varepsilon_p^*.$$

Мы показали, что

$$d^* = - \sum_{p=1}^n \iota_p \frac{\partial}{\partial x_p}.$$

Из антикоммутативности внешнего произведения следует, что

$$\varepsilon_p \varepsilon_q + \varepsilon_q \varepsilon_p = 0 \quad \text{для всех } p \text{ и } q.$$

Следовательно (или непосредственно из свойств внутреннего произведения),

$$\iota_p \iota_q + \iota_q \iota_p = 0 \quad \text{для всех } p \text{ и } q.$$

Если  $p \neq q$ , то внешнее умножение на  $dx_p$  и внутреннее умножение на  $i(\partial/\partial x_q)$  удовлетворяют равенству

$$i\left(\frac{\partial}{\partial x_q}\right)\varepsilon(dx_p) + \varepsilon(dx_p)i\left(\frac{\partial}{\partial x_q}\right) = 0.$$

Действительно, достаточно проверить это равенство для  $p = 1$ ,  $q = 2$ . Это простое вычисление мы оставляем читателю. С другой стороны, для базисных элементов  $\omega$  мы имеем

$$i\left(\frac{\partial}{\partial x_1}\right)\varepsilon(dx_1)\omega = \begin{cases} 0, & \text{если } \omega = dx_1 \wedge \dots, \\ \omega, & \text{если } \omega = dx_2 \wedge \dots, \end{cases}$$

а

$$\varepsilon(dx_1)i\left(\frac{\partial}{\partial x_1}\right)\omega = \begin{cases} \omega, & \text{если } \omega = dx_1 \wedge \dots, \\ 0, & \text{если } \omega = dx_2 \wedge \dots. \end{cases}$$

Таким образом,

$$i\left(\frac{\partial}{\partial x_1}\right)\varepsilon(dx_1) + \varepsilon(dx_1)i\left(\frac{\partial}{\partial x_1}\right) = \text{id}.$$

Естественно обобщая со случая  $x_1$  на случай  $x_p$ , можно сделать вывод, что

$$\iota_p \varepsilon_p + \varepsilon_p \iota_p = \pm \text{id}, \quad \text{где } \pm \text{ соответствует знаку } (dx_p, dx_p).$$

И наконец, поскольку операторы  $\iota_p$  и  $\varepsilon_q$  постоянны (в нашей системе координат), то мы имеем

$$\begin{aligned} dd^* + d^*d &= - \left( \sum \varepsilon_p \frac{\partial}{\partial x_p} \right) \left( \sum \iota_q \frac{\partial}{\partial x_q} \right) \\ &\quad - \left( \sum \iota_q \frac{\partial}{\partial x_q} \right) \left( \sum \varepsilon_p \frac{\partial}{\partial x_p} \right) \\ &= - \sum_{p,q} (\varepsilon_p \iota_q + \iota_q \varepsilon_p) \frac{\partial^2}{\partial x_p \partial x_q} \\ &= - \sum \pm \frac{\partial^2}{\partial x_p^2}. \end{aligned}$$

Итак, мы доказали формулу для  $\square$  в евклидовых координатах.

В случае положительно определенного скалярного произведения операторы  $\varepsilon_p$  и  $\iota_q$  имеют очень важное значение в квантовой физике, где они известны как операторы *рождения* и *уничтожения* фермионов. Равенства

$$\begin{aligned} \varepsilon_p \varepsilon_q + \varepsilon_q \varepsilon_p &= 0, \quad \iota_p \iota_q + \iota_q \iota_p = 0, \\ \varepsilon_p \iota_q + \iota_q \varepsilon_p &= \begin{cases} 0, & p \neq q, \\ 1, & p = q \end{cases} \end{aligned}$$

называются *антикоммутационными соотношениями* для этих операторов.

Алгебра, порожденная этими операторами, т.е. множество всех сумм и произведений  $\varepsilon_i$  и  $\iota_i$ , является частным случаем семейства алгебр, построенного Клиффордом в девятнадцатом веке. Эти алгебры, т.е. множества объектов, для которых определены операции сложения и умножения, были разработаны Клиффордом как обобщение систем комплексных чисел и кватернионов Гамильтона. Они играют очень важную роль в современной физике, поэтому мы и расскажем здесь о них.

Начнем с переформулировки предыдущего примера антикоммутационных соотношений между операторами рождения и уничтожения фермионов. Пусть  $V$  — векторное пространство, а  $V^*$  — его дуальное пространство. Рассмотрим пространство прямой суммы  $W = V \oplus V^*$ . Тогда вектором пространства  $W$  является

$\mathbf{w} = \begin{pmatrix} \mathbf{v} \\ \alpha \end{pmatrix}$ , где  $\mathbf{v} \in V$  и  $\alpha \in V^*$ . Определим скалярное произведение в пространстве  $W$  формулой

$$(\mathbf{w}, \mathbf{w}') = \frac{1}{2}(\alpha(\mathbf{v}') + \alpha'(\mathbf{v})),$$

где

$$\mathbf{w} = \begin{pmatrix} \mathbf{v} \\ \alpha \end{pmatrix} \quad \text{и} \quad \mathbf{w}' = \begin{pmatrix} \mathbf{v}' \\ \alpha' \end{pmatrix}.$$

Пусть  $\mathbf{v}_1, \dots, \mathbf{v}_n$  образуют базис пространства  $V$ , а  $\alpha^1, \dots, \alpha^n$  — дуальный базис пространства  $V^*$ . Тогда вектора

$$\begin{pmatrix} \mathbf{v}_1 \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} \mathbf{v}_n \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \alpha^1 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \alpha^n \end{pmatrix}$$

образуют базис пространства  $W$ . Обозначим эти вектора  $\mathbf{e}_p$  и  $\mathbf{i}_q$ , т. е. положим

$$\mathbf{e}_p = \begin{pmatrix} \mathbf{v}_p \\ 0 \end{pmatrix}, \quad \mathbf{i}_q = \begin{pmatrix} 0 \\ \alpha^q \end{pmatrix}.$$

В этом случае

$$\begin{aligned} (\mathbf{e}_p, \mathbf{e}_q) &= 0, & (\mathbf{i}_p, \mathbf{i}_q) &= 0, \\ (\mathbf{e}_p, \mathbf{i}_q) &= \begin{cases} \frac{1}{2}, & \text{если } p = q, \\ 0, & \text{если } p \neq q \end{cases} \end{aligned}$$

Теперь мы можем записать антикоммутиационные соотношения для  $\varepsilon_i$  и  $\iota_i$  в краткой форме. Будем считать, что  $\varepsilon_p$  соответствует  $\mathbf{e}_p$ , а  $\iota_q$  соответствует  $\mathbf{i}_q$ , т. е. рассмотрим линейное отображение  $\gamma$ , определенное формулами

$$\begin{aligned} \gamma(\mathbf{e}_p) &= \varepsilon_p, \\ \gamma(\mathbf{i}_q) &= \iota_q. \end{aligned}$$

Тогда для

$$\mathbf{w} = v_1 \mathbf{e}_1 + \dots + v_n \mathbf{e}_n + \alpha^1 \mathbf{i}_1 + \dots + \alpha^n \mathbf{i}_n$$

получаем, что

$$\gamma(\mathbf{w}) = v_1 \varepsilon_1 + \cdots + v_n \varepsilon_n + \alpha^1 \iota_1 + \cdots + \alpha^n \iota_n$$

является элементом нашей алгебры и

$$\gamma(\mathbf{w})\gamma(\mathbf{w}') + \gamma(\mathbf{w}')\gamma(\mathbf{w}) = 2(\mathbf{w}, \mathbf{w}')\text{id}.$$

Это можно обобщить на *любое* векторное пространство  $W$ , имеющее скалярное произведение  $(\cdot, \cdot)$  (возможно, вырожденное). Итак, каждому векторному пространству  $W$  со скалярным произведением  $(\cdot, \cdot)$  соответствует ассоциативная алгебра<sup>3</sup>  $C(W)$ , называемая *алгеброй Клиффорда* для  $W$  и  $(\cdot, \cdot)$ . Должно существовать такое линейное отображение  $\gamma : W \rightarrow C(W)$ , что

(i) каждый элемент  $C(W)$  может быть записан в виде суммы произведений элементов  $\gamma(W)$  (и кратных  $\mathbb{I}$ );

(ii)  $\gamma(\mathbf{w})\gamma(\mathbf{w}') + \gamma(\mathbf{w}')\gamma(\mathbf{w}) = 2(\mathbf{w}, \mathbf{w}')\mathbb{I}$ .

(Когда мы говорим «алгебра Клиффорда для ...», то подразумевается теорема — для данного пространства  $W$  со скалярным произведением  $(\cdot, \cdot)$  существует алгебра  $C(W)$ , удовлетворяющая условиям (i) и (ii). При этом  $C(W)$  единственна с точностью до изоморфизма. В приложении к этой главе мы дадим краткое доказательство этой теоремы.)

Давайте рассмотрим несколько примеров.

(а) Пусть одномерное пространство  $W$  имеет отрицательно определенное скалярное произведение. Это значит, что существует вектор  $\mathbf{e}$ , для которого

$$(\mathbf{e}, \mathbf{e}) = -1$$

и каждый элемент пространства  $V$  кратен  $\mathbf{e}$ . Согласно (ii),

$$\gamma(\mathbf{e})^2 = -1,$$

---

<sup>3</sup>Под ассоциативной алгеброй  $A$  мы подразумеваем, что  $A$  является векторным пространством с билинейным отображением  $A \times A \rightarrow A$ , называемым умножением. Предполагается, что это умножение удовлетворяет ассоциативному закону и что для него существует нейтральный элемент  $\mathbb{I}$ .

а условие (i) говорит нам, что каждый элемент  $C(W)$  можно записать в форме

$$a\mathbb{I} + b\gamma(\mathbf{e}).$$

Если положить  $\gamma(\mathbf{e}) = i$ , то сразу видно, что  $C(W)$  оказывается алгеброй комплексных чисел.

(b) Пусть  $V = \mathbb{R}^2$ , причем скалярное произведение противоположно обычному. Тем самым

$$(\mathbf{v}, \mathbf{v}) = -x^2 - y^2, \quad \text{если} \quad \mathbf{v} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Пусть

$$i = \gamma \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{и} \quad j = \gamma \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Условие (ii) говорит, что

$$i^2 = -\mathbb{I},$$

$$j^2 = -\mathbb{I}$$

и

$$ij + ji = 0.$$

Тогда условие (i) означает, что каждый элемент  $C(W)$  можно записать в форме

$$a\mathbb{I} + bi + cj + dk,$$

где мы положили

$$k = ij.$$

Тогда

$$k^2 = ijij = -iijj = -\mathbb{I},$$

так что

$$i^2 = j^2 = k^2 = -\mathbb{I}.$$

Кроме того,

$$ij = k,$$

$$jk = jij = -iij = i,$$

$$ki = iji = iij = j,$$

а

$$kj = ijj = -i$$

и

$$ik = iij = -j.$$

Подведем итоги:

$$\begin{aligned} i^2 = j^2 = k^2 &= -\mathbb{I}, \\ ij = k, \quad jk = i, \quad ki = j, \\ ij = -ji, \quad jk = -kj, \quad ik = -ki. \end{aligned}$$

Эта алгебра называется алгеброй *кватернионов*. Впервые она была найдена Гамильтоном.

В общем случае для проверки условия (ii) достаточно показать, что

$$\gamma(\mathbf{w})^2 = (\mathbf{w}, \mathbf{w})\mathbb{I}.$$

Действительно,

$$\begin{aligned} \gamma(\mathbf{w} + \mathbf{w}')^2 &= [\gamma(\mathbf{w}) + \gamma(\mathbf{w}')]^2 \\ &= \gamma(\mathbf{w})^2 + \gamma(\mathbf{w})\gamma(\mathbf{w}') + \gamma(\mathbf{w}')\gamma(\mathbf{w}) + \gamma(\mathbf{w}')^2 \end{aligned}$$

и

$$\begin{aligned} \gamma(\mathbf{w} + \mathbf{w}')^2 &= (\mathbf{w} + \mathbf{w}', \mathbf{w} + \mathbf{w}')\mathbb{I} \\ &= (\mathbf{w}, \mathbf{w})\mathbb{I} + 2(\mathbf{w}, \mathbf{w}')\mathbb{I} + (\mathbf{w}', \mathbf{w}')\mathbb{I}, \end{aligned}$$

т. е. условие (ii) доказано.

Приняв это во внимание, возьмем  $W = \mathbb{R}^{1,3}$ , т. е. пусть  $W$  будет пространством-временем с лоренцевой метрикой. Положим

$$\mathbf{w} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix} \quad \text{и} \quad \gamma(\mathbf{w}) = \begin{pmatrix} 0 & 0 & t+z & x+iy \\ 0 & 0 & x-iy & t-z \\ t-z & -x-iy & 0 & 0 \\ -x+iy & t+z & 0 & 0 \end{pmatrix}.$$

Тогда

$$\gamma(\mathbf{w})^2 = (t^2 - x^2 - y^2 - z^2) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Итак, если единичную матрицу  $4 \times 4$  обозначить  $\mathbb{I}$ , то условие (ii) будет выполнено. По существу мы можем рассматривать  $C(W)$  как определенную подалгебру алгебры комплексных матриц  $4 \times 4$ . Эта алгебра  $C(W)$  называется *алгеброй Дирака*. Она была разработана Дираком, когда он исследовал электрон. Сейчас она играет важную роль в теории элементарных частиц.

(с) Пусть  $W$  — произвольное векторное пространство, а скалярное произведение  $(\cdot, \cdot)$  тождественно равно нулю. Тогда условие (ii) говорит, что

$$\gamma(\mathbf{w})\gamma(\mathbf{w}') = -\gamma(\mathbf{w}')\gamma(\mathbf{w})$$

для любых  $\mathbf{w}$  и  $\mathbf{w}'$ . В этом случае  $C(W)$  оказывается в точности внешней алгеброй  $\Lambda(W)$ .

## 18.6. Звездный оператор и геометрия

В главе 16 мы видели, что звездный оператор, действующий из  $\Lambda^1(\mathbb{R}^3)$  в  $\Lambda^2(\mathbb{R}^3)$ , определяет евклидово скалярное произведение в пространстве  $\mathbb{R}^3$ . Это не что иное, как математическое выражение того факта, что диэлектрические свойства вакуума определяют евклидову геометрию пространства.

Сейчас мы можем поставить следующий вопрос. Пусть  $V$  обозначает векторное пространство с заданной ориентацией, и пусть  $(\cdot, \cdot)$  и  $(\cdot, \cdot)'$  — два невырожденных скалярных произведения в пространстве  $V$ . Предположим, что при некотором  $k$ , где  $1 \leq k < n = \dim V$ , они дают *один и тот же* звездный оператор из  $\Lambda^k(V^*)$  в  $\Lambda^{n-k}(V^*)$ . Значит ли это, что  $(\cdot, \cdot) = (\cdot, \cdot)'$ ? Ответ почти утвердительный. Рассмотрим несколько примеров.

Возьмем  $V = \mathbb{R}^2$  с обычным скалярным произведением. Тогда мы можем  $\Lambda^1(\mathbb{R}^{2*})$  идентифицировать с  $\mathbb{R}^2$ . Как мы уже видели, звездный оператор  $\star : \Lambda^1(\mathbb{R}^{2*}) \rightarrow \Lambda^1(\mathbb{R}^{2*})$  в этом случае является просто поворотом на девяносто градусов. Всякое другое скалярное произведение в пространстве  $\mathbb{R}^2$  (или в  $\mathbb{R}^{2*}$ ) определяется формулой

$$(\mathbf{u}, \mathbf{v})' = (A\mathbf{u}, \mathbf{v}),$$

где  $A$  — симметричная матрица. Если  $(\cdot, \cdot)'$  определяет тот же самый оператор, то мы должны иметь равенство

$$(\star \mathbf{u}, \star \mathbf{v})' = (\mathbf{u}, \mathbf{v}).$$

Откуда

$$R^{-1}AR = A.$$

Однако, если  $RAR^{-1} = A$  для некоторого нетривиального вращения  $R$  (не на 0 или 180 градусов), то  $A$  должна быть пропорциональна единичной матрице.<sup>4</sup> Тогда

$$(\mathbf{u}, \mathbf{v})' = \lambda(\mathbf{u}, \mathbf{v}),$$

где  $\lambda$  — число, отличное от нуля.

Мы сейчас покажем, что  $(\cdot, \cdot)'$  при  $\lambda > 0$  определяет тот же самый оператор  $\star$ . Таким образом, на плоскости оператор  $\star$  не определяет геометрию Евклида, но определяет *конформную* геометрию. Этот факт является основой теории функций комплексного переменного. Введение в эту тему мы изложим в главе 20.

Давайте посмотрим, как в общем случае *масштабное преобразование* (замена  $\mathbf{v}$  на  $c\mathbf{v}$ , где число  $c$  отлично от нуля) действует на оператор  $\star$ . При умножении длины на  $c$  площадь умножается на  $c^2$ , объем умножается на  $c^3$  и так далее. Поэтому в пространстве  $\Lambda^k(V^*)$  мы умножаем на  $c^k$ . Тогда, чтобы масштабный множитель  $c$  не оказывал никакого эффекта, должно выполняться равенство

$$(c^k \lambda, c^{n-k} \omega) c^n \sigma = c^k \lambda \wedge c^{n-k} \omega.$$

Чтобы получить ту же линейную функцию  $\omega$ , мы должны иметь

$$c^{2k-n} = 1.$$

Итак, при  $k = n/2$  масштабное преобразование не влияет на  $\star : \Lambda^k(V^*) \rightarrow \Lambda^k(V^*)$ . Несложно доказать, подобно тому как мы делали это в главе 16 для  $\star : \Lambda^1(\mathbb{R}^{3*}) \rightarrow \Lambda^2(\mathbb{R}^{3*})$ , что во всех остальных случаях звездный оператор определяет скалярное произведение однозначно.

---

<sup>4</sup>Аналогичное утверждение справедливо в евклидовом пространстве любой размерности; доказательство оставляем в качестве упражнения. — Прим. ред.

Вернемся к очень важному вопросу, который мы здесь только слегка затронем. Вспомним, что обсуждая диэлектрическую постоянную, мы допускали возможность, что  $\epsilon$  может быть переменной симметричной матрицей

$$\begin{pmatrix} \epsilon_{xx} & \epsilon_{xy} & \epsilon_{xz} \\ \epsilon_{yx} & \epsilon_{yy} & \epsilon_{yz} \\ \epsilon_{zx} & \epsilon_{zy} & \epsilon_{zz} \end{pmatrix},$$

все элементы которой являются функциями в пространстве  $\mathbb{R}^3$ . Но тогда возникает мысль, что само понятие длины должно меняться от точки к точке, т.е. что скалярное произведение двух касательных векторов  $\xi = (x, \mathbf{v})$  и  $\eta = (x, \mathbf{w})$  в точке  $x$  дается формулой

$$\sum g_{ij}(x) v^i w^j,$$

где  $v^i$  и  $w^j$  — координаты векторов  $\mathbf{v}$  и  $\mathbf{w}$ . Идея, что скалярное произведение может меняться от точки к точке, впервые была высказана Риманом и с тех пор является одной из основ геометрии. В частности, именно отсюда пошла теория относительности Эйнштейна. Теория звездного оператора без проблем распространяется на риманову геометрию.

## 18.7. Звездный оператор и векторный анализ

Как обычно, существование скалярного произведения устанавливает соответствие между пространством  $V$  и его дуальным пространством  $V^* = \Lambda^1(V)$ , которое можно использовать для идентификации дифференциальной формы с векторным полем. Для удобства обозначений векторное поле, связанное с дифференциальной формой, будем помечать тем же символом со стрелкой над ним, т.е. если  $A$  — 1-форма, то  $\vec{A}$  обозначает векторное поле, связанное с ней формулой

$$\omega[\vec{A}] = (A, \omega) \quad \text{для всех 1-форм } \omega.$$

Итак, если имеется векторное поле

$$\vec{A} = A_1 \mathbf{e}_1 + A_2 \mathbf{e}_2 + \dots,$$

то связанная с ним 1-форма

$$A = \pm A_1 dx^1 \pm A_2 dx^2 + \dots,$$

где знак каждого члена выбирается в соответствии со знаком  $(\mathbf{e}_i, \mathbf{e}_i) = \pm 1$ . В частности, в пространстве  $\mathbb{R}^3$  с евклидовым скалярным произведением, если

$$\vec{A} = A_x \mathbf{e}_x + A_y \mathbf{e}_y + A_z \mathbf{e}_z,$$

то

$$A = A_x dx + A_y dy + A_z dz.$$

Для четырехмерного пространства-времени с лоренцевым скалярным произведением, если

$$\vec{A} = A_t \mathbf{e}_t + A_x \mathbf{e}_x + A_y \mathbf{e}_y + A_z \mathbf{e}_z,$$

то

$$A = A_t dt - A_x dx - A_y dy - A_z dz.$$

Теперь в пространстве  $\mathbb{R}^3$  мы можем записывать операции векторного анализа через звездный оператор и оператор  $d$ . Тогда скалярное произведение<sup>5</sup> можно записать в виде

$$\vec{A} \cdot \vec{B} = \star(A \wedge \star B) = \star(B \wedge \star A),$$

а векторное произведение в виде

$$\vec{A} \times \vec{B} = \overrightarrow{\star(A \wedge B)}.$$

Дифференциальные векторные операторы принимают вид

$$\begin{aligned} \text{grad } f &= \vec{df}, \\ \text{rot } \vec{A} &= \overrightarrow{\star dA}, \\ \text{div } \vec{A} &= \star d\star A. \end{aligned}$$

---

<sup>5</sup>Читателю следует отличать скалярное произведение из векторного анализа (dot product), имеющее физико-геометрическую природу, от соответствующего термина линейной алгебры (scalar product) с его формальным определением. Это объясняет различие в английских названиях и в обозначениях. — *Прим. ред.*

В системах координат, отличных от декартовой, связь между векторным полем  $\vec{A}$  и дифференциальной формой  $A$  устанавливается, если сначала построить ортонормированный базис 1-форм в каждой точке. Например, в цилиндрических координатах  $dr$ ,  $r d\theta$  и  $dz$  являются ортонормированными 1-формами, дуальными к (единичным) базисным векторам  $\hat{e}_r$ ,  $\hat{e}_\theta$  и  $\hat{e}_z$ . Тогда

$$\vec{A} = A_r \hat{e}_r + A_\theta \hat{e}_\theta + A_z \hat{e}_z$$

соответствует 1-форме

$$A = A_r dr + A_\theta(r d\theta) + A_z dz.$$

Аналогично в сферических координатах  $dr$ ,  $r d\theta$ ,  $r \sin \theta d\phi$  — ортонормированные 1-формы, дуальные к базисным векторам  $\hat{e}_r$ ,  $\hat{e}_\theta$ ,  $\hat{e}_\phi$ . Тогда

$$\vec{A} = A_r \hat{e}_r + A_\theta \hat{e}_\theta + A_\phi \hat{e}_\phi$$

соответствует

$$A = A_r dr + A_\theta r d\theta + A_\phi r \sin \theta d\phi.$$

Теперь очень просто вычислять  $\text{div}$ ,  $\text{grad}$ ,  $\text{rot}$  в цилиндрических или сферических координатах. Правило применения оператора  $d$  одинаково во всех системах координат. Звездный оператор действует на ортонормированный базис  $\{dr, r d\theta, dz\}$  или  $\{dr, r d\theta, r \sin \theta d\phi\}$  точно так же, как и на базис  $\{dx, dy, dz\}$ . Например, мы хотим вычислить  $\text{rot } \vec{A}$  в цилиндрических координатах, используя  $\text{rot } \vec{A} = \overrightarrow{\star dA}$ :

$$\vec{A} = A_r \hat{e}_r + A_\theta \hat{e}_\theta + A_z \hat{e}_z,$$

$$A = A_r dr + A_\theta r d\theta + A_z dz,$$

$$\begin{aligned} dA = & \left( \frac{\partial A_z}{\partial \theta} - \frac{\partial(r A_\theta)}{\partial z} \right) \frac{1}{r} (r d\theta \wedge dz) + \left( \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right) (dz \wedge dr) \\ & + \left( \frac{\partial(r A_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) (dr \wedge r d\theta), \end{aligned}$$

$$\begin{aligned} \star dA = & \left( \frac{\partial A_z}{\partial \theta} - \frac{\partial(r A_\theta)}{\partial z} \right) \frac{1}{r} dr + \left( \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right) r d\theta \\ & + \left( \frac{\partial(r A_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) \frac{1}{r} dz. \end{aligned}$$

Теперь заменяем  $dr$  на  $\hat{e}_r$ ,  $r d\theta$  на  $\hat{e}_\theta$  и  $dz$  на  $\hat{e}_z$  и получаем

$$\begin{aligned} \operatorname{rot} \vec{A} = & \left( \frac{\partial A_z}{\partial \theta} - \frac{\partial(r A_\theta)}{\partial z} \right) \frac{1}{r} \hat{e}_r + \left( \frac{\partial A_r}{\partial z} - \frac{\partial A_z}{\partial r} \right) \hat{e}_\theta \\ & + \left( \frac{\partial(r A_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) \frac{1}{r} \hat{e}_z. \end{aligned}$$

Аналогично вычисляется дивергенция радиального вектора в сферических координатах, используя  $\operatorname{div} \vec{A} = \star d \star A$ :

$$\begin{aligned} \vec{A} &= A_r \hat{e}_r, \\ A &= A_r dr, \\ \star A &= A_r (r d\theta \wedge r \sin \theta d\phi) = r^2 A_r \sin \theta d\theta \wedge d\phi, \\ d \star A &= \frac{\partial(r^2 A_r)}{\partial r} \sin \theta dr \wedge d\theta \wedge d\phi \\ &= \frac{1}{r^2} \frac{\partial(r^2 A_r)}{\partial r} dr \wedge r d\theta \wedge r \sin \theta d\phi, \\ \star d \star A &= \frac{1}{r^2} \frac{\partial(r^2 A_r)}{\partial r} = \operatorname{div} \vec{A}. \end{aligned}$$

Итак, секрет состоит в том, что сначала надо выразить все дифференциальные формы через ортонормированные 1-формы и только потом действовать звездным оператором.

## Приложение: тензорные произведения

Для того, что мы делаем в этой книге, очень важно понятие тензорного произведения двух или более векторных пространств. Это одно из основополагающих понятий в теории векторных пространств. Мы должны здесь дать его краткое изложение, потому что оно отсутствует во многих учебниках по элементарной линейной алгебре.<sup>6</sup>

---

<sup>6</sup>Этим недостатком не обладают большинство учебников линейной алгебры на русском языке. С другой стороны, необходимо отметить, что в этой книге авторы так и не приводят явного определения *тензора* — столь же фундаментального понятия линейной алгебры (и не только), включающего в себя большинство изучаемых ею объектов и давшего название изучаемой в этом приложении операции над векторными пространствами. — Прим. ред.

Пусть у нас есть векторные пространства  $V, W$  и  $U$ . Отображение  $f : V \times W \rightarrow U$  называется *билинейным*, если

$$\begin{aligned} f(\mathbf{v}_1 + \mathbf{v}_2, \mathbf{w}) &= f(\mathbf{v}_1, \mathbf{w}) + f(\mathbf{v}_2, \mathbf{w}), \\ f(\mathbf{v}, \mathbf{w}_1 + \mathbf{w}_2) &= f(\mathbf{v}, \mathbf{w}_1) + f(\mathbf{v}, \mathbf{w}_2) \end{aligned}$$

и

$$f(a\mathbf{v}, \mathbf{w}) = af(\mathbf{v}, \mathbf{w}) = f(\mathbf{v}, a\mathbf{w}) \quad \text{для скаляра } a.$$

Другими словами, отображение  $f$  называется билинейным, если оно линейно по  $\mathbf{v}$  при фиксированном  $\mathbf{w}$  и линейно по  $\mathbf{w}$  при фиксированном  $\mathbf{v}$ .

Известным примером, когда  $V = W = U$ , является *векторное произведение* в обычном трехмерном пространстве.

Взяв пространства  $V$  и  $W$ , мы хотим построить векторное пространство  $Z$  и билинейное отображение  $b : V \times W \rightarrow Z$ , которое будет универсальным в следующем смысле. Предположим, что  $f : V \times W \rightarrow U$  — произвольное билинейное отображение. Тогда существует *единственное* линейное отображение  $l_f : Z \rightarrow U$  такое, что

$$f = l_f \circ b.$$

На диаграмме это можно изобразить так:

$$\begin{array}{ccc} V \times W & \xrightarrow{b} & Z \\ & \searrow f & \swarrow l_f \\ & U & \end{array}$$

Таким образом, для любого  $f$  существует единственное линейное отображение  $l_f$ , для которого  $l_f \circ b = f$ , т. е. диаграмма коммутативна.

Заметим, что если  $b$  и  $Z$  существуют, то они единственны с точностью до изоморфизма. Действительно, предположим, что  $b'$  и  $Z'$  — другое *универсальное* билинейное отображение. Тогда, положив  $U = Z'$  и  $f = b'$  в нашей диаграмме, мы получим

$$\begin{array}{ccc} V \times W & \xrightarrow{b} & Z \\ & \searrow b' & \swarrow l_{b'} \\ & Z' & \end{array}$$

Таким образом, имеем отображение  $l_{b'} : Z \rightarrow Z'$  с

$$b' = l_{b'} \circ b.$$

Аналогично получаем линейное отображение  $l_b : Z' \rightarrow Z$  с

$$b = l_b \circ b'.$$

Но тогда

$$b = (l_b \circ l_{b'}) \circ b.$$

В то же время

$$\begin{array}{ccc} V \times W & \xrightarrow{b} & Z \\ & \searrow b & \swarrow \text{id} \\ & & Z \end{array} \quad \begin{array}{c} \\ \\ \swarrow l_b \circ l_{b'} \end{array}$$

Значит,  $l = \text{id}$  и  $l = l_b \circ l_{b'}$  удовлетворяют уравнению

$$b = lb.$$

В силу единственности мы приходим к выводу, что

$$l_b \circ l_{b'} = \text{id}$$

и аналогично

$$l_{b'} \circ l_b = \text{id}.$$

Таким образом,  $l_{b'}$  дает изоморфизм  $Z$  с  $Z'$  и  $b' = l_{b'} \circ b$ . Поэтому с точностью до изоморфизма  $b$  и  $Z$  определены однозначно.

Наша задача состоит в том, чтобы показать существование такой пары  $(b, Z)$ . Мы здесь приведем два построения, которые на первый взгляд выглядят различными. Однако в силу единственности они должны давать одинаковый результат.

Пусть  $Y$  обозначает пространство всех скалярных билинейных функций, определенных в пространстве  $V \times W$ . Тогда  $\alpha \in Y$  является билинейным отображением  $V \times W \rightarrow \mathbb{R}$ . Очевидно, что множество всех билинейных функций  $\alpha$  образуют векторное пространство  $Y$ . Возьмем

$$Z = Y^*.$$

Тогда вектор в пространстве  $Z$  является линейной функцией на пространстве  $Y$ . Отображение  $b : V \times W \rightarrow Z$  определим формулой

$$[b(\mathbf{v}, \mathbf{w})](\alpha) = \alpha(\mathbf{v}, \mathbf{w}).$$

Это значит, что  $b(\mathbf{v}, \mathbf{w})$  является такой линейной функцией, которая каждой билинейной функции  $\alpha$  приписывает значение  $\alpha(\mathbf{v}, \mathbf{w})$ . Очевидно, что  $b(\mathbf{v}, \mathbf{w})$  зависит линейно от  $\alpha$  и билинейно от  $\mathbf{v}$  и  $\mathbf{w}$ . Утверждается, что  $b$  и  $Z$  дают решение нашей общей задачи.

Действительно, пусть  $f : V \times W \rightarrow U$  — билинейное отображение. Здесь  $U$  обозначает векторное пространство. Для каждого  $\nu \in U^*$  мы получаем билинейную функцию

$$\nu \circ f \in Y.$$

Итак, мы определили линейное отображение  $l^* : U^* \rightarrow Y$ :

$$l^*(\nu) = \nu \circ f.$$

Отсюда получаем его сопряженное  $l^{**} : Z = Y^* \rightarrow U^{**}$ . По определению

$$\begin{aligned} l^{**}(b(\mathbf{v}, \mathbf{w}))(\nu) &= b(\mathbf{v}, \mathbf{w})(l^*(\nu)) \\ &= (l^*(\nu))(\mathbf{v}, \mathbf{w}) \\ &= (\nu \circ f)(\mathbf{v}, \mathbf{w}) \\ &= \nu(f(\mathbf{v}, \mathbf{w})). \end{aligned}$$

Другими словами,  $l^{**}(b(\mathbf{v}, \mathbf{w}))$  оказывается той линейной функцией в пространстве  $U^*$ , которая каждому  $\nu$  приписывает число  $\nu(f(\mathbf{v}, \mathbf{w}))$ . А сейчас вспомним, что существует канонический изоморфизм<sup>7</sup>

$$U \xrightarrow{i} U^{**},$$

где каждый элемент  $\mathbf{u} \in U$  идентифицируется с линейной функцией  $i(\mathbf{u})$  на пространстве  $U^*$ , согласно формуле

$$i(\mathbf{u})(\nu) = \nu(\mathbf{u}), \quad \mathbf{u} \in U, \quad \nu \in U^*.$$

<sup>7</sup>Напомним, что он называется отображением вычисления. В дальнейшем для значения элемента  $\alpha \in U^*$  на векторе  $\mathbf{u} \in U$  также используется «симметричное» обозначение  $\langle \alpha, \mathbf{u} \rangle = \langle \mathbf{u}, \alpha \rangle = i(\mathbf{u})(\alpha) = \alpha(\mathbf{u})$ . — Прим. ред.

Тогда мы можем написать

$$l^{**}(b(\mathbf{v}, \mathbf{w})) = i(f(\mathbf{v}, \mathbf{w})).$$

Используя эту идентификацию  $i$ , можно написать  $l^{**} = i \circ l$ , где  $l$  является отображением  $Z \rightarrow U$ . Итак, мы нашли отображение  $l: Z \rightarrow U$ , удовлетворяющее условию

$$l \circ b = f.$$

Предположим, что существует два таких отображения. Тогда для их разности  $m$  следует, что

$$(m \circ b)(\mathbf{v}, \mathbf{w}) = 0 \quad \text{для всех } \mathbf{v}, \mathbf{w} \in V.$$

Значит,

$$m^*(\nu)(b(\mathbf{v}, \mathbf{w})) = 0 \quad \text{для всех } \nu \in U^*.$$

Итак,  $m^*(\nu) \in Y$  является такой билинейной функцией на пространстве  $V \times W$ , которая приписывает 0 всем парам  $(\mathbf{v}, \mathbf{w})$ , т. е.  $m^*(\nu)$  является нулевой билинейной функцией. Поскольку  $m^*$  переводит все элементы пространства  $U^*$  в нуль, это нулевое линейное отображение. Отсюда следует, что  $m = 0$ . Таким образом, мы доказали единственность  $l$ .

Пространство  $Z$  называется *тензорным произведением* пространств  $V$  и  $W$  и обозначается  $V \otimes W$ . Кроме того, мы будем пользоваться обозначением

$$b(\mathbf{v}, \mathbf{w}) = \mathbf{v} \otimes \mathbf{w}.$$

Предположим, что пространства  $V$  и  $W$  конечномерны. Пусть  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$  — базис пространства  $V$ , а  $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$  — базис пространства  $W$ . Тогда билинейная функция  $\beta$  на пространстве  $V \otimes W$  полностью определена своими значениями  $\beta(\mathbf{e}_i, \mathbf{f}_j)$ . Мы можем записать

$$\beta = \sum \beta(\mathbf{e}_i, \mathbf{f}_j) \varepsilon_{ij},$$

где  $\varepsilon_{ij}$  — билинейная функция, определенная формулами:

$$\begin{aligned} \varepsilon_{ij}(\mathbf{v}, \mathbf{w}) &= v_i w_j, \quad \text{если } \mathbf{v} = v_1 \mathbf{e}_1 + \dots + v_m \mathbf{e}_m, \\ &\quad \mathbf{w} = w_1 \mathbf{f}_1 + \dots + w_n \mathbf{f}_n. \end{aligned}$$

Очевидно, что элементы  $\varepsilon_{ij}$  линейно независимы и порождают пространство  $Y$ , являясь его базисом. Дуальным базисом пространства  $Z = Y^*$  является  $b(\mathbf{e}_i, \mathbf{f}_j) = \mathbf{e}_i \otimes \mathbf{f}_j$ . Тогда

если  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$  является базисом пространства  $V$ , а  $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$  является базисом пространства  $W$ , то  $\{\mathbf{e}_i \otimes \mathbf{f}_j\}_{i=1, \dots, m; j=1, \dots, n}$  оказывается базисом пространства  $V \otimes W$ . Кроме того,

$$\dim(V \otimes W) = (\dim V) \times (\dim W).$$

Опишем другой способ построения пространства  $V \otimes W$ . Он более прямой и абстрактный, но имеет один недостаток — используются бесконечномерные пространства. Для произвольного множества  $M$  обозначим  $F(M)$  векторное пространство всех линейных комбинаций элементов этого множества<sup>8</sup>. Тогда  $F(M)$  будет состоять из конечных сумм вида

$$a_{m_1} m_1 + \dots + a_{m_k} m_k,$$

где  $m_i$  — элементы  $M$ , а  $a_{m_i}$  — скалярные коэффициенты. Такие выражения складываются путем приведения подобных членов. Если  $M$  — конечное множество, то  $F(M)$  будет пространством всех функций на этом множестве, причем функция  $f$  соответствует формальной сумме

$$\sum_{m \in M} f(m) m.$$

Если же  $M$  бесконечно, то можно считать, что  $F(M)$  является пространством всех функций на множестве  $M$ , которые равны нулю везде, кроме конечного числа точек. (Здесь мы отождествляем  $m \in M$  с функцией  $\delta_m$ , обладающей свойствами:  $\delta_m(n) = 0$ , если  $n \neq m$ , и  $\delta_m(m) = 1$ . Тогда общий вид элемента  $F(M)$  — это конечная линейная комбинация  $\delta_m$ . Отображение  $M \rightarrow F(M)$ , переводящее  $m \rightarrow \delta_m$ , дает возможность рассматривать  $M$  как подпространство  $F(M)$ .)

<sup>8</sup> Оно называется свободным векторным пространством над данным множеством. — *Прим. ред.*

Пространство  $F(M)$  универсально по отношению к отображениям  $M$  на векторные пространства: для данного отображения  $\phi : M \rightarrow U$ , где  $U$  — векторное пространство, существует единственное линейное отображение  $L_\phi : F(M) \rightarrow U$  такое, что

$$L_\phi \circ \delta_m = \phi(m).$$

Действительно, эта формула связывает  $L_\phi$  с элементами  $\delta_m$  и, следовательно, в силу линейности  $L_\phi$  распространяется на все элементы  $F(M)$ .

Теперь вернемся к нашей задаче. Возьмем  $M = V \times W$ . Если  $\phi$  — отображение множества  $V \times W \rightarrow U$ , то существует  $L_\phi : F(V \times W) \rightarrow U$ , которое линейно и удовлетворяет уравнению  $L_\phi \circ \delta_m = \phi(m)$  для всех  $m \in V \times W$ . Если  $f : V \times W \rightarrow U$  билинейно, то  $L_f$  должно обращаться в нуль на всех элементах

$$\begin{aligned} \delta_{(v_1+v_2, w)} - \delta_{(v_1, w)} - \delta_{(v_2, w)}, \quad \delta_{(v, w_1+w_2)} - \delta_{(v, w_1)} - \delta_{(v, w_2)}, \\ \delta_{(rv, w)} - r\delta_{(v, w)}, \quad \delta_{(v, rw)} - r\delta_{(v, w)}. \end{aligned}$$

Итак,  $L_f$  должно обращаться в нуль в подпространстве  $B$ , порожденном этими элементами. Тогда можно определить  $V \otimes W$  как факторпространство

$$V \otimes W = F(V \times W)/B$$

и отображение  $b : V \times W \rightarrow V \otimes W$  формулой

$$b(v, w) = [\delta_{(v, w)}],$$

где  $[\cdot]$  обозначает класс эквивалентности mod  $B$ . Теперь должно быть ясно, что это определение  $b$  и  $V \otimes W$  удовлетворяет сформулированным свойствам универсальности.

Давайте получим ряд следствий из этой универсальности. Пусть  $V'$  и  $W'$  обозначают два других векторных пространства. Пусть  $A : V \rightarrow V'$  и  $B : W \rightarrow W'$  — линейные преобразования. Рассмотрим отображение, переводящее  $(v, w)$  в  $Av \otimes Bw$ . Очевидно, что оно билинейно. Следовательно, существует единственное линейное отображение  $V \otimes W \rightarrow V' \otimes W'$ , обозначаемое  $A \otimes B$ , для которого справедлива формула

$$(A \otimes B)(v \otimes w) = Av \otimes Bw.$$

Предположим, что  $A' : V' \rightarrow V''$  и  $B' : W' \rightarrow W''$  — еще одна пара векторных пространств и линейных преобразований. Тогда отображение, переводящее  $(\mathbf{v}, \mathbf{w})$  в  $A'A\mathbf{v} \otimes B'B\mathbf{w}$ , обозначается

$$A'A \otimes B'B : V \otimes W \rightarrow V'' \otimes W''.$$

Кроме того, у нас есть отображения

$$A \otimes B : V \otimes W \rightarrow V' \otimes W'$$

и

$$A' \otimes B' : V' \otimes W' \rightarrow V'' \otimes W''.$$

Поэтому мы получаем

$$(A' \otimes B') \circ (A \otimes B) : V \otimes W \rightarrow V'' \otimes W''.$$

По построению  $(A' \otimes B')(A\mathbf{v} \otimes B\mathbf{w}) = A'A'\mathbf{v} \otimes B'B'\mathbf{w}$ . Следовательно, в силу единственности мы имеем

$$(A' \otimes B') \circ (A \otimes B) = A'A \otimes B'B.$$

Приведем еще одно следствие свойств тензорного произведения. Предположим, что в пространствах  $V$  и  $W$  имеются скалярные произведения  $(\cdot, \cdot)_V$  и  $(\cdot, \cdot)_W$ . Рассмотрим скалярную функцию  $f$ , определенную на пространстве  $(V \times W) \times (V \times W)$  формулой

$$f((\mathbf{v}, \mathbf{w}), (\mathbf{v}', \mathbf{w}')) = (\mathbf{v}, \mathbf{v}')_V (\mathbf{w}, \mathbf{w}')_W.$$

Для заданных  $\mathbf{v}'$  и  $\mathbf{w}'$  функция  $f$  билинейна по  $\mathbf{v}$  и  $\mathbf{w}$ , т.е. определяет линейную функцию  $l_{(\mathbf{v}', \mathbf{w}')}$  на пространстве  $V \otimes W$ , удовлетворяющую условию

$$l_{(\mathbf{v}', \mathbf{w}')}(\mathbf{v} \otimes \mathbf{w}) = (\mathbf{v}, \mathbf{v}')_V (\mathbf{w}, \mathbf{w}')_W.$$

Для заданных  $\mathbf{v}$  и  $\mathbf{w}$  это выражение билинейно по  $\mathbf{v}'$  и  $\mathbf{w}'$ . Поскольку каждый элемент  $\alpha \in V \otimes W$  является конечной суммой элементов вида  $\mathbf{v} \otimes \mathbf{w}$ , заключаем, что

$$l_{(\mathbf{v}', \mathbf{w}')}(\alpha)$$

билинейна по  $\mathbf{v}'$  и  $\mathbf{w}'$  для любого заданного  $\alpha$  в пространстве  $V \otimes W$ . Итак, существует линейная функция  $l^\alpha$  в пространстве  $V \otimes W$ , для которой

$$l^\alpha(\mathbf{v}' \otimes \mathbf{w}') = l_{(\mathbf{v}', \mathbf{w}')}(\alpha).$$

Тогда  $l^\alpha(\beta)$  линейна по  $\alpha$  и по  $\beta$ . Легко проверить, что

$$(\alpha, \beta)_{V \otimes W} = l^\alpha(\beta)$$

определяет скалярное произведение в пространстве  $V \otimes W$ . Подведем итоги.

Существует и однозначно определено скалярное произведение  $(\cdot, \cdot)_{V \otimes W}$  в пространстве  $V \otimes W$ , которое обладает свойством

$$(\mathbf{v} \otimes \mathbf{w}, \mathbf{v}' \otimes \mathbf{w}')_{V \otimes W} = (\mathbf{v}, \mathbf{v}')_V (\mathbf{w}, \mathbf{w}')_W.$$

**Тензорные произведения и пространства линейных отображений.** Здесь мы приведем еще одно полезное определение, связанное с тензорными произведениями. Пусть есть два векторных пространства  $V$  и  $W$ . Определим векторное пространство  $\text{Hom}(W, V)$  как пространство всех линейных преобразований из  $W$  в  $V$ . Это векторное пространство имеет размерность  $(\dim W)(\dim V)$ . Для каждого  $\mathbf{v} \in V$  и  $\mu \in W^*$  рассмотрим линейное преобразование  $T_\mathbf{v}^\mu : W \rightarrow V$  ранга 1, определенное формулой

$$T_\mathbf{v}^\mu(\mathbf{w}) = \langle \mu, \mathbf{w} \rangle \mathbf{v}.$$

Здесь  $\langle \mu, \mathbf{w} \rangle$  — это значение линейной функции  $\mu \in W^*$  на векторе  $\mathbf{w} \in W$ . Очевидно, что отображение  $T_\mathbf{v}^\mu$  линейно зависит от  $\mathbf{v}$  для фиксированного  $\mu$  и линейно зависит от  $\mu$  для фиксированного  $\mathbf{v}$ . Таким образом, мы имеем единственное линейное отображение

$$i : V \otimes W^* \rightarrow \text{Hom}(W, V),$$

определенное формулой

$$i(\mathbf{v} \otimes \mu)\mathbf{w} = \langle \mu, \mathbf{w} \rangle \mathbf{v}.$$

Легко проверить, что отображение  $i$  инъективное. Оно отображает  $V \otimes W^*$  на подпространство  $\text{Hom}(W, V)$ , состоящее из линейных преобразований конечного ранга. Если хотя бы одно из пространств  $W$  или  $V$  конечномерно, то это подпространство совпадает с  $\text{Hom}(W, V)$ . Если же оба пространства  $V$  и  $W$  бесконечномерны, то оно — собственное подпространство.

**Тензорные произведения более высокого порядка.** Существуют несколько обобщений и модификаций понятия тензорного произведения, на которых мы здесь кратко остановимся. Предположим, что вместо двух векторных пространств  $V$  и  $W$  мы имеем  $k$  векторных пространств  $V_1, \dots, V_k$ . Тогда можно определить  $k$ -линейное (или полилинейное) отображение  $f : V_1 \times \dots \times V_k \rightarrow U$ , которое должно быть линейным по каждой из переменных, когда все остальные фиксированы. В этом случае существует универсальное пространство  $V_1 \otimes \dots \otimes V_k$  и полилинейное отображение  $m : V_1 \times \dots \times V_k \rightarrow V_1 \otimes \dots \otimes V_k$ , которое обладает теми же свойствами, что и в случае двух сомножителей. Из свойства универсальности следует, что существует изоморфизм

$$(V_1 \otimes \dots \otimes V_k) \otimes (V_{k+1} \otimes \dots \otimes V_{k+l})$$

с

$$V_1 \otimes \dots \otimes V_{k+l},$$

поскольку они оба удовлетворяют свойству универсальности для  $(k+l)$ -линейных отображений.

Как и в случае  $k = 2$ , если нам даны линейные отображения  $A_1 : V_1 \rightarrow W_1$ ,  $A_2 : V_2 \rightarrow W_2$  и так далее, то мы получаем линейное отображение

$$A_1 \otimes \dots \otimes A_k : V_1 \otimes \dots \otimes V_k \rightarrow W_1 \otimes \dots \otimes W_k.$$

Также получается обобщение правила вычисления композиции. Кроме того, при идентификации  $(V_1 \otimes \dots \otimes V_k) \otimes (V_{k+1} \otimes \dots \otimes V_{k+l})$  с  $V_1 \otimes \dots \otimes V_{k+l}$  мы получаем идентификацию  $(A_1 \otimes \dots \otimes A_k) \otimes (A_{k+1} \otimes \dots \otimes A_{k+l})$  с  $A_1 \otimes \dots \otimes A_{k+l}$ .

Предположим, что  $V_1 = V_2 = \dots = V_k$  — одно и то же векторное пространство. Мы говорим, что полилинейное отображение  $f$

пространства  $V \times \cdots \times V$  ( $k$  раз) в (другое) векторное пространство  $W$  кососимметрично, если

$$f(\mathbf{v}_1, \dots, \mathbf{v}_i, \mathbf{v}_{i+1}, \dots, \mathbf{v}_k) = -f(\mathbf{v}_1, \dots, \mathbf{v}_{i+1}, \mathbf{v}_i, \dots, \mathbf{v}_k)$$

для всех  $i = 1, 2, \dots, k$ . Другими словами, перестановка двух любых соседних индексов изменяет знак  $f$ . Отсюда следует, что для любой перестановки  $\pi$  индексов  $(1, \dots, k)$  можно написать

$$f(\mathbf{v}_{\pi(1)}, \dots, \mathbf{v}_{\pi(k)}) = (\text{sign } \pi) f(\mathbf{v}_1, \dots, \mathbf{v}_k).$$

Теперь можно искать универсальное пространство и кососимметричное полилинейное отображение, т. е. мы ищем векторное пространство  $\Lambda^k(V)$  и такое кососимметричное полилинейное отображение  $m_k : V \times \cdots \times V$  ( $k$  раз)  $\rightarrow \Lambda^k(V)$ , чтобы для любого кососимметричного полилинейного отображения  $f : V \times \cdots \times V \rightarrow U$  существовало единственное линейное отображение  $l_f : \Lambda^k(V) \rightarrow U$ , для которого  $l_f \circ m_k = f$ . Изобразим это на диаграмме:

$$\begin{array}{ccc} \overbrace{V \times \cdots \times V}^{k \text{ раз}} & \xrightarrow{m_k} & \Lambda^k(V) \\ & \searrow f & \swarrow l_f \\ & U & \end{array}$$

Таким образом, для данного отображения  $f$  должно существовать единственное отображение  $l_f$ .

Отсюда немедленно следует, что если  $m_k$  и  $\Lambda^k(V)$  существуют, то они единственны с точностью до изоморфизма. Кроме того, любое из приведенных нами доказательств существования тензорных произведений приводит, с небольшими изменениями, к доказательству существования  $\Lambda^k(V)$  и  $m_k$ . В частности, первое доказательство показывает следующее.

Пусть  $V^*$  — дуальное пространство к  $V$ . Тогда  $\Lambda^k(V^*)$  можно идентифицировать с пространством всех  $k$ -линейных кососимметричных функций на пространстве  $V \times \cdots \times V$ .

**Внешняя алгебра.** Рассмотрим отображение  $m_{k+l}$  пространства  $V \times \cdots \times V$  ( $k + l$  раз) в пространство  $\Lambda^{k+l}(V)$ . Оно полилинейно и кососимметрично по первым  $k$  переменным, с одной стороны, и по остальным  $l$  переменным — с другой. Следовательно, оно определяет отображение

$$\Lambda^k(V) \otimes \Lambda^l(V) \rightarrow \Lambda^{k+l}(V),$$

известное как внешнее умножение и обозначаемое символом  $\wedge$ . С помощью свойства универсальности легко проверить, что  $\wedge$  ассоциативно, т. е.

$$(\omega \wedge \sigma) \wedge \tau = \omega \wedge (\sigma \wedge \tau),$$

где обе части отображают пространство  $\Lambda^k(V) \otimes \Lambda^l(V) \otimes \Lambda^p(V)$  на пространство  $\Lambda^{k+l+p}(V)$ ,  $\omega \in \Lambda^k(V)$ ,  $\sigma \in \Lambda^l(V)$  и  $\tau \in \Lambda^p(V)$ . Кроме того, вычисляя значения на базисных элементах, можно проверить, что

$$\omega \wedge \sigma = (-1)^{kl} \sigma \wedge \omega, \quad \omega \in \Lambda^k(V), \quad \sigma \in \Lambda^l(V).$$

Наконец, предположим, что  $V$  имеет скалярное произведение. Тогда функция в пространстве  $(V \times \cdots \times V) \times (V \times \cdots \times V)$  (каждый множитель входит  $k$  раз), заданная формулой

$$f(\mathbf{v}_1, \dots, \mathbf{v}_k; \mathbf{w}_1, \dots, \mathbf{w}_k) = \text{Det}((\mathbf{v}_i, \mathbf{w}_j)_V),$$

полилинейна и кососимметрична по  $\mathbf{v}_1, \dots, \mathbf{v}_k$  и полилинейна и кососимметрична по  $\mathbf{w}_1, \dots, \mathbf{w}_k$ . Тогда, как мы уже обсуждали для индуцированного скалярного произведения в пространстве  $V \otimes W$ , из универсальности следует, что  $f$  индуцирует скалярное произведение в пространстве  $\Lambda^k(V)$ . Другими словами, мы получаем единственное скалярное произведение  $(\cdot, \cdot)_{\Lambda^k(V)}$ , для которого справедливо равенство

$$(\mathbf{v}_1 \wedge \cdots \wedge \mathbf{v}_k, \mathbf{w}_1 \wedge \cdots \wedge \mathbf{w}_k)_{\Lambda^k(V)} = \text{Det}((\mathbf{v}_i, \mathbf{w}_j)_V).$$

Теперь идею «универсального объекта» можно использовать для построения алгебр, которые мы изучали в этой главе, т. е. для внешней алгебры и для алгебры Клиффорда.

Приведем некоторые детали. Векторное пространство  $A$  называется *алгеброй*, если задано отображение  $m$

$$m : A \otimes A \rightarrow A,$$

называемое умножением. Образ  $m(a \otimes b)$  будем обозначать  $a \cdot b$  или просто  $ab$ . Алгебра называется *ассоциативной*, если

$$a(bc) = (ab)c$$

для всех  $a, b$  и  $c$ . Говорят, что она является алгеброй с единицей, если существует элемент, обозначаемый  $1_A$  (или просто  $1$ , если нет риска путаницы), для которого выполняется равенство

$$1_A a = a$$

для всех элементов  $A$ . Если это специально не оговаривается, то все алгебры ассоциативны и имеют единицу.

**Тензорная алгебра.** Пусть  $V$  обозначает векторное пространство. Рассмотрим следующую универсальную задачу: надо найти такую алгебру  $U$  и линейное отображение  $i : V \rightarrow U$ , чтобы для произвольной алгебры  $A$  и любого линейного отображения  $f : V \rightarrow A$ , существовал единственный гомоморфизм  $\phi : U \rightarrow A$ , для которого справедливо равенство

$$\phi(1_U) = 1_A$$

и

$$\phi(iv) = f(v) \quad \text{для всех } v \in V.$$

$$\begin{array}{ccc} V & \xrightarrow{i} & U \\ & \searrow f & \downarrow \phi \\ & & A \end{array}$$

(Здесь слово гомоморфизм означает, что  $\phi$  — линейное отображение, удовлетворяющее условию  $\phi(ab) = \phi(a)\phi(b)$ .) Если  $U$  и  $i$  существуют, то пара  $(U, i)$  единственна с точностью до взаимно однозначного гомоморфизма в силу тех же аргументов, которые

применялись для всех универсальных построений. (Они действительно являются аргументами, принадлежащими теории категорий.) Как всегда, наша задача — построить одну такую алгебру. Рассмотрим (бесконечную) прямую сумму

$$T(V) = \mathbb{R} \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \dots$$

Превратим ее в алгебру, определив отображение

$$m = m_{pq} : \overbrace{(V \otimes \dots \otimes V)}^{p \text{ раз}} \otimes \overbrace{(V \otimes \dots \otimes V)}^{q \text{ раз}} \rightarrow \overbrace{V \otimes \dots \otimes V}^{p+q \text{ раз}},$$

которое должно быть единственным и удовлетворять формуле

$$m[(\mathbf{v}_1 \otimes \dots \otimes \mathbf{v}_p) \otimes (\mathbf{w}_1 \otimes \dots \otimes \mathbf{w}_q)] = \mathbf{v}_1 \otimes \dots \otimes \mathbf{v}_p \otimes \mathbf{w}_1 \otimes \dots \otimes \mathbf{w}_q$$

(где просто убраны скобки). Полилинейность гарантирует, что это отображение корректно определено. Поэтому мы знаем, как умножать  $a_p$  на  $b_q$ , где

$$a_p \in T_p(V) \stackrel{\text{def}}{=} \overbrace{V \otimes \dots \otimes V}^{p \text{ раз}}$$

и

$$b_q \in T_q(V).$$

(Для  $r \in \mathbb{R}$  мы полагаем  $r \cdot a = ra$ .) Поскольку каждый элемент

$$T(V) = \bigoplus_{r=0}^{\infty} T_r(V)$$

является конечной суммой  $a = a_0 + a_1 + \dots + a_r$ , где элемент  $a_p \in T_p(V)$ , то дистрибутивный закон (билинейность) определяет умножение в  $T(V)$ . Возьмем  $U = T(V)$  и пусть  $i : V \rightarrow T(V)$  — отображение, которое идентифицирует  $V$  как кусок прямой суммы  $T(V)$  для  $r = 1$ , обозначаемый  $T_1(V)$ .

Возьмем линейное отображение  $f : V \rightarrow A$ . Тогда

$$f(\mathbf{v}_1)f(\mathbf{v}_2)$$

(умножение в  $A$ ) билинейно по  $\mathbf{v}_1$  и  $\mathbf{v}_2$ . Следовательно, оно определяет отображение

$$\phi_2 : V \otimes V \rightarrow A.$$

Аналогично

$$\phi_r : T_r(V) \rightarrow A$$

однозначно определяется формулой

$$\phi_r(\mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_r) = f(\mathbf{v}_1) \cdots f(\mathbf{v}_r).$$

Наконец, положим

$$\phi(a_0 + \cdots + a_r) = \phi_0(a_0) + \cdots + \phi_r(a_r),$$

где  $\phi_0(a_0) = a_0 \cdot 1_A$ . Легко проверить, что  $\phi$  является гомоморфизмом, причем единственным, и удовлетворяющим условию

$$\phi(i\mathbf{v}) = f(\mathbf{v}) \quad \text{для всех } \mathbf{v} \in V.$$

Таким образом,  $(T(V), i)$  является «универсальной» алгеброй над пространством  $V$ .

Пусть  $A$  является алгеброй. Тогда подпространство  $I \subset A$  называется (двусторонним) *идеалом*, если

$$a \in A, b \in I \quad \text{означает, что} \quad ab \in I \text{ и } ba \in I.$$

Теперь мы можем определить умножение в факторпространстве

$$B = A/I$$

формулой

$$[a/I] \cdot [a'/I] = [aa'/I],$$

где мы класс эквивалентности  $a \bmod I$  обозначаем  $[a/I]$ . Оказывается, что это умножение корректно определено, т. е. оно не зависит от частного выбора  $a$  или  $a' \bmod I$ .

**Алгебра Клиффорда.** Теперь мы можем воспользоваться этим определением факторалгебры, чтобы решить другие вопросы об

универсальных объектах. Например, давайте рассмотрим алгебры Клиффорда. Пусть  $V$  обозначает векторное пространство, где имеется квадратичная форма  $Q$  и связанное с ней скалярное произведение  $(\cdot, \cdot)$ . Пусть  $A$  обозначает соответствующую алгебру с единицей  $1_A$ . Отображение Клиффорда — это такое линейное отображение

$$f : V \rightarrow A,$$

для которого

$$f(\mathbf{v})^2 = Q(\mathbf{v})1_A \quad \text{для всех } \mathbf{v} \in V.$$

Это уравнение можно переписать в форме

$$f(\mathbf{u})f(\mathbf{v}) + f(\mathbf{v})f(\mathbf{u}) = 2(\mathbf{u}, \mathbf{v})1_A.$$

Алгеброй Клиффорда над пространством  $V$  называется алгебра  $C(V, Q)$  вместе с отображением Клиффорда  $j : V \rightarrow C(V, Q)$ , которая является универсальной в обычном смысле: для любого заданного отображения Клиффорда  $f : V \rightarrow A$  существует единственный гомоморфизм  $\phi : C(V, Q) \rightarrow A$ , удовлетворяющий условию  $f = \phi \circ j$ .

$$\begin{array}{ccc} V & \xrightarrow{j} & C(V, Q) \\ & \searrow f & \downarrow \phi \\ & & A \end{array}$$

Как обычно, если алгебра  $C(V, Q)$  существует, то она должна быть единственной с точностью до изоморфизма. Наша задача построить такую алгебру. Построение диктуется свойствами тензорной алгебры. Если существует такое отображение  $j$  пространства  $V$  в алгебру  $C(V, Q)$ , то мы должны иметь единственный гомоморфизм  $\phi : T(V) \rightarrow C(V, Q)$ , для которого

$$j = \phi \circ i.$$

$$\begin{array}{ccc} V & \xrightarrow{i} & T(V) \\ & \searrow j & \downarrow \phi \\ & & C(V, Q) \end{array}$$

По-прежнему предполагая, что  $C(V, Q)$  и  $j$  существуют, обозначим через  $J \subset T(V)$  множество всех  $b$  в  $T(V)$ , для которых

$$\psi(b) = 0.$$

Тогда  $J$  является идеалом, и мы должны получить изоморфизм алгебр

$$\bar{\psi} : T(V)/J \rightarrow C(V, Q),$$

индуцированный  $\psi$ :

$$\bar{\psi}[a/J] = \psi(a) \quad (\text{независимо от выбора } a \text{ в } [a/J]).$$

Тогда элементы

$$\mathbf{v} \otimes \mathbf{v} - Q(\mathbf{v})1_{T(V)}$$

обязаны лежать в  $J$  в силу определения отображений Клиффорда. Следовательно,  $J$  должно содержать все суммы правых и левых кратных этих элементов. Поэтому давайте начнем сначала. Итак, пусть  $I$  является идеалом в  $T(V)$ , порожденным элементами  $\mathbf{v} \otimes \mathbf{v} - Q(\mathbf{v})1_{T(V)}$ , т. е.  $I$  состоит из всех (конечных) сумм выражений вида

$$a \cdot (\mathbf{v} \otimes \mathbf{v} - Q(\mathbf{v})1_{T(V)})b,$$

где  $a$  и  $b$  принадлежат  $T(V)$ , а  $\mathbf{v}$  — пространству  $V$ . В этом случае  $I$  является идеалом. Определим

$$C(V, Q) = T(V)/I$$

и

$$j : V \rightarrow C(V, Q), \quad j(\mathbf{v}) = [i(\mathbf{v})/I] = [\mathbf{v}/I].$$

По построению  $j$  является отображением Клиффорда. Читателю рекомендуется проверить, что все свойства универсальности выполняются. В частном случае, когда  $Q \equiv 0$ , условие Клиффорда принимает вид

$$f(\mathbf{u})f(\mathbf{v}) = -f(\mathbf{v})f(\mathbf{u}).$$

Тогда читателю рекомендуется проверить, что алгебра Клиффорда оказывается внешней алгеброй  $\Lambda(V)$ . Подробный анализ структуры алгебры Клиффорда приводится в этой главе после задач.

## Резюме

### А. Скалярные произведения и звездный оператор

Для данного ортонормированного базиса в пространстве  $\Lambda^1(V^*)$  надо уметь вычислять скалярные произведения в пространстве  $\Lambda^k(V^*)$ .

Для данного векторного пространства со скалярным произведением (не обязательно положительно определенным) и ориентацией с помощью определения звездного оператора надо уметь вычислять  $\star\lambda$  для произвольной базисной  $k$ -формы  $\lambda$ .

Вы должны уметь определять оператор Лапласа и выражать его через частные производные в декартовых координатах.

### В. Векторный анализ

Вы должны уметь определять  $\operatorname{div}$ ,  $\operatorname{grad}$  и  $\operatorname{rot}$  через  $d$  и  $\star$  и использовать эти определения для доказательства тождеств векторного анализа и для записи дифференциальных операторов в ортогональных системах координат.

## Задачи

- 18.1. В двумерном пространстве-времени имеются аффинные координаты  $t, x$  и скалярное произведение  $(\mathbf{e}_t, \mathbf{e}_t) = +1$ ,  $(\mathbf{e}_x, \mathbf{e}_x) = -1$ ,  $(\mathbf{e}_t, \mathbf{e}_x) = 0$ . Базисная 2-форма имеет вид  $dt \wedge dx$ .

- (а) Используя определение звездного оператора

$$\lambda \wedge \omega = (\star\lambda, \omega)\sigma,$$

вычислите  $\star 1$ ,  $\star dt$ ,  $\star dx$  и  $\star(dt \wedge dx)$ .

- (б) В этом пространстве вводится другая пара аффинных координатных функций

$$u = t - x, \quad v = t + x.$$

Вычислите  $du$ ,  $dv$ ,  $\star du$ ,  $\star dv$ .

- (с) Пусть  $f$  — дважды дифференцируемая функция в двумерном пространстве-времени. Вычислите ее лапласиан

$$\square f = -d\star df$$

через частные производные  $f$  по  $t$  и  $x$ .

(d) Вычислите  $\square f$  через частные производные  $f$  по  $u$  и  $v$ .

18.2. На плоскости вводятся полярные координаты. Вектора  $\mathbf{e}_r$  и  $\mathbf{e}_\theta$  образуют ортонормированный базис во всех точках плоскости, кроме начала координат, где угол  $\theta$  не определен. Пусть 1-формы  $dr$  и  $r d\theta$  дуальны этим базисным векторам.

(a) Используя определение звездного оператора

$$\omega \wedge \lambda = (\star \omega, \lambda) \sigma,$$

вычислите  $\star dr$  и  $\star d\theta$ .

(b) Пусть  $f$  — дважды дифференцируемая функция на плоскости. Вычислите ее лапласиан

$$\square f = -d\star df$$

через частные производные  $f$  по  $r$  и  $\theta$ .

18.3. Перепишите следующие тождества в обозначениях векторного анализа ( $f$  обозначает функцию,  $A$  — 1-форму, но ответы должны содержать вектор  $\mathbf{A}$ ).

(a)  $d(df) = 0$ .

(b)  $d(dA) = 0$ . (В пространстве  $\mathbb{R}^3$  оператор  $\star\star$  является тождественным.)

(c)  $d(fA) = df \wedge A + f dA$ . (Примените  $\star$  к обеим частям.)

(d)  $d(A \wedge B) = dA \wedge B - A \wedge dB$ . (Посмотрите на  $(\star d\star)(A \wedge B)$ .)

(e)  $d(f \star A) = df \wedge \star A + f d\star A$ . (Примените  $\star$  к обеим частям.)

18.4. Пусть  $f(t, x, y, z)$  — дважды дифференцируемая функция в четырехмерном пространстве-времени. Вычислите лапласиан  $\square f$  через частные производные  $f$ .

18.5. Векторное поле задано в сферических координатах  $\vec{B} = B_r \mathbf{e}_r + B_\theta \mathbf{e}_\theta + B_\phi \mathbf{e}_\phi$ . Вычисляя  $\star d\star B$ , где  $B$  — соответствующая 1-форма, получите выражение для  $\operatorname{div} \vec{B}$ .

18.6. (a) Пусть  $\vec{A} = A_x \mathbf{e}_x + A_y \mathbf{e}_y + A_z \mathbf{e}_z$ . Вычисляя  $(dd^\star + d^\star d)A$ , где  $A$  — соответствующая 1-форма, получите формулу для векторного лапласиана от  $\vec{A}$  через частные производные  $A_x, A_y, A_z$ .

(b) Прделайте то же самое в цилиндрических координатах, когда  $\vec{A} = A_r \mathbf{e}_r + A_\phi \mathbf{e}_\phi + A_z \mathbf{e}_z$ . Будьте внимательны: коэффициент при  $d\theta$  будет  $r A_\theta$ !

18.7. Получите выражение для  $\text{rot } \vec{A}$  в сферических координатах, т. е. начните с

$$\vec{A} = A_r \mathbf{e}_r + A_\theta \mathbf{e}_\theta + A_\phi \mathbf{e}_\phi,$$

где соответствующая 1-форма имеет вид

$$A = A_r dr + A_\theta r d\theta + A_\phi r \sin \theta d\phi.$$

Вычислите  $\star dA$  и представьте результат в виде вектора.

Цель последующего обсуждения — описать все (конечномерные) алгебры Клиффорда над вещественными числами. Вспомним, что алгебра Клиффорда  $C(V, W)$  полностью определена векторным пространством  $V$  и квадратичной формой  $Q$ . Отсюда и из свойства универсальности алгебры Клиффорда следует вывод.

Пусть имеется линейное отображение  $f : V \rightarrow V'$ , которое «сохраняет длину», т. е. удовлетворяет условию

$$Q'(f(\mathbf{v})) = Q(\mathbf{v}) \quad \text{для всех } \mathbf{v} \in V.$$

Тогда  $f$  индуцирует единственный гомоморфизм  $\phi_f : C(V, Q) \rightarrow C(V', Q')$ , для которого

$$\phi_f(\mathbf{v}) = f(\mathbf{v}) \quad \text{для } \mathbf{v} \in V,$$

если мы считаем, что  $V$  является подпространством  $C(V, Q)$ . В частности, если  $f$  — изометрия, то есть  $f$  также является линейным изоморфизмом  $V$  с  $V'$ , то  $\phi_f$  является изоморфизмом  $C(V, Q)$  с  $C(V', Q)$ . Таким образом, классификация алгебр Клиффорда совпадает с классификацией векторных пространств, снабженных квадратичными формами. Ответ на эту задачу классификации мы знаем: квадратичная форма полностью классифицирована ее сигнатурой  $(p, q)$ . Более точно, пусть  $V^\perp \subset V$  обозначает подпространство  $V$ , состоящее из всех векторов  $\mathbf{u}$ , которые удовлетворяют условию:  $(\mathbf{u}, \mathbf{v})_Q = 0$  для всех  $\mathbf{v} \in V$ . Тогда  $Q$  индуцирует невырожденное скалярное произведение в факторпространстве  $V/V^\perp$ . Из процедуры Грама–Шмидта мы знаем, что это индуцированное скалярное произведение определено с точностью до изоморфизма числами  $p$  и  $q$  знаков  $+$  или  $-$  в своей записи в ортонормированном базисе. Если мы выберем векторное пространство  $W$ , дополнительное к  $V^\perp$  в  $V$ , так что  $V = W \oplus V^\perp$ ,

то  $W$  будет изометрично  $V/V^\perp$ . Если  $\dim V = n$ , то  $V$  изометрично  $\mathbb{R}^n$ , где стандартный базис  $\delta_1, \dots, \delta_n$  ортогонален и выполняются равенства:

$$\begin{aligned} Q(\delta_i) &= 1, & 1 \leq i \leq p, \\ Q(\delta_i) &= -1, & p+1 \leq i \leq q, \\ Q(\delta_i) &= 0, & p+q < i \leq n. \end{aligned}$$

Итак, алгебра Клиффорда полностью задана целыми числами  $(n, p, q)$ , удовлетворяющими условию  $p + q \leq n$ . Нам хотелось бы увидеть, что собой представляет каждая из этих алгебр. Начнем с элементарного, но очень важного свойства. Изометрия  $f : V \rightarrow V'$  индуцирует изоморфизм  $\phi_f$  соответствующей алгебры Клиффорда. Рассмотрим отображение пространства  $V$  в себя, определенное формулой

$$f(\mathbf{v}) = -\mathbf{v}.$$

Это, очевидно, изометрия и  $f^2 = \text{id}$ . Поэтому оно индуцирует изоморфизм  $\phi_f$  алгебры Клиффорда  $C = C(V, Q)$  на себя. Этот изоморфизм будем обозначать  $\omega$ . Тогда

$$\omega = \phi_f.$$

Из общих принципов следует, что если  $f : V \rightarrow V'$  и  $g : V' \rightarrow V''$  — изометрии, то

$$\phi_{g \circ f} = \phi_g \circ \phi_f.$$

Возьмем  $V = V' = V''$  и  $g = f$ . Тогда получим

$$\phi_{f^2} = \omega^2.$$

Но  $f^2 = \text{id}$ . В силу единственности  $\phi$ ,  $\phi_{\text{id}}$  должно быть тождественным изоморфизмом  $C$ . Тогда

$$\omega^2 = \text{id}.$$

Мы говорим, что  $\omega$  является *инволюцией*  $C$ .

**$\mathbb{Z}_2$ -градуировка.** Пусть  $C_0$  является подпространством  $C$ , состоящим из всех  $c$ , которые удовлетворяют условию

$$\omega(c) = c,$$

и пусть  $C_1$  является подпространством, состоящим из  $c$ , которые удовлетворяют условию

$$\omega(c) = -c.$$

Тогда любое  $c$  можно записать в виде

$$c = c_0 + c_1, \quad \text{где} \quad c_0 = \frac{1}{2}(c + \omega(c)), \quad c_1 = \frac{1}{2}(c - \omega(c)),$$

при этом

$$c_0 \in C_0 \quad \text{и} \quad c_1 \in C_1.$$

Итак,

$$C = C_0 \oplus C_1$$

как векторное пространство. Кроме того, если

$$\omega(c) = (-1)^i c \quad \text{и} \quad \omega(c') = (-1)^{i'} c',$$

то

$$\omega(cc') = (-1)^{i+i'} cc'.$$

Поэтому

$$\begin{aligned} c_0 c'_0 &\in C_0, & \text{если} & \quad c_0 \in C_0 \text{ и } c'_0 \in C_0, \\ c_0 c'_1 &\in C_1, & \text{если} & \quad c_0 \in C_0 \text{ и } c'_1 \in C_1, \\ c_1 c'_0 &\in C_1, & \text{если} & \quad c_1 \in C_1 \text{ и } c'_0 \in C_0, \\ c_1 c'_1 &\in C_0, & \text{если} & \quad c_1 \in C_1 \text{ и } c'_1 \in C_1. \end{aligned}$$

Мы можем подвести итог:

$$\begin{aligned} C_0 \cdot C_0 &\subset C_0, \\ C_0 \cdot C_1 &\subset C_1, \\ C_1 \cdot C_0 &\subset C_1 \\ C_1 \cdot C_1 &\subset C_0. \end{aligned}$$

и

Мы говорим, что  $C$  является  $\mathbb{Z}_2$ -градуированной алгеброй. Теперь  $V \subset C_1$  и  $1 \in C_0$ , а каждый элемент  $C$  может быть записан (возможно, несколькими способами) как суммы произведений (единицы и) элементов пространства  $V$ . Таким образом,  $C_0$  состоит из тех элементов, которые можно записать в виде сумм произведений элементов  $V$ , имеющих четное число множителей, а  $C_1$  состоит из сумм произведений нечетного числа множителей.

**Скрученные тензорные произведения.** Возьмем две  $\mathbb{Z}_2$ -градуированные алгебры:  $A = A_0 \oplus A_1$  и  $B = B_0 \oplus B_1$ . Определим их *скрученное тензорное произведение*, обозначаемое  $A \hat{\otimes} B$ , следующим образом. Как векторное пространство  $A \hat{\otimes} B$  — это обычное тензорное произведение  $A \otimes B$  для  $A$  и  $B$ , т. е. в данном случае

$$\begin{aligned} A \hat{\otimes} B &= (A_0 \oplus A_1) \otimes (B_0 \oplus B_1) \\ &= A \otimes B_0 \oplus A_1 \otimes B_0 \oplus A_0 \otimes B_1 \oplus A_1 \otimes B_1. \end{aligned}$$

Умножение определяется правилами:

$$\begin{aligned} (a_0 \otimes b_0) \cdot (a' \otimes b') &= aa' \otimes b_0b', \\ (a \otimes b) \cdot (a'_0 \otimes b') &= aa'_0 \otimes bb', \\ (a \otimes b_1) \cdot (a_1 \otimes b') &= -aa'_1 \otimes b_1b', \end{aligned}$$

где нижний индекс обозначает, четный или нечетный элемент. Отсутствие нижнего индекса означает, что уравнение выполняется независимо от четности данного элемента. Запишем эти правила в сжатой форме:

$$(a \otimes b_i)(a'_j \otimes b') = (-1)^{ij} aa'_j \otimes b_i b'.$$

Другими словами, при перестановке  $a'_j$  и  $b_i$  надо умножать на  $(-1)^{ij}$ . Полезность понятия скрученного тензорного произведения состоит в следующем. Пусть у нас есть два векторных пространства  $V$  и  $W$ , где имеются квадратичные формы  $Q_V$  и  $Q_W$  соответственно. На прямой сумме пространств  $V \oplus W$  возьмем прямую сумму  $Q_{V \oplus W}$  этих форм. Это значит, что квадратичная форма  $Q_{V \oplus W}$  совпадает с  $Q_V$  на пространстве  $V$  и с  $Q_W$  на пространстве  $W$ , а  $V$  и  $W$  относительно нее ортогональны. Тогда мы имеем изоморфизм

$$C(V \oplus W, Q_{V \oplus W}) \cong C(V, Q_V) \hat{\otimes} C(W, Q_W).$$

Опишем это подробнее. Отображения  $f : V \rightarrow V \oplus W$ , где  $f(\mathbf{v}) = \mathbf{v} \oplus 0$ , и  $g : W \rightarrow V \oplus W$ , где  $g(\mathbf{w}) = 0 \oplus \mathbf{w}$ , «сохраняют длину» и поэтому индуцируют гомоморфизмы

$$\phi_f : C(V, Q_V) \rightarrow C(V \oplus W, Q_{V \oplus W})$$

и

$$\phi_g : C(W, Q_W) \rightarrow C(V \oplus W, Q_{V \oplus W}).$$

Отображение

$$\Phi : C(V, Q_V) \hat{\otimes} C(W, Q_W) \rightarrow C(V \oplus W, Q_{V \oplus W})$$

определим формулой

$$\Phi(a \hat{\otimes} b) = \phi_f(a) \cdot \phi_g(b).$$

Мы утверждаем, что  $\phi$  является изоморфизмом. Сначала мы должны показать, что  $\phi$  — гомоморфизм, т.е. что

$$\phi((a \otimes b) \cdot (a' \otimes b')) = \phi(a \otimes b) \cdot \phi(a' \otimes b').$$

Это достаточно проверить для

$$b = \mathbf{w}_1 \dots \mathbf{w}_i$$

и

$$a' = \mathbf{v}_1 \dots \mathbf{v}_j.$$

Далее, для  $\mathbf{v} \in V$  и  $\mathbf{w} \in W$  мы имеем  $(\mathbf{v}, \mathbf{w}) = 0$  в пространстве  $V \oplus W$ . (Мы пишем  $\mathbf{v}$  вместо  $f(\mathbf{v}) = \mathbf{v} \oplus 0$  и аналогично для  $\mathbf{w}$ .) Поэтому для алгебры Клиффорда  $C(V \oplus W, Q_{V \oplus W})$  мы имеем

$$\mathbf{v} \cdot \mathbf{w} + \mathbf{w} \cdot \mathbf{v} = 0$$

и, следовательно,

$$ba' = (-1)^{ij} a'b.$$

Это как раз и доказывает, что  $\phi$  является гомоморфизмом. Чтобы показать, что  $\phi$  является изоморфизмом, рассмотрим отображение  $h : V \oplus W \rightarrow C(V, Q_V) \hat{\otimes} C(W, Q_W)$ , заданное формулой

$$h(\mathbf{v} \oplus \mathbf{w}) = 1_V \hat{\otimes} \mathbf{w} + \mathbf{v} \hat{\otimes} 1_W,$$

где  $1_V$  — единица из  $C(V, Q_V)$ , а  $1_W$  — единица из  $C(W, Q_W)$ . Из определения  $\hat{\otimes}$  немедленно следует, что

$$h(\mathbf{v} \oplus 0)h(0 \oplus \mathbf{w}) + h(0 \oplus \mathbf{w})h(\mathbf{v} \oplus 0) = 0,$$

в то время как  $h(\mathbf{v} \oplus 0)^2 = Q(\mathbf{v})1_V$  и  $h(0 \oplus \mathbf{w})^2 = Q(\mathbf{w})1_W$ . Таким образом, в силу свойства универсальности  $C(V \oplus W, Q_{V \oplus W})$  отображение  $h$  индуцирует гомоморфизм

$$\phi_h : C(V \oplus W, Q_{V \oplus W}) \rightarrow C(V, Q_V) \hat{\otimes} C(W, Q_W)$$

и

$$\begin{aligned} \phi \circ \phi_h(\mathbf{v} \otimes \mathbf{w}) &= \phi(\mathbf{v} \hat{\otimes} 1_W + 1_V \hat{\otimes} \mathbf{w}) \\ &= \phi_f(\mathbf{v} \hat{\otimes} 1_W) + \phi_g(1_V \hat{\otimes} \mathbf{w}) \\ &= \mathbf{v} + \mathbf{w}. \end{aligned}$$

Итак,  $\phi \circ \phi_h = \text{id}$  в пространстве  $V \oplus W$ . Следовательно, в силу единственности

$$\phi \circ \phi_h = \text{id} \quad \text{на } V \otimes W,$$

т. е.  $\phi_h$  является правым обратным к  $\phi$ . Аналогичные соображения показывают, что  $\phi_h \circ \phi = \text{id}$ , а это значит, что  $\phi$  — изоморфизм.

Например, если  $U$  — векторное пространство с квадратичной формой  $Q \equiv 0$ , то его алгебра Клиффорда есть просто  $\Lambda(U)$  (это следует из определений). Пусть  $V$  — произвольное векторное пространство с квадратичной формой  $Q$ . Вспомним, что мы можем написать

$$V = V^\perp \oplus W,$$

где ограничение  $Q_W$  формы  $Q$  на подпространство  $W$  невырождено. Тогда  $\phi$  устанавливает изоморфизм между  $C(V, Q)$  и  $\Lambda(V^\perp) \hat{\otimes} C(W, Q_W)$ . Следовательно, с точностью до включения в скрученное тензорное произведение внешней алгебры, мы можем ограничиться только невырожденными квадратичными формами. Введем обозначение

$$C(p, q) = C(\mathbb{R}^n, Q_{p,q}),$$

где  $Q_{p,q}$  обозначает стандартную квадратичную форму в пространстве  $\mathbb{R}^n$  ( $p + q = n$ ), в которой  $p$  плюсов и  $q$  минусов. Нам хотелось бы понять структуру  $C(p, q)$ .

**Расщепленный случай.** В качестве примера рассмотрим  $p = q$ . В этом случае моделью для  $C(p, q)$  может быть  $U + U^*$ , где  $U$  —  $p$ -мерное векторное пространство, в котором возьмем  $(\mathbf{u}, \mathbf{u}') = 0$ ,  $(\mathbf{u}^*, \mathbf{u}'^*) = 0$  для  $\mathbf{u}, \mathbf{u}' \in U$ ,  $\mathbf{u}^*, \mathbf{u}'^* \in U^*$  и

$$(\mathbf{u}^*, \mathbf{u}) = \langle \mathbf{u}^*, \mathbf{u} \rangle \quad (\text{отображение вычисления}).$$

Как уже обсуждалось ранее в основном тексте, у нас есть описание соответствующей алгебры Клиффорда через операторы рождения и уничтожения в  $\Lambda(U)$ . Дальнейшее изучение покажет, что алгебра Клиффорда идентифицируется с алгебрами *всех* линейных преобразований на  $\Lambda(U)$ . Поскольку размерность  $\Lambda(U)$  равна  $2^p$ , то алгебру Клиффорда  $C(p, p)$  можно идентифицировать с  $\mathbb{R}(2^p)$ , где  $\mathbb{R}(n)$  обозначает алгебру всех вещественных матриц  $n \times n$ .

В случае  $p \neq q$  замечаем, что  $\mathbb{R}^n$  можно записать как прямую сумму  $\mathbb{R}^p$  с положительно определенной формой и  $\mathbb{R}^q$  с отрицательно определенной формой. Следовательно, по теореме о прямой сумме мы имеем

$$C(p, q) = C(p, 0) \hat{\otimes} C(0, q).$$

Если ввести обозначения  $C(p) = C(p, 0)$  и  $C(-q) = C(0, q)$ , то это уравнение переписывается в виде

$$C(p, q) = C(p) \hat{\otimes} C(-q).$$

Теперь надо понять структуру  $C(p)$  и  $C(-q)$ .

**Невырожденный случай.** Однако сначала мы рассмотрим произвольное векторное пространство  $V$  с невырожденной квадратичной формой  $Q$ . Положим

$$C = C(V, Q).$$

Пусть  $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  будет ортонормированным базисом  $V$  (при этом  $Q(\mathbf{e}_1) = \dots = Q(\mathbf{e}_p) = 1$  и  $Q(\mathbf{e}_i) = -1$  для  $i = p + 1, \dots, n$ ). Рассмотрим элемент

$$\gamma = \mathbf{e}_1 \cdots \mathbf{e}_n \quad \text{в } C.$$

Если

$$x_i = \sum_{j=1}^n a_{ij} \mathbf{e}_j, \quad i = 1, \dots, n,$$

— произвольные элементы пространства  $V$ , то из ортогональности  $\mathbf{e}_j$  следует, что

$$x_1 \cdots x_n = \text{Det}(a_{ij}) \mathbf{e}_1 \cdots \mathbf{e}_n = \text{Det}(a_{ij}) \gamma.$$

В частности, если  $\{\mathbf{f}_1, \dots, \mathbf{f}_n\}$  — другой ортогональный базис, то

$$\mathbf{f}_1 \cdots \mathbf{f}_n = \pm \mathbf{e}_1 \cdots \mathbf{e}_n.$$

Итак, с точностью до знака элемент  $\gamma$  не зависит от выбора базиса. (В алгебре Дирака  $C(1, 3)$  элемент  $\gamma$  в физической литературе обычно обозначается  $\gamma_5$ . Индекс 5 соответствует тому факту, что в четырехмерном случае  $\gamma$  является произведением четырех элементов.) Заметим, что

$$\gamma^2 = (-1)^{n(n-1)/2+q}.$$

Число  $n(n-1)/2$  возникает из-за перестановок  $\mathbf{e}_1$  после  $\mathbf{e}_2 \cdots \mathbf{e}_n$ ,  $\mathbf{e}_2$  после  $\mathbf{e}_3 \cdots \mathbf{e}_n$  и так далее. Число  $q$  появилось потому, что  $\mathbf{e}_{p+1}^2 = \cdots = \mathbf{e}_n^2 = -1$ , где  $n-p=q$ .

Наконец, заметим, что для любого  $i$  ( $1 \leq i \leq n$ ) справедливы равенства

$$\gamma \mathbf{e}_i = \mathbf{e}_1 \cdots \mathbf{e}_n \mathbf{e}_i = (-1)^{n-i} \mathbf{e}_1 \cdots \mathbf{e}_i \mathbf{e}_i \mathbf{e}_{i+1} \cdots \mathbf{e}_n$$

и

$$\mathbf{e}_i \gamma = \mathbf{e}_i \mathbf{e}_1 \cdots \mathbf{e}_n = (-1)^{i-1} \mathbf{e}_1 \cdots \mathbf{e}_i \mathbf{e}_i \mathbf{e}_{i+1} \cdots \mathbf{e}_n.$$

Тогда

$$\gamma \mathbf{e}_i = (-1)^{n-1} \mathbf{e}_i \gamma.$$

Поскольку  $\mathbf{e}_i$  образуют базис пространства  $V$ , мы имеем

$$\gamma \mathbf{v} = (-1)^{n-1} \mathbf{v} \gamma \quad \text{для всех } \mathbf{v} \in V.$$

Следовательно,

$$\gamma a = \omega(a)^{n-1} a \gamma \quad \text{для всех } a \in C.$$

**Примеры пространств малой размерности.** Рассмотрим два важных примера. Пусть  $V = \mathbb{R}^2$ . Мы утверждаем, что

$$C(2) \cong M(2) = \text{алгебре всех матриц } 2 \times 2,$$

где символ  $\cong$  обозначает изоморфизм. Чтобы это доказать, возьмем отображение

$$\phi : \mathbb{R}^2 \rightarrow M(2),$$

заданное формулой

$$\phi \left( \begin{pmatrix} x \\ y \end{pmatrix} \right) = \begin{pmatrix} x & y \\ y & -x \end{pmatrix}.$$

Тогда

$$\left( \phi \left( \begin{pmatrix} x \\ y \end{pmatrix} \right) \right)^2 = \begin{pmatrix} x^2 + y^2 & 0 \\ 0 & x^2 + y^2 \end{pmatrix} = (x^2 + y^2) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Откуда получаем

$$\phi(\mathbf{u})^2 = Q(\mathbf{v}) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

и, таким образом,  $M(2)$  обладает всеми свойствами алгебры Клиффорда  $C(2)$ . Если взять ортогональный базис  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ , то соответствующие элементы алгебры Клиффорда  $M(2)$  имеют вид

$$\mathbf{e}_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{и} \quad \mathbf{e}_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

так что

$$\gamma = \mathbf{e}_1 \mathbf{e}_2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Заметим, что  $\gamma^2 = -1$ , как и требуется. Четная часть  $C_0$  алгебры  $C$  состоит из всех матриц  $2 \times 2$ , которые коммутируют с  $\gamma$ . Это — множество конформных матриц. Нечетная часть  $C_1$  алгебры  $C$  в этом частном случае состоит из самого подпространства  $\mathbb{R}^2$ , представленного пространством симметричных матриц, имеющих нулевой след.

Мы уже рассмотрели случай  $C(-2)$  и показали, что эта алгебра изоморфна кватернионам. Давайте еще раз посмотрим на эту конструкцию. Пусть элементом  $i$  в  $C(-2)$  будет вектор  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ , а элементом  $j$  — вектор  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ . Тогда

$$i^2 = -1, \quad j^2 = -1, \quad ij + ji = 0$$

в силу свойств алгебры Клиффорда. Далее, положим

$$k = \gamma = ij.$$

Тогда

$$k^2 = -1$$

и

$$ki + ik = 0 = kj + jk.$$

Таким образом,  $C(-2)$  задается 1,  $i$ ,  $j$  и  $k$ , а приведенные выше соотношения определяют свойства кватернионов. Такая алгебра часто обозначается  $\mathbb{H}$ .

Пусть  $A$  и  $B$  — алгебры. Определим их тензорное произведение  $A \otimes B$  как алгебру, которая является векторным пространством  $A \otimes B$ , а закон умножения в ней имеет вид

$$(a \otimes b) \cdot (a' \otimes b') = aa' \otimes bb'.$$

(Заметим, что в отличие от скрученного тензорного произведения здесь не меняется знак.)

**Периодичность.** Мы хотим доказать следующие основные формулы:

$$\begin{aligned} C(p, q) \otimes C(2) &\cong C(q + 2, p), \\ C(p, q) \otimes C(1, 1) &\cong C(p + 1, q + 1), \\ C(p, q) \otimes C(-2) &\cong C(q, p + 2), \end{aligned}$$

где  $\cong$  обозначает изоморфизм.

Сначала давайте докажем первую формулу. Пусть  $V$   $(p+q)$ -мерное векторное пространство (скажем,  $\mathbb{R}^n$ ), снабженное квадратичной формой типа  $(p, q)$ . Далее пусть  $W = \mathbb{R}^2$  с положительно определенным скалярным произведением. Обозначим  $\gamma = \mathbf{e}_1 \mathbf{e}_2$  для  $C(W) = C(2)$ . Рассмотрим отображение

$$\psi : V \oplus W \rightarrow C(p, q) \otimes C(2),$$

заданное формулами

$$\psi(\mathbf{v}) = \mathbf{v} \otimes \gamma, \quad \mathbf{v} \in V$$

и

$$\psi(\mathbf{w}) = 1 \otimes \mathbf{w}.$$

Далее,

$$(\mathbf{v} \otimes \gamma)(1 \otimes \mathbf{w}) + (1 \otimes \mathbf{w})(\mathbf{v} \otimes \gamma) = \mathbf{v} \otimes (\gamma \mathbf{w} + \mathbf{w} \gamma) = 0,$$

откуда

$$\psi(\mathbf{v})\psi(\mathbf{w}) + \psi(\mathbf{w})\psi(\mathbf{v}) = 0.$$

Кроме того,

$$\psi(\mathbf{w})^2 = 1 \otimes \mathbf{w}^2 = \|\mathbf{w}\|^2 1$$

и

$$\begin{aligned} \psi(\mathbf{v})^2 &= (\mathbf{v} \otimes \gamma)(\mathbf{v} \otimes \gamma) = \mathbf{v}^2 \otimes \gamma^2 \\ &= \mathbf{v}^2 \otimes (-1) \\ &= -Q(\mathbf{v}) 1 \otimes 1. \end{aligned}$$

Итак, меняем  $Q$  на  $-Q$  в пространстве  $V$  и берем прямую сумму  $-Q$  на  $V$  и  $\|\cdot\|^2$  на  $W$ . Для этой квадратичной формы в пространстве  $V \oplus W$  отображение  $\psi$  удовлетворяет тождествам Клиффорда. Эта квадратичная форма имеет  $q+2$  плюсов и  $p$  минусов. Таким образом, отображение  $\psi$  продолжается до гомоморфизма

$$\psi : C(q+2, p) = C(V \oplus W) \rightarrow C(p, q) \otimes C(2).$$

В качестве упражнения читателю рекомендуем построить обратное отображение и проверить, что  $\psi$  является изоморфизмом. Те

же самые аргументы (только теперь  $\|\mathbf{w}\|^2$  отрицательно) дают изоморфизмы

$$C(p, q) \otimes C(1, 1) \cong C(p+1, q+1)$$

и

$$C(p, q) \otimes C(-2) \cong C(q, p+2).$$

Приведем ряд приложений этих изоморфизмов.

$$\begin{aligned} \mathbb{H} \otimes \mathbb{H} &= C(-2) \otimes C(-2) \\ &= C(0, 2) \otimes C(-2) \\ &\cong C(2, 2) \\ &\cong \mathbb{R}(2^2) = \mathbb{R}(4), \end{aligned}$$

следовательно,

$$\mathbb{H} \otimes \mathbb{H} \cong \mathbb{R}(4).$$

Далее,

$$C(1) \cong \mathbb{R} \oplus \mathbb{R}$$

для изоморфизма, переводящего

$$x1 + ye \rightarrow (x - y, x + y),$$

что можно легко проверить (упражнение). Таким образом,

$$\begin{aligned} C(-3) \sim C(1) \otimes C(-2) &\cong (\mathbb{R} \oplus \mathbb{R}) \otimes \mathbb{H} \\ &\cong \mathbb{H} \oplus \mathbb{H}. \end{aligned}$$

Мы уже проверили, что  $C(-1) \cong \mathbb{C}$ . Следовательно,

$$C(3) \cong C(-1) \otimes C(2) = \mathbb{C} \otimes \mathbb{R}(2).$$

Это алгебра комплексных матриц  $2 \times 2$ , обозначаемая  $\mathbb{C}(2)$ . Аналогично

$$\begin{aligned} \mathbb{C} \otimes \mathbb{H} &\cong C(-1) \otimes C(-2) \cong C(0, 1) \otimes C(0, 2) \\ &\cong C(1, 2) \\ &\cong C(0, 1) \otimes C(1, 1) \\ &\cong \mathbb{C} \otimes \mathbb{R}(2) = \mathbb{C}(2). \end{aligned}$$

Откуда  $\mathbb{C} \otimes \mathbb{H} \sim \mathbb{C}(2)$ .

Эти изоморфизмы мы можем использовать для заполнения таблицы всех элементов  $C(p, q)$ :

	0	1	2	3	4	...
0	$\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{H} \oplus \mathbb{H}$		
1	$\mathbb{R} + \mathbb{R}$	$\mathbb{R}(2)$				
2	$\mathbb{R}(2)$		$\mathbb{R}(4)$			
3	$\mathbb{C}(2)$					
4						
$\vdots$						

Например, чтобы найти  $C(1, 2)$ , мы элемент  $\mathbb{C} \cong C(0, 1)$  тензорно умножаем на  $\mathbb{R}(2) \cong C(1, 1)$  и показываем, что  $C(1, 2) \cong \mathbb{C}(2)$ . Аналогично, чтобы найти  $C(3, 1)$ , мы элемент  $\mathbb{R}(2) \cong C(2, 0)$  тензорно умножаем на  $\mathbb{R}(2)$  и находим, что  $C(3, 1) \cong \mathbb{R}(2) \otimes \mathbb{R}(2)$ . Однако, по определению матричной алгебры  $\mathbb{R}(r) \otimes \mathbb{R}(l) \cong \mathbb{R}(rl)$ . Поэтому  $\mathbb{R}(2) \otimes \mathbb{R}(2) \cong \mathbb{R}(4)$ . Аналогично

$$C(1, 3) \cong \mathbb{H} \otimes \mathbb{R}(2) = \mathbb{H}(2),$$

т.е. алгебра всех матриц  $2 \times 2$  с кватернионными элементами. (Заметим, что  $C(1, 3)$  и  $C(3, 1)$  не изоморфны.) Аналогично,

$$C(4, 0) \cong C(0, 2) \otimes C(2, 0) = \mathbb{H} \otimes \mathbb{R}(2) \cong \mathbb{H}(2)$$

и

$$C(0, 4) \cong C(2, 0) \otimes C(0, 2) \cong \mathbb{H}(2).$$

Продолжая таким образом, мы получаем следующую таблицу:

	$q$							
0	1	2	3	4	5	6	7	8
0 $\mathbb{R}$	$\mathbb{C}$	$\mathbb{H}$	$\mathbb{H} \oplus \mathbb{H}$	$\mathbb{H}(2)$	$\mathbb{C}(4)$	$\mathbb{R}(8)$	$\mathbb{R}(8) + \mathbb{R}(8)$	$\mathbb{R}(16)$
1 $\mathbb{R} \oplus \mathbb{R}$								
2 $\mathbb{R}(2)$								
3 $\mathbb{C}(2)$								
4 $\mathbb{H}(2)$								
5 $\mathbb{H}(2) \oplus \mathbb{H}(2)$								
6 $\mathbb{H}(6)$								
7 $\mathbb{C}(8)$								
8 $\mathbb{R}(16)$								

Остальные элементы таблицы могут быть получены, двигаясь вдоль диагонали и на каждом шагу тензорно умножая на  $\mathbb{R}(2)$ . Эта таблица  $8 \times 8$  определяет все алгебры Клиффорда, так как  $C(n+8, 0) \cong C(n, 0) \otimes \mathbb{R}(16)$  и  $C(0, n+8) \cong C(0, n) \otimes \mathbb{R}(16)$ , что следует из четырехкратного применения базисных изоморфизмов. Например, чтобы вычислить  $C(9)$ , мы начнем с  $C(1)$ , последовательно проводя тензорное умножение на  $C(-2)$ ,  $C(2)$ ,  $C(-2)$ ,  $C(2)$ . Это то же самое, что и

$$\begin{aligned} C(2) \otimes C(-2) \otimes C(2) \otimes C(-2) &\cong C(-4) \otimes C(2) \otimes C(-2) \\ &\cong C(6) \otimes C(-2) \\ &\cong C(8) \cong C(-8) \cong \mathbb{R}(16). \end{aligned}$$

Такие же соображения показывают, что

$$C(k+8) \cong C(k) \otimes \mathbb{R}(16)$$

и

$$C(-k-8) \cong C(-k) \otimes \mathbb{R}(16).$$

Таким образом,

$$C(8r+k) \cong C(k) \otimes \mathbb{R}(16^r) \quad (\text{периодичность Ботта})$$

и

$$C(-8r-k) \cong C(-k) \otimes \mathbb{R}(16^r).$$

Это полностью определяет структуру всех конечномерных алгебр Клиффорда.

## Глава 19

---

# Уравнения Максвелла

---

Главу 19 можно считать кульминацией курса. Результаты предыдущих глав применяются к изучению уравнений Максвелла и соответствующих волновых уравнений.

### 19.1. Уравнения

Электростатика и магнитостатика являются всего лишь приближенными теориями. Для полей, зависящих от времени, их надо заменить теорией Максвелла. Начнем с *закона индукции Фарадея*, согласно которому для любой поверхности  $S$ , натянутой на кривую  $\gamma$ , производная по времени от поля  $B$  связана с  $E$  равенством

$$\frac{d}{dt} \int_S B = - \int_\gamma E.$$

Перепишем это выражение в форме, удобной для пространства четырех измерений. Рассмотрим временной отрезок  $[a, b]$  и трехмерный цилиндр  $S \times [a, b]$ , границей которого является боковая поверхность — двумерный цилиндр  $\gamma \times [a, b]$  и торцевые поверхности — нижняя  $S \times \{a\}$  и верхняя  $S \times \{b\}$  (см. рис. 19.1). Проинтегрировав закон Фарадея по времени  $t$  от  $a$  до  $b$ , получаем уравнение

$$\int_{S \times \{b\}} B - \int_{S \times \{a\}} B + \int_{\gamma \times [a, b]} E \wedge dt = 0.$$

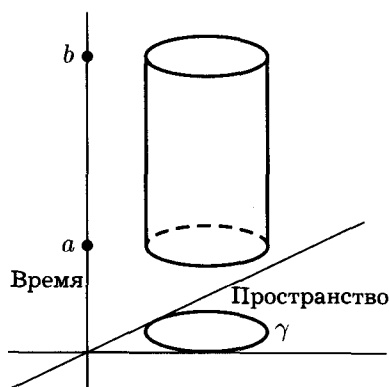


Рис. 19.1

Пусть

$$F = B + E \wedge dt,$$

так что  $F$  — это 2-форма в четырехмерном пространстве. И пусть  $C$  обозначает трехмерный цилиндр  $C = S \times [a, b]$ , у которого

$$\partial C = S \times \{b\} - S \times \{a\} + \gamma \times [a, b].$$

Здесь  $B$  — 2-форма, содержащая только пространственные производные и потому обращающаяся в нуль на боковой поверхности цилиндра  $\gamma \times [a, b]$ , а дифференциал  $dt$  и, следовательно,  $E \wedge dt$ , обращается в нуль на обоих торцах нашего цилиндра. Таким образом, закон индукции Фарадея можно записать в виде

$$\int_{\partial C} F = 0.$$

Кроме того, мы можем рассмотреть трехмерную область  $C$ , содержащую только пространственные координаты при фиксированном времени. В этом случае  $\partial C$  будет поверхностью, на которой  $dt = 0$ , так что

$$\int_{\partial C} F = \int_{\partial C} B$$

и, в силу отсутствия истинного магнетизма<sup>1</sup>, поверхностный интеграл должен быть равен нулю. Таким образом,  $\int_{\partial C} F = 0$  для

<sup>1</sup>т. е. отсутствия «магнитных зарядов». — Прим. ред.

всех трехмерных кубов, ребра которых параллельны любым трем из четырех координатных осей. Этого достаточно, чтобы выполнялось равенство  $dF = 0$ , где, конечно,  $d$  обозначает внешнюю производную в четырехмерном пространстве. Поскольку

$$F = B_x dy \wedge dz - B_y dx \wedge dz + B_z dx \wedge dy \\ + E_x dx \wedge dt + E_y dy \wedge dt + E_z dz \wedge dt,$$

уравнение  $dF = 0$  эквивалентно четырем уравнениям

$$\begin{aligned} \frac{\partial B_x}{\partial x} + \frac{\partial B_y}{\partial y} + \frac{\partial B_z}{\partial z} = 0, & \quad \frac{\partial B_y}{\partial t} - \frac{\partial E_z}{\partial x} + \frac{\partial E_x}{\partial z} = 0, \\ \frac{\partial B_x}{\partial t} - \frac{\partial E_y}{\partial z} + \frac{\partial E_z}{\partial y} = 0, & \quad \frac{\partial B_z}{\partial t} - \frac{\partial E_x}{\partial y} + \frac{\partial E_y}{\partial x} = 0. \end{aligned}$$

Мы воспользуемся законом Фарадея для определения единицы измерения  $B$

$$\frac{\text{напряжение} \cdot \text{время}}{\text{площадь}} = \frac{\text{энергия} \cdot \text{время}}{\text{заряд} \cdot (\text{длина})^2}.$$

Закон Ампера связывает ток с магнетизмом — поток электрического тока через поверхность  $S$ , имеющую границу  $\gamma$ , равняется циркуляции магнитного поля по контуру  $\gamma$ . Согласно великому открытию Максвелла, мы должны писать поток электрического тока в виде двух слагаемых  $\partial D / \partial t + 4\pi J$ , где  $D$  — электрическое смещение, а  $J$  — плотность тока движущихся зарядов. (Для медленно меняющихся полей первый член пренебрежимо мал по сравнению со вторым, и поэтому его нет в первоначальной формулировке закона Ампера.) Циркуляция магнитного поля получается интегрированием линейной дифференциальной формы  $H$ , называемой *напряженностью магнитного поля*, по  $\gamma$ . Тогда модифицированный Максвеллом закон Ампера имеет вид

$$\int_S \left( \frac{\partial D}{\partial t} + 4\pi J \right) = \int_\gamma H.$$

Рассмотрим трехмерный цилиндр  $C = S \times [a, b]$  и положим

$$G = D - H \wedge dt.$$

Проинтегрируем закон Ампера по  $t$  от  $a$  до  $b$ . Тогда получаем

$$\int_{\partial C} G = -4\pi \int_C J \wedge dt.$$

Если мы рассмотрим трехмерную область  $R$  при постоянном времени, то  $dt$  обращается в нуль на границе  $\partial R$  и интеграл от  $G$  по  $\partial R$  точно такой же, как интеграл от  $D$ , что равняется  $4\pi \times (\text{полный заряд внутри области } R) = 4\pi \int \rho dx \wedge dy \wedge dz$ . Таким образом, для областей при постоянном времени мы имеем

$$\int_{\partial R} G = 4\pi \int_R \rho dx \wedge dy \wedge dz.$$

Положим

$$j = \rho dx \wedge dy \wedge dz - J \wedge dt,$$

тогда мы видим, что

$$\int_{\partial C} G = 4\pi \int_C j$$

для любого трехмерного куба, ребра которого параллельны осям координат. Тогда

$$dG = 4\pi j,$$

откуда следует, что  $dj = 0$ .

Итак, подведем итоги:

$$\begin{array}{ll} \text{Положим} & \text{тогда из уравнений Максвелла имеем} \\ F = B + E \wedge dt, & dF = 0, \\ G = D - H \wedge dt & dG = 4\pi j. \end{array}$$

Заметим, что уравнения Максвелла инвариантны относительно гладкой замены координат.

Воспользуемся законом Ампера для определения  $H$ . Электрическое смещение  $D$  измеряется в единицах заряд/площадь, так что  $\partial D / \partial t$  измеряется в единицах заряд/(площадь · время), а плотность тока  $J$  — в единицах ток/площадь = заряд/(площадь · время). Таким образом,  $H$  измеряется в единицах заряд/(время · длина). В вакууме мы имеем конститутивные соотношения

$$D = \varepsilon_0 \star E \quad \text{и} \quad B = \mu_0 \star H,$$

где  $\varepsilon_0$  измеряется в единицах

$$\frac{\text{заряд}}{\text{площадь}} \times \frac{\text{длина}}{\text{напряжение}} = \frac{(\text{заряд})^2}{\text{энергия} \cdot \text{длина}},$$

а  $\mu_0$  — в единицах

$$\frac{\text{энергия} \cdot \text{время}}{\text{заряд} \cdot (\text{длина})^2} \times \frac{\text{время} \cdot \text{длина}}{\text{заряд}} = \frac{\text{энергия} \cdot (\text{время})^2}{(\text{заряд})^2 \cdot \text{длина}}.$$

При этом величина  $1/\varepsilon_0\mu_0$  имеет размерность

$$(\text{длина})^2/(\text{время})^2 = (\text{скорость})^2.$$

Таким образом, в теорию электромагнетизма принципиально встроена фундаментальная скорость. Великое открытие Максвелла состоит в том, что эта скорость является не чем иным, как скоростью света  $c$ ! Поэтому вместо  $dt$  введем  $c dt$ . Тогда четырехмерный оператор  $\star$  принимает вид

$$\begin{aligned}\star(dx \wedge dy) &= -c dt \wedge dz, \\ \star(dx \wedge dz) &= c dt \wedge dy, \\ \star(dy \wedge dz) &= -c dt \wedge dx, \\ \star(dx \wedge c dt) &= -dy \wedge dz, \\ \star(dy \wedge c dt) &= dx \wedge dz, \\ \star(dz \wedge c dt) &= -dx \wedge dy.\end{aligned}$$

В этом случае

$$\begin{aligned}\star F &= c(B_z dz \wedge dt + B_y dy \wedge dt + B_x dx \wedge dt) \\ &\quad - (1/c)(E_x dy \wedge dz - E_y dx \wedge dz + E_z dx \wedge dy).\end{aligned}$$

Тогда конститутивные соотношения можно записать в виде

$$G = -\frac{\varepsilon_0}{\mu_0} \star F.$$

Наконец, в вакууме, где величина  $\varepsilon_0/\mu_0$  постоянна, можно переопределить  $j$ , включив в него все константы, и записать уравнения Максвелла в виде

$$dF = 0 \quad \text{и} \quad d\star F = 4\pi j.$$

Из последней главы мы знаем процедуру решения такой системы уравнений: предположив соответствующее топологическое условие, мы ищем 1-форму  $A$ , удовлетворяющую условиям

$$dA = F, \quad d\star A = 0$$

и

$$-\square A = 4\pi\star j.$$

Таким образом, решения уравнений Максвелла тесно связаны с изучением *волнового оператора*

$$-\square = \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2},$$

если используется система координат, в которой  $c = 1$ .

## 19.2. Одномерное однородное волновое уравнение

В этом параграфе в качестве подготовки к изучению уравнений Максвелла рассмотрим уравнение, представляющее самостоятельный интерес<sup>2</sup>:

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} = 0.$$

Если мы сделаем замену переменных  $p = x + t$ ,  $q = x - t$ , то уравнение переписется в виде

$$\frac{\partial^2 u}{\partial p \partial q} = 0,$$

решениями которого являются функции

$$u = u_1(p) + u_2(q)$$

или

$$u(x, t) = u_1(x + t) + u_2(x - t). \quad (19.1)$$

---

<sup>2</sup>Оно известно также под названием уравнения колебаний струны и именно в этой интерпретации обычно изучается одним из первых в курсах уравнений с частными производными. — *Прим. ред.*

Очевидно, что любая функция такого вида будет решением. Функция  $u_2(x - t)$  может рассматриваться динамически. В каждый момент времени  $t$  график функции  $u_2(x - t)$  является графиком  $u_2(x)$ , смещенным вправо на величину  $t$ . Мы говорим, что функция  $u_2(x - t)$  описывает волну, движущуюся вправо без искажений. Таким образом, общее решение одномерного однородного волнового уравнения является суммой двух волн, распространяющихся вправо и влево без искажений.

Как правило, представляет интерес описание волнового движения, соответствующего некоторому начальному возмущению, скажем в момент  $t = 0$ . Поскольку мы имеем дело с уравнением второго порядка, то можно задать начальные значения  $u_0$  и  $v_0$ , где

$$u_0 = u(x, 0) \quad \text{и} \quad v_0 = \frac{\partial u}{\partial t}(x, 0).$$

Действительно, пусть  $A(x, t)$  обозначает операцию усреднения по интервалу длины  $2t$  (или радиуса  $t$ ) с центром в точке  $x$ . Тогда, по определению, для любой функции  $w$

$$A(x, t)w = \frac{1}{2} \int_{x-t}^{x+t} w(s) ds.$$

Утверждается, что

$$u(x, t) = A(x, t)v_0 + \frac{\partial}{\partial t}[A(x, t)u_0]$$

является единственным решением волнового уравнения с заданными начальными условиями. Действительно, записав  $u(x, t) = u_1(x + t) + u_2(x - t)$ , мы видим, что

$$u(x, 0) = u_1(x) + u_2(x)$$

и

$$\frac{\partial u}{\partial t}(x, 0) = u'_1(x) - u'_2(x).$$

Подставим  $v_0 = \partial u / \partial t$  в последнее уравнение и проинтегрируем. В результате получаем

$$u_1(x) - u_2(x) = \int_0^x v_0(s) ds + c,$$

где  $c$  — некоторая константа. Складывая и вычитая уравнение  $u_0 = u_1(x) + u_2(x)$ , видим, что

$$\begin{aligned} 2u_1(x) &= u_0(x) + \int_0^x v_0(s) ds + c, \\ 2u_2(x) &= u_0(x) - \int_0^x v_0(s) ds - c. \end{aligned}$$

Поэтому из формулы (19.1) имеем

$$u(x, t) = \frac{1}{2} \left( u_0(x+t) + u_0(x-t) + \int_{x-t}^{x+t} v_0(s) ds \right).$$

Это еще один способ записи решения.

Заметим, что мы доказали теорему существования и единственности для нашей задачи с начальными условиями. Из явной формулы также видно, что малые изменения начальных значений приводят к небольшим изменениям в решении. В этом случае мы говорим, что задача *корректна*. Кроме того, значения функции  $u$  в точке  $(x_0, t_0)$  зависят не от всех значений  $u$  и  $\partial u / \partial t$  при  $t = 0$ , а только от значений в интервале  $[x_0 - t_0, x_0 + t_0]$  (рис. 19.2).

С другой стороны, решение в точке  $(x, 0)$  влияет только на те точки пространства-времени, которые лежат внутри конуса  $\|x_0\| = \|t_0\|$  (рис. 19.3). Это — известный *принцип причинности*.

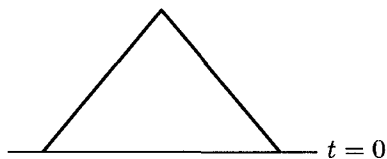


Рис. 19.2. Область зависимости

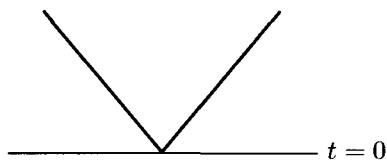


Рис. 19.3. Область влияния

### 19.3. Однородное волновое уравнение в пространстве $\mathbb{R}^3$

Мы хотим найти решение уравнения

$$\frac{\partial^2 u}{\partial t^2} - \frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} - \frac{\partial^2 u}{\partial z^2} = 0$$

с начальными условиями

$$\begin{aligned} u(\mathbf{x}, 0) &= u_0(\mathbf{x}), \\ \frac{\partial u}{\partial t}(\mathbf{x}, 0) &= v_0(\mathbf{x}). \end{aligned}$$

Пусть  $A(t, \mathbf{x})$  обозначает усреднение по сфере радиуса  $|t|$  с центром в точке  $\mathbf{x}$ . Тогда

$$A(t, \mathbf{x})f = \int_S f(\mathbf{x} + t\cdot)\tau.$$

Здесь  $f(\mathbf{x}, t\cdot)$  обозначает функцию на сфере  $S$  единичного радиуса, значение которой на любом векторе  $\mathbf{v}$  с  $\|\mathbf{v}\| = 1$  равно  $f(\mathbf{x} + t\mathbf{v})$  и  $\tau$  — элемент телесного угла, умноженный на  $1/4\pi$ . Утверждается, что функция  $u$ , заданная формулой

$$u(\mathbf{x}, t) = tA(t, \mathbf{x})v_0 + \frac{\partial}{\partial t}[tA(t, \mathbf{x})u_0],$$

является единственным решением нашей задачи.

Прежде всего, покажем, что эта функция является решением. Для этого проверим, что она удовлетворяет начальным условиям. Для малых значений  $t$  можно написать

$$u(\mathbf{x}, t) = A(t, \mathbf{x})u_0 + O(t).$$

Очевидно, что предел  $A(t, \mathbf{x})$  при усреднении по сферам с центром в точке  $\mathbf{x}$ , радиус которых стремится к нулю, равен  $u_0(\mathbf{x})$ . Таким образом,

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}).$$

Кроме того,

$$\frac{\partial u}{\partial t}(\mathbf{x}, 0) = \lim_{t \rightarrow 0} \left[ A(t, \mathbf{x})v_0 + \frac{\partial A}{\partial t}(t, \mathbf{x})u_0 \right],$$

поскольку все остальные члены суть  $O(t)$ . Однако,

$$\frac{\partial A}{\partial t}(0, \mathbf{x})u_0 = \int_S \sum \frac{\partial u}{\partial x_i}(\mathbf{x})x_i \tau = \sum \frac{\partial u}{\partial x_i}(\mathbf{x}) \int_S x_i \tau = 0,$$

поскольку среднее значение нечетной функции (такой, как координата  $x_i$ ) по единичной сфере должно обращаться в нуль.

Чтобы доказать, что  $u$  удовлетворяет волновому уравнению, достаточно доказать, для любой функции  $\omega$ , зависящей от  $\mathbf{x}$ , функция

$$W(\mathbf{x}, t) = tA(t, \mathbf{x})\omega$$

удовлетворяет волновому уравнению. Действительно, наше выражение для  $u$  является суммой двух членов: один вида  $W$ , а другой вида  $\partial W/\partial t$ . Если  $W$  является решением волнового уравнения, то и  $\partial W/\partial t$  также является его решением.

Для любой дважды дифференцируемой функции  $\omega$  мы имеем

$$\begin{aligned} \int_{|\mathbf{y}-\mathbf{x}| \leq |t|} \left( \frac{\partial^2 \omega}{\partial x^2} + \frac{\partial^2 \omega}{\partial y^2} + \frac{\partial^2 \omega}{\partial z^2} \right) dx dy dz &= \int_{|\mathbf{y}-\mathbf{x}| \leq |t|} d\star d\omega \\ &= \int_{S_{\mathbf{x},t}} \star d\omega, \end{aligned}$$

где  $S_{\mathbf{x},t}$  обозначает сферу с радиусом  $|t|$  и центром в точке  $\mathbf{x}$ .

На поверхности сферы с радиусом  $r$  мы можем написать

$$(\star d\omega) \wedge dr = \frac{\partial \omega}{\partial r} dx \wedge dy \wedge dz$$

и

$$dx \wedge dy \wedge dz = \tau \wedge dr.$$

Тогда

$$\frac{1}{4\pi t^2} \int_{S_{\mathbf{x},t}} \star d\omega = \frac{\partial}{\partial t} A(t, \mathbf{x})\omega.$$

Таким образом,

$$\begin{aligned} \frac{\partial}{\partial t} A(t, \mathbf{x})\omega &= \frac{1}{4\pi t^2} \int_{|\mathbf{y}-\mathbf{x}| \leq |t|} (\Delta \omega)(\mathbf{y}) dy \\ &= \frac{1}{4\pi t^2} \int_0^{|t|} d\rho \int_{|\mathbf{y}-\mathbf{x}|=\rho} (\Delta \omega)(\mathbf{y}) dS. \end{aligned}$$

Дифференцируя это выражение по  $t$ , получаем

$$\frac{\partial^2 A}{\partial t^2} = -\frac{1}{2\pi t^3} \int_0^t d\rho \int_{|\mathbf{y}-\mathbf{x}|=\rho} (\Delta\omega)(\mathbf{y}) dS + \frac{1}{4\pi t^2} \int_{|\mathbf{y}-\mathbf{x}|=\rho} \Delta\omega(\mathbf{y}) dS.$$

Теперь положим  $W(\mathbf{x}, t) = tA(t, \mathbf{x})\omega$ . В этом случае

$$\frac{\partial^2 W}{\partial t^2} = 2\frac{\partial}{\partial t} A(t, \mathbf{x})\omega + t\frac{\partial^2 A}{\partial t^2}(t, \mathbf{x})\omega = \frac{1}{4\pi t} \int_{|\mathbf{y}-\mathbf{x}|=|t|} \Delta\omega(\mathbf{y}) dS.$$

Но

$$\Delta W = \frac{t}{4\pi} \int_{|\mathbf{v}|=1} \Delta_{\mathbf{x}}\omega(\mathbf{x} + t\mathbf{v}) dS = \frac{1}{4\pi t} \int_{|\mathbf{y}-\mathbf{x}|=|t|} \Delta\omega(\mathbf{y}) dS,$$

следовательно,  $W$  — решение волнового уравнения.

Можно выстроить эти рассуждения по-другому и доказать единственность нашего решения, сведя трехмерное однородное волновое уравнение к одномерному. Мы доказали интегральную формулу

$$\int_{|\mathbf{y}-\mathbf{x}|\leq r} \Delta u(\mathbf{y}) d\mathbf{y} = r^2 \int_{|\mathbf{v}|=1} \frac{\partial}{\partial r} u(\mathbf{x} + \mathbf{vr}) d\omega.$$

Поэтому для функции  $u(\mathbf{x}, t)$  мы имеем

$$\int_{|\mathbf{y}-\mathbf{x}|\leq r} \Delta u(\mathbf{y}, t) d\mathbf{y} = 4\pi r^2 \frac{\partial}{\partial r} A(r, \mathbf{x})u.$$

Если  $u$  удовлетворяет уравнению

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

то это можно подставить в предыдущее уравнение и получить

$$4\pi r^2 A(r, \mathbf{x})u = \int_{|\mathbf{y}-\mathbf{x}|\leq r} \frac{\partial^2 u}{\partial t^2}(\mathbf{y}, t) d\mathbf{y} = \int_0^r d\rho \int_{|\mathbf{y}-\mathbf{x}|=\rho} \frac{\partial^2 u}{\partial t^2}(\mathbf{y}, t) dS.$$

Дифференцирование по  $r$  показывает, что функция

$$Z(r, t) = rA(r, \mathbf{x})u(\mathbf{x}, v)$$

удовлетворяет одномерному волновому уравнению

$$\frac{\partial^2 Z}{\partial r^2} - \frac{\partial^2 Z}{\partial t^2} = 0.$$

Таким образом,

$$Z(r, t) = rA(r, \mathbf{x})u(\mathbf{x}, v) = w_1(r + t) + w_2(r - t).$$

Поскольку  $Z(0, t) = 0$ , мы имеем

$$w_1(t) = -w_2(-t).$$

Написав  $w$  вместо  $w_1$ , получаем

$$Z(r, t) = w(r + t) - w(t - r).$$

Дифференцируем это выражение по  $r$  и переходим к пределу  $r \rightarrow 0$ . В результате получаем

$$u(\mathbf{x}, t) = A(0, \mathbf{x})u(\mathbf{x}, t) = 2w'(t).$$

Кроме того,

$$\frac{\partial Z}{\partial r} + \frac{\partial Z}{\partial t} = 2w'(r + t),$$

следовательно,

$$u(\mathbf{x}, r) = 2w'(r) = \lim_{t \rightarrow 0} 2w'(r + t) = \lim_{t \rightarrow 0} \frac{\partial Z}{\partial r} + \lim_{t \rightarrow 0} \frac{\partial Z}{\partial t}.$$

Подстановка выражения для  $Z$  в эту формулу дает

$$\begin{aligned} & u(\mathbf{x}, r) \\ &= \frac{\partial}{\partial r} \left( \frac{r}{4\pi} \int_{|\gamma|=1} u(\mathbf{x} + r\gamma, 0) dS + \frac{r}{4\pi} \int_{|\gamma|=1} \frac{\partial u}{\partial t}(\mathbf{x} + r\gamma, 0) dS \right). \end{aligned}$$

Но

$$u(\mathbf{x} + r\gamma, 0) = u_0(\mathbf{x} + r\gamma)$$

и

$$\frac{\partial u}{\partial t}(\mathbf{x} + r\gamma) = v_0(\mathbf{x} + r\gamma).$$

Если в этом соотношении положить  $r = t$ , то получим наше исходное решение, что доказывает его единственность.

## 19.4. Неоднородное волновое уравнение в пространстве $\mathbb{R}^3$

Мы можем решать уравнения Максвелла, решая последовательность уравнений вида

$$\square u = f, \quad \square = \frac{\partial^2}{\partial t^2} - \Delta.$$

Давайте сначала покажем, как решить это уравнение с начальными условиями

$$u(\mathbf{x}, 0) \equiv 0, \quad \frac{\partial u}{\partial t}(\mathbf{x}, 0) \equiv 0.$$

Неоднородное волновое уравнение можно решить, сводя его к решению однородного волнового уравнения с последовательностью начальных условий следующим образом. Пусть  $v(\mathbf{x}, t, \tau)$  — решение задачи с начальными условиями

$$\square v = 0, \quad v(\mathbf{x}, \tau, \tau) = 0, \quad \frac{\partial v}{\partial t}(\mathbf{x}, \tau, \tau) = f(\mathbf{x}, \tau).$$

Положим

$$u(\mathbf{x}, t) = - \int_0^t v(\mathbf{x}, t, \tau) d\tau.$$

Тогда очевидно, что  $u(\mathbf{x}, 0) = 0$  и

$$\frac{\partial u}{\partial t}(\mathbf{x}, t) = -v(\mathbf{x}, t, t) - \int_0^t \frac{\partial v}{\partial t}(\mathbf{x}, t, \tau) d\tau.$$

Но  $v(\mathbf{x}, t, t) = 0$  в силу начальных условий для  $v$ . Поэтому

$$\frac{\partial u}{\partial t}(\mathbf{x}, t) = - \int_0^t \frac{\partial v}{\partial t}(\mathbf{x}, t, \tau) d\tau$$

и, следовательно,

$$\frac{\partial u}{\partial t}(\mathbf{x}, 0) = 0.$$

Кроме того, дифференцирование уравнения для  $\partial u / \partial t$  дает

$$-\frac{\partial^2 u}{\partial t^2}(\mathbf{x}, t) = \frac{\partial v}{\partial t}(\mathbf{x}, t, t) + \int_0^t \frac{\partial^2 v}{\partial t^2}(\mathbf{x}, t, \tau) d\tau,$$

а

$$\Delta u = - \int_0^t v(\mathbf{x}, t, \tau) d\tau,$$

так что

$$\left( \frac{\partial^2}{\partial t^2} - \Delta \right) u = f.$$

Теперь функция  $v$  определяется выражением

$$v(\mathbf{x}, t, \tau) = (t - \tau) A(\mathbf{x}, t - \tau) f = \frac{1}{4\pi} (t - \tau) \int_{|\gamma|=1} f(\mathbf{x} + (t - \tau)\gamma, \tau) d\gamma,$$

следовательно,

$$\begin{aligned} u(\mathbf{x}, t) = & -\frac{1}{4\pi} \int_0^t (t - \tau) \int_{|\gamma|=1} f(\mathbf{x} + (t - \tau)\gamma, \tau) d\gamma d\tau \\ & - \frac{1}{4\pi} \int_0^t \tau' \int_{|\gamma|=1} f(\mathbf{x} + \tau'\gamma, t - \tau') d\gamma d\tau'. \end{aligned}$$

В этом выражении положим  $\mathbf{y} = \mathbf{x} + \tau'\gamma$ , тогда последний интеграл принимает вид

$$\begin{aligned} -\frac{1}{4\pi} \int_0^t \frac{1}{\tau'} \int_{|\mathbf{y}-\mathbf{x}|=\tau'} f(\mathbf{y}, t - \tau') dS d\tau' \\ = -\frac{1}{4\pi} \int_{|\mathbf{y}-\mathbf{x}|=\tau'} \frac{f(\mathbf{y}, t - \|\mathbf{y} - \mathbf{x}\|)}{\|\mathbf{y} - \mathbf{x}\|} d\mathbf{y}. \end{aligned}$$

Итак, мы получили формулу для запаздывающих потенциалов:

$$u(\mathbf{x}, t) = -\frac{1}{4\pi} \int \frac{f(\mathbf{y}, t - \|\mathbf{y} - \mathbf{x}\|)}{\|\mathbf{y} - \mathbf{x}\|} d\mathbf{y}.$$

Она похожа на формулу Пуассона для потенциала, но вклад от сферы с радиусом  $a$  и центром в точке  $\mathbf{x}$  задается величиной  $1/a$ , умноженной на значение  $f$  в момент времени, равный  $a$  единиц.

Раз мы получили решение неоднородного уравнения с нулевыми начальными условиями, то теперь решение неоднородного уравнения с произвольными начальными условиями можно свести к решению однородного уравнения путем вычитания уже полученного решения.

Воспользуемся последней формулой, чтобы решить неоднородное волновое уравнение для каждой компоненты 4-потенциала  $A$  в уравнениях Максвелла. Получим

$$A(\mathbf{x}) = - \int \frac{\star j(\mathbf{x}, t - |\mathbf{x} - \mathbf{x}'|)}{|\mathbf{x} - \mathbf{x}'|} d\mathbf{x}.$$

Это решение называется *потенциалом Льенара-Вихерта* для заданного распределения плотности тока  $j$ . Общее решение уравнений Максвелла получается прибавлением к  $A$  потенциала, компоненты которого удовлетворяют однородному волновому уравнению.

## 19.5. Лагранжиан электромагнитного поля и тензор энергии-импульса

Пусть  $A$  — 1-форма и  $j$  — 3-форма в пространстве Минковского  $\mathbb{R}^{1,3}$ . Рассмотрим 4-форму

$$\mathcal{L} = \mathcal{L}(A, j) = -\frac{1}{2} dA \wedge \star dA - 4\pi A \wedge \star j.$$

Проинтегрируем  $\mathcal{L}$  по любой ограниченной четырехмерной области  $N$  и получим функцию

$$L_{N,j}(A) = \int_N \mathcal{L}.$$

Здесь мы считаем, что  $N$  и  $j$  заданы и фиксированы, так что  $L$  является только функцией  $A$ . Мы говорим, что  $A$  является *экстремалью*  $L$ , если

$$\frac{d}{ds} L_{N,j}(A_s)|_{s=0} = 0$$

для любого дифференцируемого семейства 1-форм  $A_s$ , удовлетворяющих условиям  $A_0 = A$  и  $A_s = A$  на границе  $\partial N$ . Пусть  $\frac{dA_s}{ds}|_{s=0} = B$ . Тогда

$$\frac{d}{ds} (A_s, j)|_{s=0} = -\frac{1}{2} (dB \wedge \star dA + dA \wedge \star dB) - 4\pi B \wedge j.$$

Теперь

$$d(B \wedge \star dA) = dB \wedge \star dA + B \wedge d\star dA$$

и

$$dB \wedge \star dA = dA \wedge \star dB,$$

следовательно,

$$\left. \frac{d}{ds}(A_s, j) \right|_{s=0} = B \wedge [d\star dA - 4\pi j] = d(B \wedge \star dA).$$

Интеграл  $\int_{\partial N} B \wedge \star dA$  обращается в нуль, потому что на границе  $B = 0$ . Таким образом,

$$\left. \frac{dL_{N,j}(A_s)}{ds} \right|_{s=0} = \int_N B \wedge [d\star dA - 4\pi j].$$

Если это выражение равно нулю для *всех*  $B$ , обращающихся в нуль на границе, то мы должны иметь

$$d\star dA = 4\pi j.$$

Однако, в области, где  $H^2 = 0$ , это эквивалентно уравнению Максвелла  $F = dA$ . Чтобы эти уравнения можно было решить, необходимо выполнение условия

$$dj = 0.$$

Таким образом, мы вывели уравнения Максвелла из *вариационного принципа*: уравнения Максвелла в точности совпадают с уравнениями для *критических 1-форм*  $A$  лагранжиана  $L$ . Формулировка в терминах лагранжиана обладает рядом преимуществ. Одно из них состоит в том, что изменяя лагранжиан  $L$ , мы получаем процедуру модификации уравнений для его экстремалей и, таким образом, можем открывать различные законы природы. Другое преимущество в том, что с помощью лагранжева формализма мы получаем процедуру перехода от классической механики к квантовой. Функция  $L$  играет ключевую роль в квантовой электродинамике. Третье преимущество — в рамках лагранжева формализма каждая однопараметрическая группа симметрий

лагранжиана  $L$  дает нам закон сохранения с помощью метода, известного как *теорема Нётер*. Этот метод мы объясним в частном случае электромагнетизма.

Пусть  $\xi$  — векторное поле в пространстве  $\mathbb{R}^{1,3}$ , обладающее свойством

$$D_\xi \star \omega = \star D_\xi \omega \quad \text{для всех 2-форм } \omega.$$

Как мы уже видели, любое конформное преобразование пространства-времени сохраняет действие звездного оператора на 2-формы. Поэтому  $\xi$  может быть векторным полем, связанным с любой однопараметрической группой конформных преобразований. В частности,  $\xi$  может быть любым постоянным векторным полем, что соответствует трансляциям, или любым инфинитезимальным преобразованием Лоренца.

Итак,

$$D_\xi \mathcal{L} = -\frac{1}{2} dD_\xi A \wedge \star dA - \frac{1}{2} dA D_\xi \star dA - D_\xi A \wedge 4\pi j - A \wedge 4\pi D_\xi j.$$

Поскольку

$$D_\xi \star dA = \star D_\xi dA = \star dD_\xi A$$

и

$$dD_\xi A \wedge \star dA = dA \wedge \star dD_\xi A,$$

первый и второй члены равны, поэтому они складываются в

$$-dD_\xi A \wedge \star dA = -d(D_\xi A \wedge \star dA) + D_\xi A \wedge d\star dA.$$

Таким образом,

$$D_\xi \mathcal{L} = -d(D_\xi A \wedge \star dA) + D_\xi A \wedge (d\star dA - 4\pi j) - A \wedge 4\pi D_\xi j.$$

В силу уравнений Максвелла  $d\star dA = 4\pi j$ . Поэтому

$$\begin{aligned} D_\xi \mathcal{L} &= -d(D_\xi A \wedge \star dA) - A \wedge 4\pi D_\xi j \\ &= -d((i(\xi)dA + di(\xi)A) \wedge \star dA) - A \wedge 4\pi D_\xi j \\ &= -d[i(\xi)F \wedge \star F] - di(\xi)A \wedge d\star dA - A \wedge 4\pi D_\xi j \\ &= -d[i(\xi)F \wedge \star F] - di(\xi)A \wedge 4\pi j - A \wedge 4\pi D_\xi j, \end{aligned}$$

где мы использовали  $dA = F$  и  $d\star dA = j$ . С другой стороны

$$\begin{aligned} D_{\xi}A &= i(\xi)dA + dA + di(\xi)A \\ &= i(\xi)F + di(\xi)A, \end{aligned}$$

а

$$\begin{aligned} D_{\xi}\mathcal{L} &= -\frac{1}{2}D_{\xi}(F \wedge \star F) = D_{\xi}(A \wedge 4\pi j) \\ &= \frac{1}{2}di(\xi)(F \wedge \star F) - (i(\xi)F + di(\xi)A) \wedge j - A \wedge D_{\xi}j, \end{aligned}$$

поскольку  $d(F \wedge \star F) = 0$ , потому что  $F \wedge \star F$  — 4-форма. Приравнивая эти два выражения для  $D_{\xi}\mathcal{L}$ , получаем

$$\begin{aligned} &-d[i(\xi)F \wedge \star F] - di(\xi)A \wedge 4\pi j - A \wedge D_{\xi}j \\ &= -\frac{1}{2}d[i(\xi)F \wedge \star F + F \wedge i(\xi)\star F] - i(\xi)F \wedge j - di(\xi)A \wedge 4\pi j - A \wedge 4\pi D_{\xi}j \end{aligned}$$

или

$$d\frac{1}{2}[F \wedge i(\xi)\star F - i(\xi)F \wedge \star F] = -i(\xi)F \wedge 4\pi j.$$

Если нет токов, т. е. если  $j = 0$ , оказывается, что 3-форма

$$C(\xi) = \frac{1}{2}[F \wedge i(\xi)\star F - i(\xi)F \wedge \star F]$$

замкнута. Этот факт интерпретируется как закон сохранения: интегралы от  $C(\xi)$  по двум трехмерным областям, ограничивающим некую четырехмерную область, равны.

Если при стремлении к бесконечности в пространственных направлениях  $F$  достаточно быстро убывает, то два интеграла от  $C(\xi)$  по любым трехмерным пространственноподобным поверхностям равны.

Например, в некоторой системе координат возьмем трехмерные поверхности, заданные уравнениями  $t = t_1$  и  $t = t_2$  (рис. 19.4). Полное значение величины  $C(\xi)$  в момент  $t = t_1$  равно полному значению  $C(\xi)$  в момент  $t = t_2$ . Это значит, что полное значение величины  $C(\xi)$  сохраняется во времени. Поэтому уравнение  $dC(\xi) = 0$  называется законом сохранения.

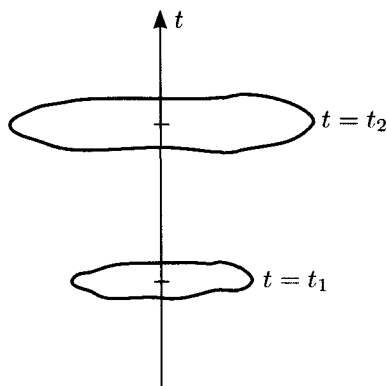


Рис. 19.4

Например, давайте возьмем  $\xi = \partial/\partial t$ , что соответствует инфинитезимальной трансляции во времени (в смысле некоторого расщепления на пространство и время). Прямое, но весьма поучительное вычисление, которое мы оставляем читателю, показывает, что

$$C(\xi) = \frac{1}{2}(\|\mathbf{E}\|^2 + \|\mathbf{B}\|^2) dx \wedge dy \wedge dz \\ + P_1 dy \wedge dz \wedge dt - P_2 dx \wedge dz \wedge dt + P_3 dx \wedge dy \wedge dt,$$

где

$$P_1 = E_y B_z - B_y E_z, \\ P_2 = E_x B_z - B_x E_z, \\ P_3 = E_x B_y - B_x E_y,$$

или, более кратко,

$$\vec{\mathbf{P}} = \vec{\mathbf{E}} \times \vec{\mathbf{B}}.$$

(Трехмерный вектор  $\vec{\mathbf{P}}$  называется вектором Пойнтинга.) Таким образом, на поверхности  $t = \text{const}$  интеграл от  $C(\xi)$  дает полную энергию электрического и магнитного полей, т.е. наш закон сохранения становится законом сохранения энергии. Остальные компоненты  $C(\xi)$  в этом случае описывают поток (локальной плотности энергии) через любую поверхность.

Точно так же, как энергия связана с временной трансляцией,  $x$ -компонента импульса связана с бесконечно малым смещением в направлении  $x$ . Это общий принцип механики (см., например, Loomis and Sternberg, глава 13, или более сложное изложение в книге Guillemin and Sternberg, глава 2). Таким образом,  $C(\partial/\partial x)$  будет замкнутой 3-формой, интеграл от которой по пространственноподобной трехмерной поверхности  $t = \text{const}$  дает полный импульс электромагнитного поля в направлении  $x$ . Интегрирование по поверхностям вида  $[a, b] \times S$ , где  $S$  — поверхность в пространстве  $\mathbb{R}^3$ , а  $[a, b]$  — временной интервал, величины  $C(\partial/\partial x)$  дает поток  $x$ -компоненты импульса через поверхность  $S$  в течение времени  $[a, b]$ .

Для любого постоянного вектора  $\xi$  величина  $C(\xi)$  является 3-формой, а  $\star C(\xi)$  — 1-формой. Таким образом, равенство  $T(\xi, \eta) = \star C(\xi)(\eta)$  имеет смысл для некоторого другого постоянного вектора  $\eta$ . Функция  $T = T(\xi, \eta)$  называется *тензором энергии-импульса*, потому что в ней закодированы все компоненты плотности и потока энергии и импульса.

При наличии тока  $j$  формы  $C(\xi)$  уже не будут замкнутыми, но они удовлетворяют условию

$$dC(\xi) = i(\xi)F \wedge 4\pi j.$$

Это называется *силой Лоренца* и представляет передачу энергии-импульса от электромагнитного поля току зарядов.

## 19.6. Волновые фронты и принцип Гюйгенса

Решение волнового уравнения

$$\frac{\partial^2 u}{\partial t^2} - \Delta u = 0$$

вида

$$u(\mathbf{x}, t) = f(\mathbf{x})g(h(\mathbf{x}) \pm t)$$

называется *распространяющейся волной*. Выражение  $h(\mathbf{x} \pm t)$  называется *фазой* волны, а поверхности уровня  $h = \text{const}$  — *волновыми фронтами*.

Мы уже видели, что в одномерном пространстве, взяв  $f \equiv 1$ ,  $h(\mathbf{x}) = x$  и произвольную функцию  $g$ , получаем решение, которое в общем случае может быть записано как сумма двух таких выражений, одно со знаком  $+$ , а другое со знаком  $-$ .

В трехмерном пространстве в качестве  $h$  возьмем функцию, линейную по  $\mathbf{x}$

$$h(\mathbf{x}) = \mathbf{k} \cdot \mathbf{x},$$

$f \equiv 1$  и произвольную функцию  $g$ . Чтобы решить уравнение, накладывается единственное условие  $\|\mathbf{k}\| = 1$ . (По существу мы свели задачу к одномерному случаю: с помощью вращений можно сделать так, что  $\mathbf{k} = (1, 0, 0)$ . Поэтому решение зависит только от  $x$  и  $t$ , что равносильно решению волнового уравнения в пространстве одного измерения.)

В частности, если мы возьмем  $g(\theta) = e^{i\theta}$ , то получим функции

$$e^{i(\mathbf{k} \cdot \mathbf{x} - t)}, \quad \|\mathbf{k}\| = 1.$$

Важность этих решений с синусоидальными волнами объясняется следующим: это теорема, согласно которой каждая гладкая функция в пространстве  $\mathbb{R}^{1,3}$  (или в общем случае в пространстве  $\mathbb{R}^n$ ), которая на бесконечности достаточно быстро убывает, может быть записана в виде суперпозиции функций вида  $e^{i\mathbf{k} \cdot \mathbf{x}}$ , т. е. любую такую функцию можно записать в форме

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{3/2}} \int \hat{f}(\boldsymbol{\xi}) e^{i\boldsymbol{\xi} \cdot \mathbf{x}} d\mathbf{x},$$

где  $\hat{f}$  — подходящая функция. В главе 21 мы дадим очень краткое доказательство этого факта. Оказывается, что эта формула *обратного преобразования Фурье* верна и в более общем случае. В нашем же случае общее решение волнового уравнения может быть записано как суперпозиция функций  $e^{i(\mathbf{k} \cdot \mathbf{x} - t)}$ , где  $\|\mathbf{k}\| = 1$ .

В качестве второго примера рассмотрим сферически симметричные решения волнового уравнения (определенного вне прямой  $r = 0$ ). Запишем лапласиан в сферических координатах:

$$\Delta u = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial u}{\partial r} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial u}{\partial \theta} + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 u}{\partial \phi^2},$$

так что для сферически симметричной функции  $u = u(r, t)$  волновое уравнение принимает вид

$$\frac{\partial^2 u}{\partial t^2} = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial u}{\partial r} = \frac{1}{r} \left[ 2 \frac{\partial u}{\partial r} + \frac{\partial^2 u}{\partial r^2} \right] = \frac{1}{r} \frac{\partial^2}{\partial r^2} (ru).$$

Таким образом, функция  $v = ru$  удовлетворяет одномерному волновому уравнению

$$\left( \frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial r^2} \right) v = 0.$$

Общее решение этого уравнения дается формулой

$$v(r, t) = f(r + t) + g(r - t),$$

так что общее решение сферически симметричного волнового уравнения имеет вид

$$u(r, t) = \frac{f(r + t)}{r} + \frac{g(r - t)}{r}.$$

Здесь первый член представляет падающую волну, а второй — испускаемую. В частности, возьмем  $f = 0$  и  $g(s) = e^{iks}$ . Тогда

$$w_k(t, r) = \frac{e^{ik(r-t)}}{r}$$

представляет испускаемую (синусоидальную) волну с частотой  $k$ . Действительно, легко проверить, что с точностью до нормировочной постоянной

$$E_k(r) = \frac{e^{ikr}}{r}$$

является «фундаментальным решением» приведенного волнового оператора  $\Delta + k^2$ , т. е.

$$(\Delta + k^2)E_k = 0$$

при  $r \neq 0$ , и

$$-\frac{1}{\pi} \int E_k(\Delta + k^2)\phi \, d\mathbf{x} = \phi(0)$$

для любой гладкой функции  $\phi$ . В главе 15 мы доказали этот результат в случае  $k = 0$  из формулы Грина. Аналогичное доказательство работает и здесь.

Далее, если  $D$  — любая ограниченная область в  $\mathbb{R}^3$  и  $u$  удовлетворяет приведенному волновому уравнению (уравнению Гельмгольца)

$$\Delta u + k^2 u = 0,$$

то из формулы Грина можно видеть, что верна следующая формула Гельмгольца:

$$\frac{1}{4\pi} \iint_{\partial D} \left[ \frac{e^{ikr_P}}{r_P} \star du - u \star d \left( \frac{e^{ikr_P}}{r_P} \right) \right] = \begin{cases} u(P), & \text{если } P \in D \\ 0, & \text{если } P \notin D, \end{cases}$$

где  $r_P$  обозначает расстояние от  $P$ .

Во многих приложениях нам интересна ситуация, когда вместо ограниченности области  $D$  она является внешней для некоторой поверхности  $S$ . Сначала воспользуемся полученной формулой для ограниченной области  $D_R$ , содержащей пересечение  $D$  с шаром радиуса  $R$  и центром в точке  $P$ . Если  $R$  выбирается достаточно большим (рис. 19.5), то интеграл становится суммой двух интегралов по поверхностям, и мы получаем формулу

$$\frac{1}{4\pi} \iint_S + \frac{1}{4\pi} \iint_{\Sigma_R} = \begin{cases} u(P), & \text{если } P \in D, \\ 0, & \text{если } P \notin D. \end{cases}$$

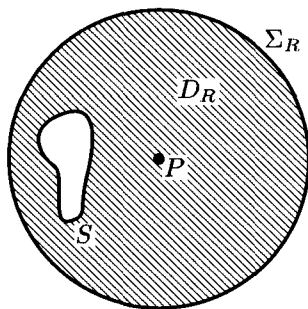


Рис. 19.5

Здесь  $\Sigma_R$  — сфера радиуса  $R$ . Тогда

$$d\left(\frac{e^{ikr}}{r}\right) = \frac{e^{ikr}}{r} \left[ ik - \frac{1}{r} \right] dr$$

и  $\star dr = R^2 d\omega$  на поверхности  $\Sigma_R$ , где  $d\omega$  — элемент телесного угла на этой поверхности. Таким образом, второй интеграл принимает вид

$$\iint e^{ikr} \left[ r \left( \frac{\partial u}{\partial r} - iku \right) + u \right]_{r=R} d\omega.$$

При  $R \rightarrow \infty$  интеграл по сфере стремится к нулю, если

$$\iint |u| d\omega = o(1) \quad \text{и} \quad \iint \left| \frac{\partial u}{\partial r} - iku \right| d\omega = o(R^{-1}).$$

Здесь интегралы вычисляются при  $r = R$ . Эти условия называются условиями излучения Зоммерфельда. Их значимость определяется тем, что они представляют ситуацию, когда решение  $u$  содержит только испускаемые волны, а падающих волн нет<sup>3</sup>. Предположим, что это условие выполнено. Тогда значение  $u$  вне некоторой поверхности  $S$  задается формулой

$$u(P) = \frac{1}{4\pi} \int_S \left[ \frac{e^{ikr}}{r} \star du - u \star d\left(\frac{e^{ikr}}{r}\right) \right]. \quad (19.2)$$

<sup>3</sup>Строгое математическое объяснение условий излучения Зоммерфельда можно найти в книге Lax and Phillips, с. 120–128. Идея доказательства состоит в следующем. Пусть  $f = \{f_1, f_2\}$  — данные Коши для волнового уравнения. Мы ищем решение волнового уравнения

$$\frac{\partial^2 w}{\partial t^2} - \Delta w = 0$$

с начальными условиями  $w(x, 0) = f_1$  и  $(\partial w / \partial t)(x, 0) = f_2$ . Мы говорим, что  $f$  описывает *расходящуюся волну*, если существует некоторая константа  $s$ , такая что  $w = 0$  для  $|x| < t - s$ . Если мы ищем решение с фиксированной частотой, то соответствующие начальные условия Коши имеют вид  $\{w, ikw\}$ . Предположим, что  $w$  — решение приведенного волнового уравнения вне некоторой ограниченной области. Тогда условие  $\{w, ikw\}$  обеспечивает расходящееся решение тогда и только тогда, когда выполняются условия излучения Зоммерфельда.

Таким образом, решение за пределами области  $S$  описывается расходящимися сферическими волнами, излученными поверхностью  $S$ . Именно Гюйгенсу принадлежит идея, что распространение возмущений в волновой теории представляется суперпозицией *вторичных возмущений* от промежуточной поверхности, скажем от  $S$ . Однако, он не дал адекватного объяснения, почему не существует *обратной волны* (падающей волны), т. е. почему волны распространяются только от поверхности. Идея, что обратные волны могут уничтожать друг друга в силу разности их фаз, принадлежит Френелю. Френель думал, что если все источники находятся внутри поверхности  $S$ , то *вторичное излучение* (т. е. подинтегральное выражение в формуле Гельмгольца) в силу интерференции для каждого элемента поверхности будет создавать нулевой эффект в каждой внутренней точке. Приведенные соображения, принадлежащие по существу Гельмгольцу, были первым строго математическим подходом к проблеме. Они показывают, что волны уничтожаются внутри области благодаря суммарному эффекту от границы. И тем не менее, как мы увидим в главе 21, Френель был прав с точностью до членов порядка  $1/k$ . Это будет сделано с помощью асимптотической оценки интегралов в формуле Гельмгольца.

До сих пор мы работали с *монохроматическим излучением*  $u$ , соответствующим зависящей от времени функции  $v(x, y, z, t) = u(x, y, z)e^{-ikt}$ . Для фиксированной точки  $P$  введем функцию

$$v_P(x, y, z, t) = v(x, y, z, t - r),$$

где  $r$  — расстояние от  $P$  до  $(x, y, z)$ . Подстановка в формулу Гельмгольца дает

$$v(P, T) = \frac{1}{4\pi} \int_S \left( v_P \star d\left(\frac{1}{r}\right) - \frac{1}{r} \frac{\partial u}{\partial t_P} \star dr - \frac{1}{r} \star dv_P \right).$$

Это формула Кирхгофа. Поскольку она линейна по  $v$  не содержит в явном виде частоту, то она верна для любой суперпозиции монохроматических волн с различными частотами, а следовательно, для произвольного решения волнового уравнения. В этой форме связь с принципом Гюйгенса очевидна.

## Резюме

### А. Уравнения Максвелла

Вы должны знать законы Фарадея и Ампера в интегральной и дифференциальной форме и уметь доказывать, что 2-формы

$$F = B + E \wedge dt \quad \text{и} \quad G = D - H \wedge dt$$

удовлетворяют условию  $G = \star F$  в вакууме.

Вы должны уметь приводить уравнения Максвелла к виду  $\square A = 4\pi j$  и описывать решения этого волнового уравнения.

Вы должны знать, как формулируется закон сохранения энергии-импульса для электромагнитного поля через дифференциальные формы.

## Задачи

19.1. В четырехмерном пространстве-времени (используются единицы, когда  $c = 1$ )  $(dt, dt) = 1$ ,  $(dx, dx) = (dy, dy) = (dz, dz) = -1$  и  $\sigma = dt \wedge dx \wedge dy \wedge dz$ . 2-форма  $F$ , представляющая электромагнитное поле, может быть выражена через координаты  $x, y, z, t$  в виде  $F = B_z(x, y, t) dx \wedge dy - E_x(x, y, t) dt \wedge dx - E_y(x, y, t) dt \wedge dy$ . Все другие члены равны нулю.

(а) Какое соотношение должно выполняться для частных производных  $B_z$ ,  $E_x$  и  $E_y$ , чтобы выполнялось равенство  $dF = 0$ ?

(б) Выразите  $\star d\star F$  через частные производные  $B_z$ ,  $E_x$  и  $E_y$ .

19.2. Преобразование Лоренца  $\alpha$ , соответствующее скорости  $v$  вдоль оси  $x$  можно записать в терминах переноса (для удобства мы положили  $c = 1$ ) следующим образом:

$$\begin{aligned}\alpha^* t' &= \gamma(t - vx), \\ \alpha^* x' &= \gamma(x - vt), \\ \alpha^* y' &= y, \\ \alpha^* z' &= z,\end{aligned}$$

где  $\gamma = 1/\sqrt{1 - v^2}$ . Если  $x, y, z, t$  — координаты системы отсчета наблюдателя, стоящего на платформе, а  $x', y', z', t'$  — координаты системы отсчета наблюдателя, сидящего в поезде, который

движется со скоростью  $v$  вдоль оси  $x$ , то  $\alpha^*$  преобразует координаты события  $x', y', z', t'$ , наблюдаемое в поезде, в координаты того же события в системе отсчета наблюдателя на платформе.

- (а) Покажите явно, что  $\alpha^*$  сохраняет звездный оператор, т.е. что  $\alpha^* \star \omega = \star(\alpha^* \omega)$ . Используя звездный оператор для координат  $x, y, z, t$ , т.е.  $\star dt = dx \wedge dy \wedge dz$  и т.д., проверьте следующее:

$$\begin{aligned}\alpha^* \star dx' \wedge dy' \wedge dz' &= \star \alpha^* (dx' \wedge dy' \wedge dz'), \\ \alpha^* \star dt' \wedge dy' \wedge dz' &= \star \alpha^* (dt' \wedge dy' \wedge dz'), \\ \alpha^* \star dt' \wedge dx' &= \star \alpha^* (dt' \wedge dx'), \\ \alpha^* \star dx' \wedge dy' &= \star \alpha^* (dx' \wedge dy').\end{aligned}$$

- (б) Потенциал  $\mathbf{A}$  может быть записан через координаты пассажира поезда в форме

$$\mathbf{A} = A'_t dt' + A'_x dx' + A'_y dy' + A'_z dz'.$$

Вычислите  $\alpha^* \mathbf{A}$  и приравняйте полученный результат

$$A_t dt + A_x dx + A_y dy + A_z dz,$$

откуда получите выражения для  $A_t, A_x, \dots$  через  $A'_t, A'_x, \dots$

### 19.3. Выполните упражнение 19.2 для 2-формы

$$\begin{aligned}F &= B' + E' \wedge dt', \\ F &= B'_x dy' \wedge dz' + \dots + E'_z dz' \wedge dt'.$$

Вычислите  $\star F$  и сравните результат с

$$F = B_x dy \wedge dz + \dots + E_z dz \wedge dt.$$

Получите выражения для компонент поля  $B_x, B_y, B_z, E_x, E_y, E_z$  через  $B'_x, B'_y, B'_z, E'_x, E'_y, E'_z$ .

### 19.4. Точечный заряд покоится в начале координат в системе отсчета поезда. Для этого заряда, как известно,

$$F = B' + E' \wedge dt' = \frac{x' dx' + y' dy' + z' dz'}{(x'^2 + y'^2 + z'^2)^{3/2}} \wedge dt'.$$

Вычислите  $\alpha^* F$  и опишите поля заряда, совершающего равномерное движение вдоль оси  $x$ .

19.5. Из 2-формы  $F$  можно построить две различные 0-формы:

$$f_1 = \star(F \wedge F), \quad f_2 = \star(F \wedge \star F).$$

Выразите эти величины, инвариантные относительно преобразований Лоренца, через компоненты  $F$ , т. е. через  $B_x, B_y, \dots, E_z$ .

## Глава 20

---

# Комплексный анализ

---

В этой главе приводится достаточно стандартное изложение теории функций комплексного переменного.

### Введение

Настоящая глава посвящена изучению теории функций комплексного переменного. Эта теория, в ее различных модификациях, является главным достижением математики XIX века. Здесь мы кратко рассмотрим только ключевые результаты. Вспомним основные положения теории функций нескольких переменных, в частности, двух переменных (главы 7–8). Пусть  $x$  и  $y$  — координаты в пространстве  $\mathbb{R}^2$ . Основные объекты дифференциального исчисления — функции, линейные дифференциальные формы и формы второй степени. Каждый из них может быть определен только на открытом подмножестве пространства  $\mathbb{R}^2$ . Функция  $f$  приписывает число  $f(x, y)$  точке  $(x, y)$ . Если специально не оговорено, то все функции предполагаются дифференцируемыми. Линейная дифференциальная форма имеет вид

$$\omega = a dx + b dy,$$

где  $a$  и  $b$  — функции. Если  $\gamma$  — ориентированная кривая (кусочно-дифференцируемая и лежащая в области определения  $\omega$ ), то можно вычислить криволинейный интеграл  $\int_{\gamma} \omega$ . Если имеется суперпозиция (объединение)  $\gamma = \gamma_1 \dots \gamma_n$ , где  $\gamma_i$  — дифференцируемые кривые, параметризованные  $x_i(t)$ ,  $y_i(t)$  при  $\alpha_i \leq t \leq \beta_i$ ,

то

$$\int_{\gamma} \omega = \int_{\gamma_1} \omega + \dots + \int_{\gamma_n} \omega,$$

где

$$\int_{\gamma_i} \omega = \int_{\alpha_i}^{\beta_i} [a(x_i(t), y_i(t))x'_i(t) + b(x_i(t), y_i(t))y'_i(t)] dt.$$

Заданная функция  $f$  имеет дифференциал  $df$ , являющийся линейной дифференциальной формой вида

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy.$$

Теорема Стокса для функций (фундаментальная теорема дифференциального исчисления) говорит, что если  $\gamma$  — любая кривая, соединяющая точки  $p$  и  $q$ , то  $\int_{\gamma} df = f(q) - f(p)$ .

2-форма имеет вид

$$\Omega = c dx \wedge dy,$$

где  $c$  — функция. Если  $D$  — ограниченная область на плоскости (содержащаяся в области определения  $\Omega$ ), то можно сформировать интеграл  $\int_D \Omega$ . Если при этом  $D$  имеет стандартную ориентацию, то  $\int_D \Omega$  — двойной интеграл

$$\int_D \omega = \int_D c(x, y) dx dy.$$

(Если  $D$  имеет противоположную ориентацию, то необходимо изменить знак.) Предполагается, что все области имеют кусочно-дифференцируемую границу, которая обладает индуцированной ориентацией. Эта граница (с ориентацией) обозначается  $\partial D$ . Если  $\omega = a dx + b dy$ , то  $d\omega$  является 2-формой вида

$$d\omega = \left( \frac{\partial a}{\partial y} - \frac{\partial b}{\partial x} \right) dx \wedge dy,$$

и теорема Стокса (для 1-форм) гласит

$$\int_{\partial D} \omega = \int_D d\omega.$$

Мы предполагаем, что читатель знаком с основными свойствами комплексных чисел. Каждое комплексное число  $c$  можно записать в форме  $c = r + is$ , где  $r$  и  $s$  — вещественные числа,  $i^2 = -1$ . При этом  $r$  называется *вещественной частью* и  $s$  — *мнимой частью* комплексного числа  $c$ . Число  $\bar{c} = r - is$  называется *комплексно-сопряженным* от  $c$ , и  $c\bar{c} = |c|^2 = r^2 + s^2$ . Для комплексных чисел выполняются все законы алгебры, в частности, коммутативный и ассоциативный законы для операций сложения и умножения и дистрибутивный закон.

## 20.1. Комплекснозначные функции

Теперь мы рассматриваем функции и дифференциальные формы, которые могут иметь комплексные значения, т. е. *комплекснозначная функция*  $f$  приписывает комплексное число  $f(x, y) = u(x, y) + iv(x, y)$  каждой точке на плоскости, где  $u$  и  $v$  — вещественнозначные функции. Другими словами, задание комплекснозначной функции означает задание пары вещественнозначных функций, т. е. то же самое, что задание отображения (открытого подмножества)  $\mathbb{R}^2$  в  $\mathbb{R}^2$ . Мы говорим, что  $f$  дифференцируема, если дифференцируемо это отображение, т. е. дифференцируемы функции  $u$  и  $v$ . Таким образом, формулы  $\partial f / \partial x = \partial u / \partial x + i \partial v / \partial x$  и  $\partial f / \partial y = \partial u / \partial y + i \partial v / \partial y$  задают частные производные функции  $f$ . Можно рассматривать различные линейные комбинации этих частных производных с вещественными или комплексными коэффициентами. Например, рассмотрим выражение

$$\frac{1}{2} \frac{\partial f}{\partial x} + \frac{i}{2} \frac{\partial f}{\partial y}.$$

Можно считать, что это результат действия *дифференциального оператора*  $\frac{1}{2}(\partial/\partial x + i\partial/\partial y)$  на функцию  $f$ . В последующем изложении этот оператор чрезвычайно важен для нас, поэтому введем для него специальное обозначение  $\partial/\partial \bar{z}$ . Таким образом,

$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2} \left( \frac{\partial}{\partial x} + i \frac{\partial}{\partial y} \right). \quad (20.1)$$

Аналогично определим

$$\frac{\partial}{\partial z} = \frac{1}{2} \left( \frac{\partial}{\partial x} - i \frac{\partial}{\partial y} \right). \quad (20.2)$$

Несколько позже станет понятна причина для введения этих обозначений. Заметим, что если  $f = u + iv$ , то

$$\begin{aligned} \frac{\partial f}{\partial \bar{z}} &= \frac{\partial u}{\partial \bar{z}} + i \frac{\partial v}{\partial \bar{z}} \\ &= \frac{1}{2} \left( \frac{\partial u}{\partial x} + i \frac{\partial u}{\partial y} \right) + \frac{i}{2} \left( \frac{\partial v}{\partial x} + i \frac{\partial v}{\partial y} \right), \end{aligned}$$

так что, собирая вместе вещественную и мнимую части, мы получим

$$\frac{\partial f}{\partial \bar{z}} = \frac{1}{2} \left( \frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right) + \frac{i}{2} \left( \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \right). \quad (20.3)$$

Заметим, что для любой пары дифференцируемых функций  $f$  и  $g$  мы имеем

$$\frac{\partial}{\partial \bar{z}}(fg) = \frac{\partial f}{\partial \bar{z}} g + f \frac{\partial g}{\partial \bar{z}} \quad (20.4)$$

и аналогичное равенство справедливо для  $\partial/\partial z$ .

Рассмотрим комплекснозначную функцию  $z(x, y) = x + iy$ . Запишем это в виде

$$z = x + iy \quad (20.5)$$

и аналогично

$$\bar{z} = x - iy.$$

Тогда функция  $z^2$  задается формулой

$$z^2 = (x^2 - y^2) + i2xy, \quad \text{а} \quad \bar{z}^2 = (x^2 - y^2) - i2xy$$

и

$$z\bar{z} = |z|^2 = x^2 + y^2.$$

Аналогично можно сформировать полином от  $z$ , например,  $17z^3 - 5z^2 + 2z - 3$ , полином от  $\bar{z}$ , или смешанный полином, например,

$3z^2\bar{z} - 2z\bar{z}^3 + z^2 - 5\bar{z} + 1$  и т. д. Все это — примеры комплекснозначных функций. Если в формулу (20.3) вместо  $f$  подставить  $z$ , где  $u(x, y) \equiv x$  и  $v(x, y) \equiv y$ , то получим

$$\frac{\partial z}{\partial \bar{z}} \equiv 0 \quad (20.6)$$

и аналогично

$$\frac{\partial z}{\partial z} \equiv 1.$$

Из формулы (20.4) и линейности  $\partial/\partial z$  мы заключаем, что если  $P = P(z)$  — произвольный полином только от  $z$ , то

$$\frac{\partial P}{\partial \bar{z}} \equiv 0.$$

Если  $P(z) = a_0 + a_1z + \dots + a_nz^n$ , то

$$\frac{\partial P}{\partial z} = a_1 + 2a_2z + \dots + na_nz^{n-1}.$$

Это полином  $P'(z)$  — обычная производная полинома  $P$ . Таким образом,

$$\frac{\partial P}{\partial \bar{z}} = 0 \quad \text{и} \quad \frac{\partial P}{\partial z} = P'.$$

## 20.2. Комплекснозначные дифференциальные формы

Рассмотрим дифференциальные формы, имеющие комплексные значения. Линейная дифференциальная форма имеет вид

$$a dx + b dy,$$

где  $a$  и  $b$  — комплекснозначные функции. Сохраняются определение криволинейного интеграла и одномерный вариант теоремы Стокса, только теперь значениями будут комплексные числа. Определим линейные дифференциальные формы  $dz$  и  $d\bar{z}$

$$dz = dx + i dy, \quad \text{и} \quad d\bar{z} = dx - i dy.$$

Отметим, что эти дифференциальные формы линейно независимы в каждой точке и, кроме того,

$$dx = \frac{1}{2}(dz + d\bar{z}), \quad dy = \frac{1}{2i}(dz - d\bar{z}).$$

Таким образом, произвольная линейная дифференциальная форма может быть записана в виде

$$a dx + b dy = A dz + B d\bar{z},$$

где

$$A = \frac{1}{2}(a - ib) \quad \text{и} \quad B = \frac{1}{2}(a + ib).$$

В частности, для любой дифференцируемой комплекснозначной функции  $f$  мы имеем

$$\begin{aligned} df &= \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy \\ &= \frac{1}{2} \left( \frac{\partial f}{\partial x} - i \frac{\partial f}{\partial y} \right) dz + \frac{1}{2} \left( \frac{\partial f}{\partial x} + i \frac{\partial f}{\partial y} \right) d\bar{z}. \end{aligned}$$

Используя определения (20.1) и (20.2), мы можем это переписать в виде

$$df = \frac{\partial f}{\partial z} dz + \frac{\partial f}{\partial \bar{z}} d\bar{z}. \quad (20.7)$$

Давайте рассмотрим несколько примеров. Если взять  $f = z$ , то из (20.5) и (20.6) имеем  $d(z) = dz$ . Таким образом, наши обозначения логичны (именно поэтому мы ввели символы  $\partial/\partial z$  и  $\partial/\partial \bar{z}$ ). Если  $f = P$  — полином от  $z$ , то

$$dP = P' dz.$$

В этом выражении отсутствует член с  $d\bar{z}$ . Возьмем  $f = |z|^2 = z\bar{z}$ , тогда получим

$$d|z|^2 = \bar{z} dz + z d\bar{z}.$$

Определим комплексную экспоненциальную функцию  $e^z$  формулой

$$e^z = e^x(\cos y + i \sin y).$$

Тогда

$$\begin{aligned} de^z &= e^x(\cos y + i \sin y) dx + e^x(-\sin y + i \cos y) dy \\ &= e^x(\cos y + i \sin y)(dx + i dy), \end{aligned}$$

и, следовательно,

$$de^z = e^z dz.$$

Из этого выражения получаем

$$\frac{\partial e^z}{\partial \bar{z}} \equiv 0.$$

Комплекснозначная 2-форма имеет вид

$$c dx \wedge dy,$$

где  $c$  — комплекснозначная функция. Из двух линейных комплекснозначных дифференциальных форм  $\omega_1$  и  $\omega_2$  мы можем образовать внешнее произведение  $\omega_1 \wedge \omega_2$ . Тогда, например,

$$\begin{aligned} dz \wedge d\bar{z} &= (dx + i dy) \wedge (dx - i dy) \\ &= -2i dx \wedge dy \end{aligned}$$

или

$$dx \wedge dy = \frac{i}{2} dz \wedge d\bar{z}.$$

В частности,

$$d(A dz) \equiv 0, \quad \text{если и только если} \quad \frac{\partial A}{\partial \bar{z}} \equiv 0.$$

Теорема Стокса сохраняется без изменения, только теперь  $\omega$  — линейная комплекснозначная дифференциальная форма. Например, если  $A dz = 0$  удовлетворяет равенству  $d(A dz) = 0$ , то теорема Стокса говорит, что для любой области  $D$ , содержащейся в области определения  $A$ , мы имеем

$$\boxed{\int_{\partial D} A dz = 0.}$$

Этот важный результат называется *теоремой Коши*.

### 20.3. Голоморфные функции

Дифференцируемая функция  $f$ , определенная в некоторой области  $D$  пространства  $\mathbb{R}^2$ , называется *голоморфной*<sup>1</sup>, если она удовлетворяет тождеству

$$\frac{\partial f}{\partial \bar{z}} \equiv 0. \quad (20.8)$$

Так как это определение является основным для всей темы, то этот параграф посвящен обсуждению эквивалентных формулировок и простейшим следствиям понятия голоморфности. Поскольку  $\partial/\partial \bar{z}$  — дифференциальный оператор, то мы имеем уравнения

$$\begin{aligned} \frac{\partial}{\partial \bar{z}}(f + g) &= \frac{\partial f}{\partial \bar{z}} + \frac{\partial g}{\partial \bar{z}}, & \frac{\partial}{\partial \bar{z}}(fg) &= \frac{\partial f}{\partial \bar{z}} g + f \frac{\partial g}{\partial \bar{z}}, \\ \frac{\partial}{\partial \bar{z}} \left( \frac{f}{g} \right) &= \frac{1}{g^2} \left( \frac{\partial f}{\partial \bar{z}} g - f \frac{\partial g}{\partial \bar{z}} \right), & g &\neq 0. \end{aligned}$$

Из этих соотношений мы видим, что сумма, произведение и частное двух голоморфных функций тоже голоморфны (в случае частного знаменатель нигде не должен обращаться в нуль). В предыдущем параграфе мы уже пользовались этим замечанием, чтобы получить тождество  $(\partial/\partial \bar{z})P(z) \equiv 0$ , т. е. что любой полином (только от  $z$ ) — голоморфная функция. Аналогично, любая рациональная функция, т. е. отношение двух полиномов  $P/Q$ , — голоморфная функция в области  $D$  при условии, что  $Q$  нигде не обращается в нуль в этой области.

С учетом (20.7), функция  $f$  голоморфна, если и только если

$$df = h dz, \quad (20.9)$$

т. е. если и только если  $df$  не имеет членов с  $d\bar{z}$ . Здесь, конечно,  $h = df/dz$ . Это условие можно записать в эквивалентной форме

$$d(f dz) = 0. \quad (20.10)$$

<sup>1</sup>а также *аналитической*. При этом предполагается непрерывность частных производных первого порядка, что иногда в дальнейшем не уточняется, но используется. — *Прим. ред.*

Пусть  $f = u + iv$ , где вещественнозначные функции  $u$  и  $v$  являются вещественной и мнимой частью  $f$ . Теперь, полагая вещественную и мнимую части  $\partial f / \partial \bar{z}$  равными нулю и используя (20.3), мы видим, что

$$\boxed{\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.} \quad (20.11)$$

Это система дифференциальных уравнений в частных производных первого порядка, называемых *уравнениями Коши–Римана*. Таким образом, функция  $f$  голоморфна, если и только если выполняются уравнения Коши–Римана.

Существует другой способ записи уравнений Коши–Римана. Рассмотрим дифференциальную форму

$$u \, dx + v \, dy.$$

Тогда в евклидовой метрике на плоскости

$$\star(u \, dx + v \, dy) = -v \, dx + u \, dy.$$

Тем самым уравнения Коши–Римана эквивалентны равенствам

$$d(u \, dx + v \, dy) = 0 \quad \text{и} \quad d\star(u \, dx + v \, dy) = 0.$$

Заметим, что они аналогичны уравнениям Максвелла в пространстве  $\mathbb{R}^{1,3}$ . В уравнениях Максвелла мы имели дело с 2-формами в четырехмерном пространстве. Уравнения Коши–Римана написаны для 1-формы в двумерном пространстве. Как уже было отмечено в главе 18, звездный оператор  $\star : \Lambda^{n/2} \rightarrow \Lambda^{n/2}$  ( $n$  — четное число) зависит только от конформной структуры.

Будем полагать, что  $f$  — отображение (подпространства)  $\mathbb{R}^2$  в  $\mathbb{R}^2$ , переводящее  $(x, y)$  в  $(u(x, y), v(x, y))$ . Матрица Якоби этого отображения равна

$$\begin{pmatrix} \partial u / \partial x & \partial u / \partial y \\ \partial v / \partial x & \partial v / \partial y \end{pmatrix}.$$

Из формул (20.11) следует, что эта матрица имеет вид  $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ .

В пространстве  $\mathbb{R}^2$  матрица имеет такой вид, где  $a^2 + b^2 \neq 0$ , если

и только если она конформна (т.е. сохраняет углы) и сохраняет ориентацию. Тогда (20.11) можно переписать так:

$$\begin{pmatrix} \partial u/\partial x & \partial u/\partial y \\ \partial v/\partial x & \partial v/\partial y \end{pmatrix} \quad \text{или равна нулю, или конформна}^2 \quad (20.12) \\ \text{и сохраняет ориентацию.}$$

Пусть  $f = u + iv$  и  $g = r + is$  — комплексные функции. Будем рассматривать  $f$  и  $g$  как отображения (подпространства)  $\mathbb{R}^2$  в  $\mathbb{R}^2$ . Если образ  $g$  лежит в области определения  $f$ , то можно построить композицию  $f \circ g$ . В силу цепного правила, матрица Якоби  $f \circ g$  (рассматриваемая как отображение  $\mathbb{R}^2$  в  $\mathbb{R}^2$ ), вычисленная в  $(x, y)$ , является произведением матрицы Якоби функции  $f$ , вычисленной в точке  $(r(x, y), s(x, y))$ , на матрицу Якоби функции  $g$ , вычисленной в точке  $(x, y)$ . Произведение двух конформных сохраняющих ориентацию матриц тоже конформно и сохраняет ориентацию. Кроме того, произведение нулевой матрицы на любую матрицу равно нулю. Из (20.12) следует, что композиция двух голоморфных функций тоже голоморфна. Для композиции двух голоморфных функций мы используем стандартное выражение: если, например,  $f(z) = e^z$  и  $g(z) = 3z^2 - 2$ , то  $f \circ g = e^{3z^2 - 2}$ . Поскольку  $\partial/\partial z$  — линейная комбинация операторов  $\partial/\partial x$  и  $\partial/\partial y$  (с постоянными коэффициентами), из цепного правила следует

$$\frac{\partial}{\partial z}(f \circ g) = \left( \frac{\partial f}{\partial z} \circ g \right) \frac{\partial g}{\partial z}. \quad (20.13)$$

Поэтому для примера, приведенного выше, мы имеем

$$\frac{\partial}{\partial z} e^{3z^2 - 2} = 6ze^{3z^2 - 2}.$$

Другими словами, для голоморфных функций действие оператора  $\partial/\partial z$  можно вычислять как обычную производную функции одной переменной. Причина вскоре станет ясна.

Условие (20.12) не слишком удобно, потому что в нем есть два разных случая: матрица Якоби равна нулю или не равна нулю. Давайте запишем условие (20.12) несколько по-другому. Рассмотрим матрицу

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

---

<sup>2</sup>По этой причине голоморфная функция, для которой  $\partial f/\partial z \neq 0$  всюду в рассматриваемой области, задает *конформное отображение*. — *Прим. ред.*

Матрица

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

коммутирует с  $J$ , т. е. выполняется равенство

$$AJ = JA,$$

если и только если  $a = d$  и  $c = -b$ , как легко увидеть, перемножая матрицы. Таким образом,  $AJ = JA$ , если и только если матрица  $A$  имеет вид

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

где  $a$  и  $b$  могут быть равны нулю. Тогда условие (20.12) можно записать в виде

$$\begin{pmatrix} \partial u / \partial x & \partial u / \partial y \\ \partial v / \partial x & \partial v / \partial y \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \partial u / \partial x & \partial u / \partial y \\ \partial v / \partial x & \partial v / \partial y \end{pmatrix} \quad (20.14)$$

Мы можем дать интересную интерпретацию равенству  $AJ = JA$  и тем самым условию (20.14). Матрица  $A$  имеет размер  $2 \times 2$ . Она задает линейное преобразование вещественного двумерного векторного пространства  $\mathbb{R}^2$  в себя. Пространство  $\mathbb{R}^2$  можно идентифицировать с одномерным комплексным векторным пространством  $\mathbb{C}$ , идентифицируя вектор  $\begin{pmatrix} x \\ y \end{pmatrix}$  с комплексным числом  $x + iy$ . Умножение этого числа на  $i$  переводит его в  $-y + ix$ . Но комплексное число  $-y + ix$  соответствует вектору  $\begin{pmatrix} -y \\ x \end{pmatrix} = J \begin{pmatrix} x \\ y \end{pmatrix}$ .

Таким образом,  $J$  — матрица умножения на  $i$  с вещественной точки зрения<sup>3</sup>. Произвольное линейное преобразование одномерного комплексного векторного пространства  $\mathbb{C}$  — это линейное преобразование пространства  $\mathbb{R}^2$ . Но не каждое линейное преобразование  $A$  пространства  $\mathbb{R}^2$  над вещественными числами соответствует линейному преобразованию пространства  $\mathbb{C}$  над комплексными числами. Комплексное линейное преобразование должно

<sup>3</sup>В обычной евклидовой метрике  $\mathbb{R}^2$  это поворот на угол  $90^\circ$  (в направлении от  $e_x$  к  $e_y$ ). Умножение на число  $c = a + ib$  представляет собой композицию гомотетии с коэффициентом  $r = |c| = \sqrt{a^2 + b^2}$  и поворота на угол  $\theta$  (аргумент  $c$ ), удовлетворяющий условию  $c = r(\cos \theta + i \sin \theta)$ , т. е. конформное линейное преобразование. — *Прим. ред.*

коммутировать с умножением на комплексные числа, в частности, на  $i$ . Таким образом, чтобы матрица  $A$  соответствовала комплексному линейному преобразованию, должно выполняться равенство  $AJ = JA$ . Если  $A$  удовлетворяет этому условию, т.е.

$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$ , то легко проверить, что  $A$  соответствует умно-

жению на комплексное число  $a + ib$ , т.е.  $A \begin{pmatrix} x \\ y \end{pmatrix}$  соответствует  $(a + ib)(x + iy)$ . Очевидно, что умножение на комплексное число в пространстве  $\mathbb{C}$  комплексно линейно. Таким образом, (20.14) можно переписать в форме

$$\text{Преобразование } \begin{pmatrix} \partial u / \partial x & \partial u / \partial y \\ \partial v / \partial x & \partial v / \partial y \end{pmatrix} \text{ комплексно линейно.} \quad (20.15)$$

Ни одна из приведенных эквивалентных формулировок, за исключением уравнений Коши–Римана (20.11), не была известна математикам XIX века. А теперь дадим классическую формулировку, наиболее близкую к (20.15). Пусть  $f$  — дифференцируемая функция, не обязательно голоморфная, а  $\begin{pmatrix} x \\ y \end{pmatrix}$  — точка из обла-

сти определения  $f$ . Кроме того, пусть  $\begin{pmatrix} k \\ l \end{pmatrix}$  — малый вектор. Тогда из определения дифференцируемости мы имеем

$$f\left(\begin{pmatrix} x+k \\ y+l \end{pmatrix}\right) - f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \frac{\partial f}{\partial x} k + \frac{\partial f}{\partial y} l + o((\|k^2 + l^2\|)^{1/2}).$$

Здесь частные производные вычисляются в точке  $\begin{pmatrix} x \\ y \end{pmatrix}$ .

Пусть  $z = x + iy$  и  $h = k + il$ . Будем писать  $f(z)$  вместо  $f\left(\begin{pmatrix} x \\ y \end{pmatrix}\right)$ . Тогда предыдущее уравнение можно переписать в форме

$$f(z+h) - f(z) = \frac{\partial f}{\partial x} k + \frac{\partial f}{\partial y} l + o(|h|).$$

Используем (20.7) и перепишем еще раз:

$$f(z+h) - f(z) = \frac{\partial f}{\partial z} h + \frac{\partial f}{\partial \bar{z}} \bar{h} + o(|h|).$$

Это уравнение содержит комплексные числа. Поэтому если  $h \neq 0$ , то обе части делим на  $h$  и получаем

$$\frac{f(z+h) - f(z)}{h} = \frac{\partial f}{\partial z} + \frac{\partial f}{\partial \bar{z}} \frac{\bar{h}}{h} + o(1).$$

Заметим, что левая часть этого равенства выглядит как обычное разностное отношение, но только здесь комплексные числа. Предположим, что  $|h| \rightarrow 0$ . Величина  $\bar{h}/h$  может быть любым комплексным числом, модуль которого равен единице. (Например, если  $h$  — вещественное число, то  $\bar{h}/h = 1$ , а если  $h = is$ , т.е. чисто мнимое, то  $\bar{h}/h = -1$ .) Поэтому предельное выражение справа может зависеть от угла, при котором  $h \rightarrow 0$ . Если мы хотим, чтобы этот предел существовал, т.е. не зависел от угла, надо потребовать, чтобы  $\partial f/\partial \bar{z} = 0$  в точке  $\begin{pmatrix} x \\ y \end{pmatrix}$ . Таким образом, если  $f$  — голоморфная функция, то предел левой части равенства существует во всех точках области определения  $f$ , и этот предел равен  $\partial f/\partial z$  и является непрерывной функцией. Обратно, предположим, что в каждой точке существует предел

$$\lim_{h \rightarrow 0} \frac{f(z+h) - f(z)}{h} = f'(z)$$

(независимо от того, как  $h$  стремится к нулю) и является непрерывной функцией. Выбирая  $h$  вещественным числом, можно показать, что  $\partial f/\partial x$  существует и непрерывна, а выбирая  $h$  чисто мнимым, можно показать, что  $\partial f/\partial y$  существует и непрерывна. Тогда  $f$  — непрерывно дифференцируемая функция. Предыдущее рассуждение показывает, что  $\partial f/\partial \bar{z} = 0$  и  $f'(z) = \partial f/\partial z$ . Итак, мы доказали, что  $f$  голоморфна, если и только если существует предел

$$\lim_{h \rightarrow 0} \frac{f(z+h) - f(z)}{h} = f'(z) \quad (20.16)$$

независимо от того, как  $h$  стремится к нулю, и этот предел — непрерывная функция. Короче условие (20.16) можно сформулировать так: функция  $f$  непрерывно дифференцируема с комплексной точки зрения. Именно так рассуждали математики в

XIX веке. Такой подход довольно обманчив, потому что условие комплексной дифференцируемости имеет гораздо больше ограничений, чем обычное условие дифференцируемости. Тем не менее, мы принимаем обозначение  $f'(z)$ , как более удобное, чем  $\partial f / \partial z$ .

Как уже было отмечено, из теоремы Стокса и (20.10) следует

$$\int_{\partial D} f(z) dz = 0, \quad (20.17)$$

что является интегральной теоремой Коши для любой области  $D$ , лежащей в области определения голоморфной функции  $f$ . И обратно, если  $f$  — произвольная функция, то

$$\int_{\partial D} f(z) dz = \int_D d(f dz).$$

Этот интеграл обращается в нуль для всех  $D$ , если подынтегральное выражение тождественно равно нулю, т. е. если  $f$  — голоморфная функция.

Таким образом, мы показали, что все условия (20.8)–(20.17) эквивалентны. Функция, удовлетворяющая любому из них (а тогда и всем им), *голоморфна*.

## 20.4. Вычисление вычетов

Начнем с вычисления контурного интеграла  $\int_{\gamma} z^n dz$ , где  $\gamma$  — окружность с центром в начале координат. Если  $n \geq 0$ , то функция  $z^n$  голоморфна на всей плоскости, тогда из теоремы Коши следует, что этот интеграл равен нулю. Для отрицательных значений  $n$  функция  $z^n$  не голоморфна (и не определена) в начале координат, и следовательно, мы должны вычислять интеграл прямым способом. Если окружность  $\gamma$  имеет радиус  $r$ , то в полярных координатах можно написать  $z = re^{i\theta}$ . Тогда вдоль окружности  $\gamma$  имеем  $dz = ire^{i\theta} d\theta$ . Итак,

$$z^n dz = ir^{n+1} e^{i(n+1)\theta} d\theta$$

и

$$\int_{\gamma} z^n dz = ir^{n+1} \int_0^{2\pi} e^{i(n+1)\theta} d\theta.$$

Этот последний интеграл равен нулю всегда, кроме случая  $n = -1$ , когда он равен  $2\pi$ . В этом случае  $r^{n+1} = 1$  и поэтому

$$\int_{\gamma} z^n dz = \begin{cases} 0, & \text{если } n \neq -1, \\ 2\pi i, & \text{если } n = -1. \end{cases}$$

Пусть  $D$  — некоторая область, содержащая начало координат (рис. 20.1), и пусть  $f$  — функция, голоморфная в области  $D - \{0\}$ . Далее предположим, что вблизи начала координат можно написать

$$f = a_{-n}z^{-n} + a_{-n+1}z^{-n+1} + \dots + a_{-1}z^{-1} + g,$$

где  $g$  — голоморфная функция вблизи начала координат.

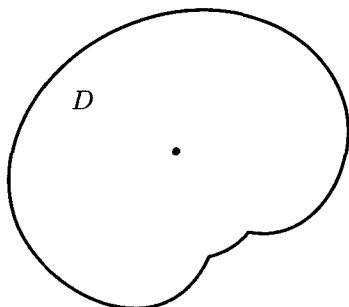


Рис. 20.1

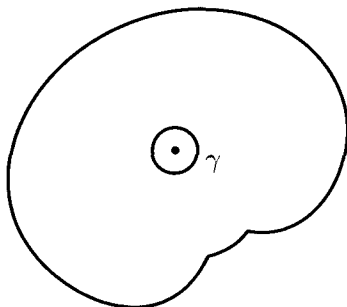


Рис. 20.2

Давайте вычислим  $\int_{\partial D} f(z) dz$ . Нарисуем окружность  $\gamma$  небольшого радиуса с центром в начале координат, целиком лежащую в  $D$  и в области, где работает выписанное разложение (рис. 20.2). Тогда  $f$  голоморфна в той части  $D$ , которая является внешней по отношению к  $\gamma$ . Поэтому, в силу интегральной теоремы Коши,

$$\int_{\partial D} f dz = \int_{\gamma} f dz.$$

Если сюда подставить разложение  $f$  и проинтегрировать каждый член, то все отрицательные степени дадут нуль, кроме  $a_{-1}z^{-1}$ . При этом  $\int_{\gamma} g dz = 0$ , поскольку  $g$  — голоморфная функция. Тогда

$$\int_{\partial D} f(z) dz = 2\pi i a_{-1}.$$

Поскольку начало координат было выбрано произвольно, то его можно заменить любой другой точкой или конечным числом точек в области  $D$ . Говорят, что функция  $f$  имеет *полюс*<sup>4</sup> в точке  $\alpha$ , если она голоморфна в области  $U - \{\alpha\}$ , где  $U$  — некоторая окрестность точки  $\alpha$ , и если в ее окрестности функция  $f$  имеет разложение

$$f(z) = a_{-n}(z - \alpha)^{-n} + \dots + a_{-1}(z - \alpha)^{-1} + g,$$

где  $g$  — голоморфная функция вблизи  $\alpha$ . Число  $a_{-1}$  называется *вычетом* функции  $f$  в точке  $\alpha$  и обозначается  $\text{res}_\alpha(f)$ . Из предыдущих рассуждений следует, что интеграл от  $f(z) dz$  по малой окружности вокруг  $\alpha$  равен  $2\pi i a_{-1} = 2\pi i \text{res}_\alpha(f)$ . Функция, голоморфная в области  $D$ , за исключением полюсов, называется *мероморфной*.

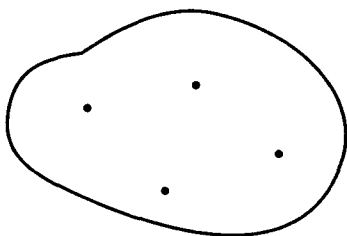


Рис. 20.3

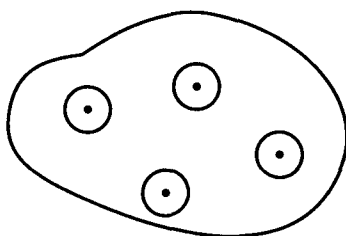


Рис. 20.4

Предположим, что  $f$  мероморфна в области  $D$  и имеет в ней конечное число полюсов (рис. 20.3). Тогда вокруг каждого полюса можно нарисовать маленькую окружность (рис. 20.4) и, применив интегральную теорему Коши, написать

$$\int_{\partial D} f(z) dz = 2\pi i \sum_{\text{по полюсам } \alpha} \text{res}_\alpha(f).$$

Эта формула называется *теоремой Коши о вычетах*. Она очень полезна при вычислении определенных интегралов. Например,

<sup>4</sup>В предположении, что старший коэффициент разложения  $a_{-n} \neq 0$ , число  $n$  называется порядком полюса. — Прим. ред.

мы хотим вычислить определенный интеграл

$$\int_0^\pi \frac{d\theta}{a + \cos \theta},$$

где  $a > 1$ . Запишем его в форме

$$\frac{1}{2} \int_0^{2\pi} \frac{d\theta}{a + \cos \theta}.$$

Теперь положим  $z = e^{i\theta}$ , тогда на единичной окружности  $dz/iz = d\theta$ . Кроме того,  $2 \cos \theta = e^{i\theta} + e^{-i\theta} = z + z^{-1}$ . Тогда наш интеграл принимает вид

$$\frac{1}{2} \int_\gamma \frac{dz}{iz(a + \frac{1}{2}(z + z^{-1}))} = \frac{1}{i} \int_\gamma \frac{dz}{(z^2 + 2az + 1)},$$

где  $\gamma$  — единичная окружность. В то же время

$$z^2 + 2az + 1 = (z - \alpha_1)(z - \alpha_2),$$

где

$$\alpha_1 = -a + \sqrt{a^2 - 1} \quad \text{и} \quad \alpha_2 = -a - \sqrt{a^2 - 1},$$

откуда мы получаем

$$\frac{1}{z^2 + 2az + 1} = \frac{1}{\alpha_1 - \alpha_2} \left( \frac{1}{z - \alpha_1} - \frac{1}{z - \alpha_2} \right).$$

Поскольку  $a > 1$ , имеем  $\alpha_2 < -1$ , и поэтому  $\alpha_2$  лежит вне единичной окружности. А так как  $\alpha_1 \alpha_2 = 1$ , то  $\alpha_1$  лежит внутри единичной окружности. Кроме того, Функция  $1/(z - \alpha_2)$  голоморфна внутри этой окружности, поэтому наша формула показывает, что вычет равен

$$\frac{1}{\alpha_1 - \alpha_2} = \frac{1}{2\sqrt{a^2 - 1}}.$$

Таким образом, согласно теореме Коши о вычетах, получаем ответ

$$\int_0^\pi \frac{d\theta}{a + \cos \theta} = \frac{1}{2i} \int_\gamma \frac{dz}{z^2 + 2az + 1} = \frac{\pi}{\sqrt{a^2 - 1}}.$$

Мы вскоре увидим, что любая рациональная функция вида  $R(z) = P(z)/Q(z)$  мероморфна и имеет полюса в точках, где  $Q$  обращается в нуль. (Это следует немедленно из разложения рациональной функции на дроби. Но здесь мы это не предполагаем.) Кроме того, мы получим эффективный способ вычисления вычетов. Тогда это позволит нам вычислять, например, интегралы вида  $\int_0^{2\pi} R(\cos \theta, \sin \theta) d\theta$ , сделав подстановку  $z = e^{i\theta}$  и учтя, что  $\cos \theta = (1/2)(z + z^{-1})$  и  $\sin \theta = (1/2i)(z - z^{-1})$ , после чего проделывая процедуру, описанную выше.

Начнем с леммы. Предположим, что  $g$  голоморфна в области  $\overline{D}$ <sup>5</sup>, кроме конечного числа точек  $\alpha_1, \dots, \alpha_k$ . Предположим, что в каждой из этих точек функция  $g$  удовлетворяет условию

$$\lim_{z \rightarrow \alpha} |z - \alpha| |g(z)| = 0.$$

Тогда

$$\int_{\partial D} g dz = 0.$$

Для доказательства этого утверждения нарисуем маленькие окружности  $\gamma_i$  вокруг каждой точки  $\alpha_i$ . Тогда, как и раньше,  $\int_{\partial D} g dz = \sum \int_{\gamma_i} g dz$ . При этом

$$\left| \int_{\gamma_i} g(z) dz \right| \leq 2\pi r M_r,$$

где  $r$  — радиус окружности  $\gamma_i$ , и  $M_r$  — максимум  $|g|$  на  $\gamma_i$ . В силу предположения  $r M_r \rightarrow 0$  при  $r \rightarrow 0$ . Таким образом, устремляя к нулю радиусы окружностей, мы получим  $\int_{\partial D} g dz = 0$ .

А теперь пусть  $f$  голоморфна в области  $\overline{D}$  и  $a$  — точка в области  $D$ . Применим доказанную лемму к функции

$$g(z) = \frac{f(z) - f(a)}{z - a}.$$

---

<sup>5</sup> $\overline{D}$  — замыкание области  $D$ : она содержит все точки области  $D$  плюс точки на границе. [Голоморфность функции на замкнутом множестве подразумевает выполнение соответствующих условий в некоторой открытой окрестности этого множества. — *Ред.*]

Функция  $g$  голоморфна в  $\bar{D}$  везде, кроме точки  $a$ . Более того,  $(z - a)g(z) = f(z) - f(a)$  и это выражение стремится к нулю при  $z \rightarrow a$ . Тогда

$$\int_{\partial D} \frac{f(z) - f(a)}{z - a} dz = 0.$$

Но  $f(a)$  — постоянная и

$$\int_{\partial D} \frac{dz}{z - a} = 2\pi i.$$

Откуда получаем

$$\int_{\partial D} \frac{f(a) dz}{z - a} = 2\pi i f(a),$$

и теперь можно написать

$$f(a) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(z)}{z - a} dz.$$

Для удобства заменим  $a$  на  $z$  и  $z$  на  $\xi$  и напомним *интегральную формулу Коши*

$$f(z) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(\xi)}{\xi - z} d\xi. \quad (20.18)$$

Эта формула является краеугольным камнем, на котором строится вся теория комплексного анализа.

Сформулируем следующую лемму. Пусть  $\phi$  — комплекснозначная функция, определенная и непрерывная на границе  $\partial D$  некоторой области  $D$ . Тогда функция

$$F_n(z) = \int_{\partial D} \frac{\phi(\xi)}{(\xi - z)^n} d\xi$$

голоморфна в области  $D$  и обладает свойством

$$F'_n(z) = nF_{n+1}(z).$$

Докажем эту лемму по индукции. Прежде всего докажем, что  $F_1$  непрерывна в  $D$ . Пусть  $z_0$  — некоторая точка в области  $D$ . Выберем  $a$  настолько малым, что окрестности  $|z'' - z_0| < 2a$  и

$|z' - z_0| < a$  целиком находятся в области  $D$ . Таким образом, расстояние от  $z'$  до всех точек границы  $\partial D$  не меньше  $a$  и, следовательно, при  $\xi \in \partial D$  неравенство

$$\frac{1}{|\xi - z'| |\xi - z_0|} < \frac{1}{2a^2} < \frac{1}{a^2}$$

выполняется для всех таких  $z$ . Теперь убираем штрих и применяем равенство

$$\frac{1}{\xi - z} - \frac{1}{\xi - z_0} = \frac{1}{(\xi - z)(\xi - z_0)}(z - z_0),$$

так что

$$\begin{aligned} |F_1(z) - F_1(z_0)| &= |z - z_0| \left| \int_{\partial D} \frac{\phi(\xi) d\xi}{(\xi - z)(\xi - z_0)} \right| \\ &\leq |z - z_0| \frac{1}{a^2} \int_{\partial D} |\phi(\xi)| d\xi. \end{aligned}$$

Это означает, что  $F_1$  — непрерывная функция. Более того,

$$\frac{F_1(z) - F_1(z_0)}{z - z_0} = \int_{\partial D} \frac{\phi(\xi) d\xi}{(\xi - z)(\xi - z_0)}.$$

Если фиксировать значение  $z_0$ , то функция  $\phi(\xi)/(\xi - z_0)$  непрерывна на границе  $\partial D$ . Следовательно, правая часть этого равенства — непрерывная функция от  $z$ , и оно выполняется для всех  $z \neq z_0$ . Устремляя  $z \rightarrow z_0$ , мы видим, что  $F_1$  — дифференцируемая функция в комплексном смысле и

$$F_1'(z_0) = \int \frac{\phi(\xi) d\xi}{(\xi - z_0)^2} = F_2(z_0).$$

Кроме того,

$$\frac{1}{(\xi - z)(\xi - z_0)} + \frac{(z - z_0)}{(\xi - z)^2(\xi - z_0)} = \frac{1}{(\xi - z)^2},$$

поэтому

$$\begin{aligned} F_2(z) - F_2(z_0) &= \int \frac{\phi(\xi) d\xi}{(\xi - z)(\xi - z_0)} - \int \frac{\phi(\xi) d\xi}{(\xi - z_0)^2} + (z - z_0) \int \frac{\phi(\xi) d\xi}{(\xi - z)^2(\xi - z_0)}, \end{aligned}$$

что показывает (с учетом предыдущих соображений) непрерывность функции  $F_2$  и голоморфность функции  $F_1$ . Далее,

$$\frac{F_2(z) - F_2(z_0)}{z - z_0} = \frac{R_1(z) - R_1(z_0)}{z - z_0} + \int \frac{\phi(\xi) d\xi}{(\xi - z)^2(\xi - z_0)},$$

где

$$R_1(z) = \int \frac{\phi(\xi) d\xi}{(\xi - z)(\xi - z_0)}.$$

Поскольку функция  $\phi(\xi)/(\xi - z_0)$  непрерывна на границе  $\partial D$ , благодаря предыдущему результату мы приходим к выводу, что  $R_1$  дифференцируема и  $R_1'(z_0) = \int (\phi(\xi)/(\xi - z_0)^3) d\xi$ . Переходя к пределу в нашем уравнении, получаем, что  $F_2$  — дифференцируемая функция в комплексном смысле и

$$F_2' = 2F_3.$$

Теперь процедура индукции уже очевидна. Предположим, что мы доказали свойство  $F_{n-1}' = (n-1)F_n$  для всех непрерывных функций  $\phi$ . Используя тождество

$$\frac{1}{(\xi - z)^{n-1}(\xi - z_0)} + \frac{z - z_0}{(\xi - z)^n(\xi - z_0)} = \frac{1}{(\xi - z_0)^n},$$

мы видим, что

$$\begin{aligned} F_n(z) - F_n(z_0) &= \int \frac{\phi(\xi) d\xi}{(\xi - z)^{n-1}(\xi - z_0)} - \int \frac{\phi(\xi) d\xi}{(\xi - z_0)^n} + (z - z_0) \int \frac{\phi(\xi) d\xi}{(\xi - z)^n(\xi - z_0)} \\ &= R_{n-1}(z) - R_{n-1}(z_0) + (z - z_0) \int \frac{\phi(\xi) d\xi}{(\xi - z)^n(\xi - z_0)}, \end{aligned}$$

где функции

$$R_k(z) = \int \frac{\phi(\xi) d\xi}{(\xi - z)^k(\xi - z_0)}$$

связаны с непрерывной функцией  $\phi(\xi)/(\xi - z_0)$ . Итак, мы доказали, что  $F_n$  непрерывна.

В силу предположения индукции  $R_{n-1}'(z_0) = (n-1)R_n(z_0) = (n-1)F_{n+1}(z_0)$ , поэтому разделим уравнение для  $F_n(z) - F_n(z_0)$

на  $z - z_0$  и перейдем к пределу  $z \rightarrow z_0$ . В результате получим  $F'_n = nF_{n+1}$ , что и требовалось доказать.

Применим этот результат к интегральной формуле Коши (20.18). Поскольку функция  $f$  непрерывна, мы приходим к удивительному заключению: если функция  $f$  голоморфна в области  $\overline{D}$ , то она в ней бесконечно дифференцируема, причем каждая производная (комплексная) голоморфна и удовлетворяет условию

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_{\partial D} \frac{f(\xi) d\xi}{(\xi - z)^{n+1}}. \quad (20.19)$$

Из формул (20.18) и (20.19) можно получить много важных следствий, что мы и сделаем в следующих параграфах. Начнем с *принципа устранимых особенностей*. Предположим, что функция  $f$  определена и голоморфна в области  $D - \{a\}$ , где  $a \in D$ , т.е. во всей области  $D$ , кроме одной точки  $a$ . Необходимым и достаточным условием того, что существует голоморфная функция, определенная в области  $D$ , которая совпадает с  $f$  в области  $D - \{a\}$ , является  $\lim_{z \rightarrow a} (z - a)f(z) = 0$ . Такая продолженная функция (если она существует) определена однозначно.

*Доказательство.* Пусть  $f(a)$  обозначает значение продолженной функции в точке  $a$ . Поскольку продолженная функция голоморфна, она непрерывна в точке  $a$  и  $\lim_{z \rightarrow a} f(z) = f(a)$ . Поэтому  $\lim_{z \rightarrow a} (z - a)f(z) = \lim_{z \rightarrow a} (z - a)f(a) = 0$ , что доказывает необходимость. Чтобы доказать достаточность, нарисуем небольшую окружность  $\gamma$  с центром в точке  $a$ , целиком лежащую в области  $D$ . Рассмотрим функции

$$\frac{1}{2\pi i} \int \frac{f(\xi) d\xi}{\xi - z}.$$

Это функция голоморфна внутри  $\gamma$ . Нарисуем другую окружность  $\gamma_1$  вокруг  $a$ , лежащую внутри  $\gamma$ . Тогда во всех точках кольца между  $\gamma_1$  и  $\gamma$ , согласно интегральной формуле Коши (20.18), имеем

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\xi) d\xi}{\xi - z} - \frac{1}{2\pi i} \int_{\gamma_1} \frac{f(\xi) d\xi}{\xi - z}.$$

Согласно нашему предположению, если фиксировать  $z$  и радиус окружности  $\gamma_1$  устремить к нулю, то второй интеграл обратится

в нуль. Поэтому

$$f(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\xi)}{\xi - z} d\xi$$

для всех  $z \neq a$  внутри  $\gamma$ . Это означает, что справа стоит голоморфная функция, определенная внутри  $\gamma$ , причем она совпадает с  $f$  при  $z \neq a$ . Очевидно, что такая продолженная функция единственна, потому что для нее должна выполняться интегральная формула Коши. В дальнейшем мы будем обозначать эту продолженную функцию также буквой  $f$ .

Пусть  $f$  голоморфна в области  $D$ , и пусть  $a$  — точка в этой области. Применим принцип устранимых особенностей к выражению

$$f_1(z) = \frac{f(z) - f(a)}{z - a}.$$

Очевидно, что функция  $f_1$  определена и голоморфна в области  $D - \{a\}$  и обладает свойством  $\lim_{z \rightarrow a} (z - a)f_1(z) = 0$ . Следовательно, ее можно продолжить до голоморфной функции, определенной во всей области  $D$ , и написать

$$f_1(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\xi) - f(a)}{(\xi - a)(\xi - z)} d\xi.$$

При этом

$$\frac{1}{(\xi - a)(\xi - z)} = (a - z)^{-1} \left( \frac{1}{\xi - a} - \frac{1}{\xi - z} \right),$$

откуда

$$\begin{aligned} (a - z)f_1(z) &= \frac{1}{2\pi i} \int_{\gamma} \frac{f(\xi) - f(a)}{\xi - a} d\xi - \frac{1}{2\pi i} \int_{\gamma} \frac{f(\xi) - f(a)}{\xi - z} d\xi \\ &= f(a) - f(z).^6 \end{aligned}$$

---

<sup>6</sup>В этой и следующей формуле учитывается, что

$$\int_{\gamma} \frac{d\xi}{(\xi - z)(\xi - a)} = 0,$$

и этот интеграл прибавляется к выражению для  $f_1(z)$  с коэффициентом  $f(a)/2\pi i$ . — Прим. ред.

Таким образом,

$$f(z) = f(a) + (z - a)f_1(z), \quad \text{где} \quad f_1(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\xi)}{(\xi - z)(\xi - a)} d\xi.$$

Конечно,

$$f_1(a) = f'(a).$$

Ту же процедуру можно применить к  $f_1$  и определить  $f_2$  из выражения  $f_1(z) = f_1(a) + (z - a)f_2(z)$ . Таким образом можно проделать  $n$  шагов

$$\begin{aligned} f(z) &= f(a) + (z - a)f_1(z), \\ f_1(z) &= f_1(a) + (z - a)f_2(z), \\ &\vdots \\ f_{n-1}(z) &= f_{n-1}(a) + (z - a)f_n(z), \end{aligned}$$

так что

$$f(z) = f(a) + (z - a)f_1(a) + \cdots + (z - a)^{n-1}f_{n-1}(a) + (z - a)^n f_n(z).$$

Дифференцируя это формулу  $k \leq n$  раз, получим  $f_k(a) = (1/k!)f^{(k)}(a)$ . В результате получается ряд Тейлора с остатком

$$f(z) = f(a) + (z - a)f'(a) + \frac{(z - a)^2}{2!}f''(a) + \cdots + (z - a)^n f_n(z), \quad (20.20)$$

где, как и раньше,

$$f_n(z) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(\xi)}{(\xi - a)^n(\xi - z)} d\xi \quad (20.21)$$

для всех  $z$  внутри  $\gamma$ .<sup>7</sup>

Из формулы (20.21) можно получить полезную оценку остатка. Предположим, что  $\gamma$  — окружность с радиусом  $R$  и центром

---

<sup>7</sup>Доказательство формулы (20.21) индукцией по  $n$  оставляем читателю в качестве упражнения. — Прим. ред.

в точке  $a$ , а  $M$  обозначает максимальное значение  $|f(\xi)|$  для  $\xi$  на окружности  $\gamma$ . Подстановка в формулу (20.21) дает

$$|f_n(z)| \leq \frac{M}{R^{n-1}(R - |z - s|)}. \quad (20.22)$$

Поскольку  $f_n(a) = (1/n!)f^{(n)}(a)$ , с помощью формул (20.21) и (20.22) мы получаем

$$f^{(n)}(a) = \frac{n!}{2\pi i} \int_{\gamma} \frac{f(\xi) d\xi}{(\xi - z)^{n+1}} \quad (20.23)$$

и

$$|f^{(n)}(a)| \leq Mn!R^{-n}. \quad (20.24)$$

(Здесь  $\gamma$  — окружность с радиусом  $R$ , центром в точке  $a$ ,  $M$  — максимум функции  $f$  на этой окружности.)

Обратите внимание на замечательное логическое развитие ситуации. Мы начали с предположения, что функция  $f$  имеет непрерывную производную первого порядка, которая удовлетворяет уравнениям Коши–Римана в частных производных. Из этих уравнений следует, что  $f$  имеет производные всех порядков, и что  $n$ -я производная функции  $f$  (в комплексном смысле) существует и определяется простым выражением (20.23) через значения самой функции  $f$ .

В следующем параграфе мы получим ряд следствий из этого факта.

## 20.5. Приложения и следствия

В качестве первого примера применения формулы (20.24) докажем *теорему Лиувилля*.

Голоморфная и ограниченная на всей плоскости функция должна быть постоянной.

*Доказательство.* Пусть  $|f(z)| \leq M$  для всех  $z$ . В формуле (20.24) положим  $n = 1$ . Значение  $R$  можно выбрать сколь угодно большим, а это значит, что  $f'(a) = 0$  для всех  $a$ . Таким образом,

обе частные производные  $f$  тождественно равны нулю, т.е.  $f$  постоянна.

Следствием теоремы Лиувилля является *основная теорема алгебры*, согласно которой

любой полином положительной степени имеет по крайней мере один корень.

*Доказательство.* Напишем  $P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_0$ , где  $n > 0$  и  $a_n \neq 0$ . Тогда

$$|P(z)| \geq [|a_n| - (|a_{n-1}| |z|^{-1} + \dots + |a_0| |z|^{-n})] |z|^n.$$

Для достаточно больших  $|z|$  выражение в квадратных скобках  $> \frac{1}{2}|a_n|$ . Таким образом,  $|P(z)| \rightarrow \infty$  при  $|z| \rightarrow \infty$ , и, следовательно, функция  $1/P$  голоморфна и ограничена вне некоторой окружности большого радиуса. Если бы  $P(z)$  нигде не обращалась в нуль, то функция  $1/P$  была бы голоморфна и ограничена на всей плоскости, т.е. была бы постоянной. Но поскольку  $P$  — не постоянная функция, то это невозможно. Таким образом, должен быть, по крайней мере, один нуль функции  $P$ , что и требовалось доказать.

Если  $P(a) = 0$ , то  $(z - a)[P(z)/(z - a)] \rightarrow 0$ , и мы можем воспользоваться принципом устранимых особенностей и утверждать, что  $P(z)/(z - a)$  — голоморфная функция. Далее, поскольку  $z^j - a^j = (z - a)(z^{j-1} + \dots + a^{j-1})$ , то  $P(z) - P(a)$  можно разделить на  $z - a$ . А поскольку  $P(a) = 0$ , то  $P(z)/(z - a)$  — полином степени  $n - 1$ . Если  $n > 1$ , то опять можно применить основную теорему алгебры и утверждать, что

полином степени  $n$  имеет ровно  $n$  корней, считая с кратностями.

Следующий пример. Покажем, что голоморфная функция, определенная в некоторой связной области  $D$ , не может обращаться в нуль вместе со всеми производными в любой точке  $a$  внутри области  $D$ , если только она не равна нулю тождественно. Действительно, предположим, что  $f(a) = 0$  и  $f^{(n)}(a) = 0$  для всех  $n$ . Выберем  $R$  так, чтобы диск радиуса  $R$  с центром в точке

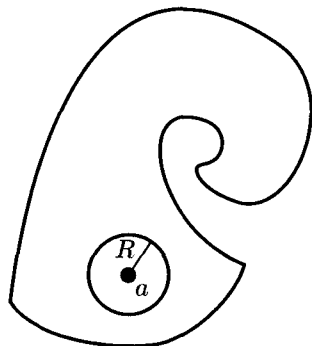


Рис. 20.5

$a$  целиком лежал в области  $D$  (рис. 20.5). Сначала покажем, что  $f$  тождественно равна нулю на этом диске. Из формул (20.20) и (20.22) мы знаем, что (20.22)

$$f(z) = (z - a)^{n+1} f_{n+1}, \quad \text{где} \quad |f_{n+1}(z)| \leq \frac{M}{R^n(R - |z - a|)},$$

откуда

$$|f(z)| \leq \frac{M}{(R - |z - a|)} \frac{|z - a|^{n+1}}{R^n}.$$

Поскольку  $|z - a| < R$  и  $n$  может быть сколь угодно большим, это неравенство означает, что  $f(z) = 0$ . А теперь возьмем произвольную точку внутри диска в качестве нового центра, увеличив, таким образом, область, где мы знаем, что  $f(z) = 0$  (рис. 20.6). Если  $b$  — любая другая точка в области  $D$ , то точки  $a$  и  $b$  можно соединить кривой, которая вся лежит в области  $D$  и, следовательно, на конечном расстоянии от границы. Можно нарисовать конечное число дисков (рис. 20.7), начав с диска, имеющего центр в точке  $a$  и закончив диском с центром в точке  $b$ , получив, таким образом,  $f(b) = 0$ . Итак, мы доказали следующее.

Предположим, что  $f$  голоморфна в некоторой связной области  $D$  и равна нулю вместе со всеми производными в точке  $a$ , лежащей внутри области  $D$ . Тогда  $f$  тождественно равна нулю.

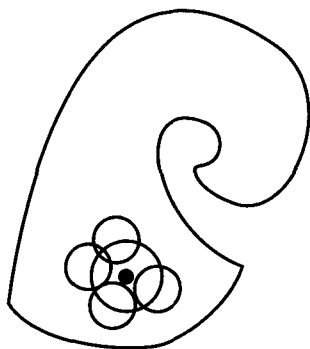


Рис. 20.6

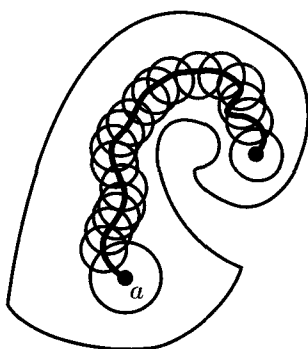


Рис. 20.7

Предположим, что  $a_n$  — последовательность точек в области  $D$ , причем  $a_n \rightarrow a$ , где  $a$  тоже лежит внутри  $D$ . Предположим, что  $f(a_n) = 0$  для всех  $n$ . Тогда  $f(a) = \lim f(a_n) = 0$  и  $f'(a) = \lim (f(a) - f(a_n))/(a - a_n) = 0$ . Утверждается, что в этом случае все производные  $f^{(k)}(a) = 0$  и, следовательно,  $f \equiv 0$ . Действительно, предположим, что  $k$  — первое положительное целое число, для которого  $f^{(k)}(a) \neq 0$ . Тогда, в силу (20.20),  $f(z) = (z - a)^k f_k(z)$ , где  $f_k(a) \neq 0$ , и, следовательно,  $f_k(z) \neq 0$  для  $z$  в окрестности  $a$ . Но поскольку  $z - a \neq 0$  для  $z \neq a$ , можно утверждать, что  $f(z) \neq 0$  для всех  $z$ , достаточно близких, но не равных  $a$ . Но это противоречит исходному предположению, что  $f(a_n) = 0$ , где  $a_n \rightarrow a$ . Только что доказанную теорему формулируют так: нули (нетривиальной) голоморфной функции должны быть изолированными. Если  $f$  и  $g$  — две голоморфные функции, то предыдущий результат можно применить к голоморфной функции  $f - g$ . Тогда утверждается следующее.

Если  $f$  и  $g$  — голоморфные функции в связной области  $D$  и если  $f(a_n) = g(a_n)$  для последовательности точек  $a_n$ , имеющих предел  $a = \lim a_n$ , лежащий внутри  $D$ , то  $f = g$  во всей области  $D$ .

Этот результат показывает, насколько сильно отличается теория голоморфных функций от теории бесконечно дифференци-

руемых функций (класса  $C^\infty$ ) вещественного переменного. Функция  $f$ , заданная формулой

$$f(x) = \begin{cases} e^{-1/x} & \text{для } x > 0, \\ 0 & \text{для } x \leq 0, \end{cases}$$

является  $C^\infty$ -функцией в пространстве  $\mathbb{R}$ . Она тождественно равна нулю для всех отрицательных  $x$ . Таким образом, поведение  $C^\infty$ -функции в одной части области ее определения не оказывает никакого влияния на ее поведение в другой части, расположенной на конечной расстоянии. Ситуация с голоморфными функциями совсем другая. Если нам задано поведение голоморфной функции в какой-то части области определения, то она полностью определена во всей области.<sup>8</sup>

Мы уже видели, что голоморфная функция  $g$  не может быть равна нулю вместе со всеми производными в любой точке внутри области определения. Поэтому, если  $g(\alpha) = 0$  в некоторой точке  $\alpha$ , то существует наименьшее число  $k$ , для которого  $g^{(k)}(\alpha) \neq 0$ . Следовательно,

$$g(z) = (z - \alpha)^k g_k(z) \neq 0 \quad \text{для } z \text{ вблизи } \alpha, \quad z \neq \alpha,$$

где  $g_k$  — голоморфная функция. Предположим, что  $f$  тоже голоморфна в той же области. Давайте рассмотрим функцию  $h = f/g$ , которая определена для  $z$  вблизи  $\alpha$ , но  $z \neq \alpha$ . Напишем разложение (20.20) функции  $f$  в окрестности  $\alpha$

$$f(z) = f(\alpha) + (z - \alpha)f'(\alpha) + \dots + (z - \alpha)^{(k-1)} \frac{f^{(k-1)}(\alpha)}{(k-1)!} + (z - \alpha)^k f_k(z).$$

Тогда

$$\begin{aligned} h(z) = f(z)/g(z) &= \frac{f(\alpha)}{(z - \alpha)^k g_k(z)} + \frac{f'(\alpha)}{(z - \alpha)^{k-1} g_k(z)} + \dots \\ &\quad + \frac{f^{(k-1)}(\alpha)}{(z - \alpha)(k-1)! g_k(z)} + \frac{f_k(z)}{g_k(z)}. \end{aligned}$$

<sup>8</sup>Под частью области определения имеется в виду ее открытое подмножество. Процедура восстановления голоморфной функции по ее *ростку*, определенному на такой части, во всю область называется *аналитическим продолжением*. — Прим. ред.

Поскольку  $1/g_k$  голоморфна вблизи  $\alpha$ , то ее также можно разложить в окрестности  $\alpha$ :

$$1/g_k(z) = b_0 + b_1(z - \alpha) + \dots + (z - \alpha)^k b_k(z),$$

где

$$b_0 = 1/g_k(\alpha) = K!/g^{(k)}(\alpha) \quad \text{и т. д.}$$

Если это разложение подставить в формулу для  $h$ , то получим

$$h(z) = a_{-k}(z - \alpha)^{-k} + a_{-k+1}(z - \alpha)^{-(k+1)} + \dots + a_{-1}(z - \alpha)^{-1} + h_0(z),$$

где

$$\begin{aligned} a_{-k} &= f(\alpha)b_0 = f(\alpha)/g_k(\alpha), \\ a_{-k+1} &= f(\alpha)b_1 + f'(\alpha)b_0, \end{aligned}$$

и т. д., а  $h_0$  голоморфна вблизи  $\alpha$ . Итак, мы доказали следующее.

Частное двух голоморфных функций — мероморфная функция.

Предыдущие формулы упрощаются, если  $g$  имеет *простой* нуль в точке  $\alpha$ , т. е. если  $k = 1$ . В этом случае утверждается следующее.

Если  $g(\alpha) = 0$  и  $g'(\alpha) \neq 0$ , то частное  $f/g$  имеет простой полюс в точке  $\alpha$ . Вычет в этой точке равен  $f(\alpha)/g'(\alpha)$ , т. е.  $f(z)/g(z) = a_{-1}(z - \alpha)^{-1} + h_0(z)$  вблизи  $\alpha$ , где  $a_{-1} = f(\alpha)/g'(\alpha)$  и  $h_0(z)$  — голоморфная функция вблизи  $\alpha$ . (20.25)

Число  $k$ , которое мы использовали сейчас, называется *порядком* нуля. Таким образом, голоморфная функция  $g$  имеет нуль первого порядка в точке  $\alpha$ , если  $g(\alpha) = 0$  и  $g'(\alpha) \neq 0$ . Она имеет нуль порядка  $k$  в точке  $\alpha$ , если  $g(\alpha) = \dots = g^{(k-1)}(\alpha) = 0$ , но при этом  $g^{(k)}(\alpha) \neq 0$ .<sup>9</sup>

Предположим, что функция  $f$  голоморфна в ограниченной области  $D$  и имеет конечное число нулей в этой области, т. е. мы предполагаем, что есть конечное число различных точек  $\alpha_1, \dots, \alpha_r$ , где функция  $f$  равна нулю. И пусть  $k_j$  — порядок нуля

<sup>9</sup>Эквивалентное условие состоит в том, что функция  $1/g$  имеет в точке  $\alpha$  полюс порядка  $k$ . — Прим. ред.

функции  $f$  в точке  $\alpha_j$ . Тогда  $f(z) = (z - \alpha_1)^{k_1} g(z)$ , где функция  $g$  обращается в нуль только в точках  $\alpha_2, \dots, \alpha_r$  и имеет тот же порядок нулей в этих точках, что и функция  $f$ . Далее можно написать  $g(z) = (z - \alpha_2)^{k_2} h(z)$ , где функция  $h$  обращается в нуль в точках  $\alpha_3, \dots, \alpha_r$ . Продолжая таким образом, мы получим

$$f(z) = (z - \alpha_1)^{k_1} (z - \alpha_2)^{k_2} \dots (z - \alpha_r)^{k_r} F(z),$$

где  $F$  — голоморфная функция в области  $D$ , которая не равна нулю ни в одной точке этой области. Продифференцировав это выражение и разделив его на  $f$ , получаем

$$\frac{f'(z)}{f(z)} = \frac{k_1}{z - \alpha_1} + \dots + \frac{k_r}{z - \alpha_r} + \frac{F'(z)}{F(z)},$$

где  $F'/F$  — голоморфная функция в области  $D$ . В частности, если  $g$  — другая голоморфная в области  $D$  функция, то  $gf'/f$  имеет только простые полюса в точках  $\alpha_j$ , вычеты в которых равны  $k_j g(\alpha_j)$ . Используя формулу Коши, получаем следующий результат.

Предположим, что  $f$  и  $g$  — голоморфные функции в ограниченной области  $D$ , причем  $f$  имеет в области  $D$  конечное число нулей в точках  $\alpha_1, \dots, \alpha_r$ , порядок которых равен  $k_1, \dots, k_r$  соответственно. Пусть, кроме того,  $f \neq 0$  на границе  $\partial D$ .

$$\frac{1}{2\pi i} \int_{\partial D} g(\xi) \frac{f'(\xi)}{f(\xi)} d\xi = k_1 g(\alpha_1) + \dots + k_r g(\alpha_r). \quad (20.26)$$

В частности, если взять  $g \equiv 1$ , то имеем

$$\begin{aligned} \frac{1}{2\pi i} \int_{\partial D} \frac{f'(\xi)}{f(\xi)} d\xi &= k_1 + \dots + k_r \\ &= \text{числу нулей функции } f \\ &\quad \text{с учетом их кратности.} \end{aligned} \quad (20.27)$$

Заметим, что если  $f$  — голоморфная функция в большей области  $E$ , такой, что  $\overline{D}$  содержится в  $E$ , тогда  $f$  может иметь только конечное число нулей в области  $D$ . В противном случае  $f$  была бы

равна нулю в последовательности точек, сходящейся к пределу, расположенному в области  $E$ .<sup>10</sup> А это означало бы, что  $f$  тождественно равна нулю. Таким образом, можно утверждать, что в этом случае предположения для (20.26) и (20.27) справедливы.

*Теорема Руше* утверждает следующее.

Если  $f$  и  $h$  — голоморфные функции в ограниченной области  $D$  и непрерывны в  $\bar{D} = D \cup \partial D$ , и если всюду на  $D$   $|f(z) - h(z)| < |f(z)|$ , то  $f$  и  $h$  имеют одинаковое (20.28)  
число нулей (с учетом их кратности) в этой области.

Из посылок теоремы следует, что  $f$  не равна нулю на границе  $\partial D$ , и, следовательно, не равна нулю в полосе вблизи этой границы. Поэтому  $f$  имеет конечное число нулей в области  $D$ . То же самое можно сказать и о функции  $h$ . На самом деле можно утверждать, что для любого  $0 \leq t \leq 1$  функция  $h_t$ , заданная формулой

$$h_t(z) = f(z) + t(h(z) - f(z)),$$

голоморфна в области  $D$ , не равна нулю на границе  $\partial D$  и имеет конечное число нулей в  $D$ , которые можно сосчитать по формуле (20.27). Формула (20.26) непрерывно зависит от  $t$  и дает в результате целое число. Следовательно, это число является константой. При этом  $h_0 = f$  и  $h_1 = h$ , так что  $f$  и  $h$  должны иметь одинаковое число нулей, что и требовалось доказать.

Предположим, что функция  $f$  голоморфна во всех точках  $z \neq \beta$  в окрестности  $\beta$ , но  $|f(z)| \rightarrow \infty$  при  $z \rightarrow \beta$ . Тогда функция  $f \neq 0$  для  $z$  достаточно близких к  $\beta$  и поэтому  $1/f(z)$  — голоморфная для  $z$  достаточно близких к  $\beta$ , кроме  $z = \beta$ , и  $1/f(z) \rightarrow 0$  при  $z \rightarrow \beta$ . Это означает, что  $1/f$  имеет устранимую особенность в точке  $\beta$ . Таким образом, она становится голоморфной не только вблизи  $\beta$ , но и в самой этой точке, если мы положим ее там равной нулю. Поскольку теперь  $1/f$  голоморфна в точке  $\beta$  и в ее окрестности, и она — не тождественный нуль, то она должна иметь в точке  $\beta$  нуль конечного порядка, т.е.  $1/f(z) = (z - \beta)^j g_j(z)$ , где  $g_j$  голоморфна вблизи  $\beta$  и  $g_j(\beta) \neq 0$ . Но тогда

$$f(z) = (z - \beta)^{-j} h(z), \quad \text{где } h = 1/g_j \text{ голоморфна вблизи } \beta.$$

<sup>10</sup>Точнее, на границе  $\partial D$ , которая в силу ограниченности  $D$  является компактным множеством. — *Прим. ред.*

Если для функции  $h$  использовать разложение Тейлора в точке  $\beta$ , то увидим, что  $f$  имеет полюс порядка  $j$  в этой точке. Итак, мы доказали, что

если  $f$  голоморфна всюду при  $z \neq \beta$  в окрестности  $\beta$  и  $|f(z)| \rightarrow \infty$ , когда  $z \rightarrow \beta$ , то она имеет полюс конечного порядка в точке  $\beta$ .

Предположим, что  $f$  мероморфна в ограниченной области  $D$ , непрерывна на границе  $\partial D$  и не обращается нигде на ней в нуль. Тогда эта функция может иметь лишь конечное число нулей и полюсов в области  $D$ . Прделав вышеописанную процедуру с каждым нулем и с каждым полюсом, получим

$$f(z) = \frac{(z - \alpha_1)^{k_1} \dots (z - \alpha_r)^{k_r}}{(z - \beta_1)^{j_1} \dots (z - \beta_p)^{j_p}} F(z),$$

где  $F$  голоморфна и не имеет нулей в области  $D$ . (Здесь  $\alpha_i$  — нули порядка  $k_i$ , и  $\beta_i$  — полюса порядка  $j_i$ .) Далее мы можем рассуждать так же, как и при доказательстве (20.25) и (20.26). В результате получаем следующее.

Предположим, что  $f$  — голоморфная функция в ограниченной области  $\bar{D}$ , непрерывна в  $D$  и вблизи границы  $\partial D$ . Далее предположим, что  $g$  — голоморфная функция в области  $\bar{D}$  и непрерывна на ней. Тогда

$$\begin{aligned} \frac{1}{2\pi i} \int_{\partial D} g(\xi) \frac{f'(\xi)}{f(\xi)} d\xi \\ = k_1 g(\alpha_1) + \dots + k_r g(\alpha_r) - (j_1 g(\beta_1) + \dots + j_p g(\beta_p)). \end{aligned}$$

Здесь  $\alpha_1, \dots, \alpha_r$  — нули функции  $f$  порядка  $k_1, \dots, k_r$  соответственно, и  $\beta_1, \dots, \beta_p$  — полюса этой функции  $f$  порядка  $j_1, \dots, j_p$ . Пусть  $g \equiv 1$ , тогда

$$\frac{1}{2\pi i} \int_{\partial D} \frac{f'(\xi)}{f(\xi)} d\xi = \begin{aligned} &\text{числу нулей функции } f \\ &\quad - \text{число ее полюсов} \\ &\quad (\text{с учетом их кратности}). \end{aligned}$$

Предположим, что  $f$  голоморфна для  $z \neq \gamma$  в окрестности точки  $\gamma$ . И пусть  $\gamma$  — не устранимая особенность функции, и не является ее полюсом (так что у нас нет условия  $|f(z)| \rightarrow \infty$  при  $z \rightarrow \gamma$ ). В этом случае  $\gamma$  называется *существенной особенностью*  $f$ . Некоторое представление о сложном поведении функции вблизи существенной особенности дает следующий результат.

Допустим, что  $f$  голоморфна вблизи  $\gamma$ , но не в самой точке, где эта функция имеет существенную особенность. Тогда для любого заданного комплексного (20.29) числа  $c$  можно найти такую последовательность точек  $a_i \rightarrow \gamma$ , что  $f(a_i) \rightarrow c$ .

Докажем от противного. Предположим, что это утверждение не выполняется для некоторого числа  $c$ . Это значит, есть такая окрестность  $U$  точки  $\gamma$ , что  $f(z)$  отличается от  $c$  на конечное значение при всех  $z$  внутри области  $U$ ,  $z \neq \gamma$ . Скажем, пусть  $|f(z) - c| > 1/M$  для некоторого числа  $M$ . Определим функцию  $g(z) = 1/(f(z) - c)$ . Она голоморфна в области  $U$  везде, кроме точки  $\gamma$ , и  $|g(z)| < M$ . Поэтому функция  $g$  в точке  $\gamma$  имеет устранимую особенность. Но тогда  $f = 1/g + c$  мероморфна в точке  $\gamma$  (в худшем случае имеет там полюс), что противоречит нашему предположению. Примером функции, имеющей существенную особенность в точке 0, является  $e^{1/z}$ . Покажем в явном виде, что для этой функции справедливо условие (20.29). Мы должны показать, что значение функции  $e^{1/z}$  можно сделать сколь угодно близким к любому комплексному числу, если  $z \rightarrow 0$ . Сделаем замену переменной  $w = 1/z$ . Тогда мы хотим показать, что значение  $e^w$  можно сделать сколь угодно близким к любому комплексному числу  $c$  при условии  $|w| > R$  для произвольного значения  $R$ . Напишем  $c = re^{i\theta}$  и  $w = u + iv$ . Тогда условие  $|w| > R$  выполняется, если выбрать  $v > R$ . Теперь  $e^w = e^u \cdot e^{iv}$ . Поэтому, если  $r > 0$ , то можно взять  $u = \log r$  и  $v = \theta + 2\pi n$ , где  $n$  — достаточно большое целое число, чтобы выполнялось условие  $v > R$ . Мы нашли точное решение уравнения  $e^w = c$  при сколь угодно большом  $n$ , так что  $z = 1/w$  приближается к нулю. Есть только одно значение комплексного числа  $c$ , в которое мы не можем попасть — это  $c = 0$ . Но можно взять последовательность  $c_k \rightarrow 0$ ,  $c_k \neq 0$  и соответствующую последовательность  $w_k$ , где  $|w_k| \rightarrow \infty$

и  $e^{w_k} = c_k$ . Тогда  $z_k = w_k$  — искомая последовательность точек, стремящихся к нулю, когда  $e^{z_k} \rightarrow 0$ .

Такое поведение функции  $e^{1/z}$  типично для голоморфных функций вблизи существенной особенности, т.е. мы можем решить точно уравнение  $f(z) = c$  в любой окрестности  $z = \gamma$  для всех комплексных чисел  $c$ , кроме одного. (Для функции  $e^{1/z}$  это было  $c = 0$ .) Это факт весьма глубокий, и доказательства его мы здесь не приводим.

## 20.6. Локальное отображение

В этом параграфе мы исследуем более подробно локальные свойства голоморфных функций. Первый результат — это очевидное следствие теоремы Руше:

Предположим, что  $f$  голоморфна и не постоянна в окрестности точки  $a$ , причем  $f(a) = b$ . Тогда существует такое  $\varepsilon > 0$ , что для всех  $w$ , удовлетворяющих (20.30) условию  $|w - b| < \varepsilon$ , найдется такое  $z$  вблизи  $a$ , что  $f(z) = w$ .

*Доказательство.* Пусть  $g(z) = f(z) - b$ . Тогда  $g$  имеет нуль конечного порядка в точке  $a$ . Мы можем найти достаточно малое  $r$ , для которого  $g(z) \neq 0$ , если  $|z - a| \leq r$ , кроме самой точки  $a$ . Пусть область  $D$  — это диск радиуса  $r$  с центром в точке  $a$ . Выберем такое  $\varepsilon > 0$ , что  $|g(z)| > \varepsilon$  для  $z \in \partial D$ . Предположим, что  $|w - b| < \varepsilon$ , и определим  $g_w$  условием

$$g_w(z) = f(z) - w = g(z) + (b - w).$$

В силу теоремы Руше (20.28), функция  $g_w$  имеет такое же число нулей (с учетом их кратности), в области  $D$ , что и функция  $g$ . Поскольку  $g$  имеет, по крайней мере один нуль, то и  $g_w$  имеет, по крайней мере, один нуль, поэтому можно найти, по крайней мере, одно значение  $z$  в области  $D$ , где  $f(z) = w$ .

Утверждение (20.30) иногда формулируется по-другому: голоморфная функция определяет *открытое* отображение, т.е. если точка принадлежит образу, то в нем целиком содержится и некоторая окрестность этой точки.

Утверждение (20.30) демонстрирует еще одно поразительное отличие теории голоморфных функций комплексной переменной от теории гладких функций вещественной переменной. Рассмотрим отображение  $\mathbb{R} \rightarrow \mathbb{R}$ , преобразующее  $x$  в  $u = x^2$ . Тогда значения  $u$  могут быть только неотрицательными (рис. 20.8). Поэтому 0 принадлежит образу этого отображения, но никакая окрестность 0 в нем не лежит. Мы можем легко изменить этот пример, рассмотрев отображение плоскости в плоскость, т. е. пусть  $(x, y)$  преобразуется в  $(u, v)$ , где  $u = x^2$  и  $v = y$ . При таком отображении плоскость складывается вдоль оси  $y$ , так что в образе нет точек  $(u, v)$  с отрицательными  $u$  (рис. 20.9).

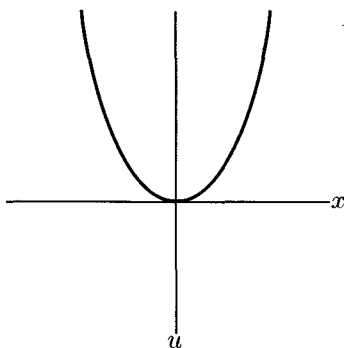


Рис. 20.8

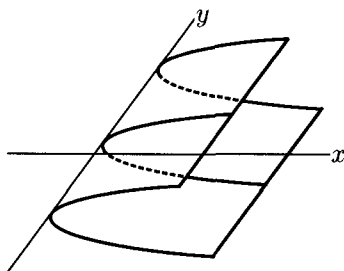


Рис. 20.9

Утверждение (20.30) говорит, что такого не бывает в случае отображений, соответствующих голоморфным функциям. Например, если взять  $f(z) = z^2$ , то для  $w = re^{i\theta} \neq 0$  можно найти два значения  $z$ , а именно  $z = r^{1/2}e^{i\theta/2}$  и  $z = r^{1/2}e^{i(\pi+\theta/2)}$ , для которых  $z^2 = w$ . Поэтому вместо отсутствия решений уравнения  $z^2 = w$  в некоторой области  $w$  мы имеем два решения.

Это также типичное свойство функции  $f(z) = z^2$ . Предположим, что  $f$  — голоморфная функция в окрестности  $a$ , где  $f(a) = b$ . Вернемся к доказательству (20.30). Функция  $g$ , введенная при доказательстве, имеет нуль конечного порядка в точке  $a$ . Пусть  $k$  — порядок этого нуля. Если  $k > 1$ , то  $g'(a) = 0$ . Можно выбрать такое  $r$  (возможно, меньшее, чем при доказательстве (20.30)), так что не только  $g$ , но и  $g'$  не имеет нулей в области  $D$ .

Таким образом, в обоих случаях  $k = 1$  и  $k > 1$  функция  $g$  вообще не имеет нулей в области  $D$ , за исключением точки  $a$ , а функция  $g'$  не имеет нулей в  $D$ , кроме, возможно, точки  $a$ . Заметим, что функция  $g_w$ , введенная при доказательстве (20.30), отличается от  $g$  на константу. Таким образом,  $g'_w = g'$  и, следовательно,  $g'_w(z) \neq 0$  в области  $D$  при  $z \neq a$ . Однако, для  $w \neq b$  функция  $g_w$  не имеет нуля в точке  $a$ . Это означает, что каждый из нулей  $g_w$  должен быть простым и, следовательно, у этой функции должно быть  $k$  различных нулей. Итак, мы доказали следующее.

Пусть  $f$  — голоморфная функция в окрестности точки  $a$  и  $f(a) = b$ . Предположим, что  $g$ , определенная формулой  $g(z) = f(z) - b$ , имеет нуль порядка  $k$  в точке  $a$ . Тогда можно найти  $r > 0$  и  $\varepsilon > 0$  такие, что (20.31) для любого  $w \neq b$  из окрестности  $|w - b| < \varepsilon$  существует  $k$  различных значений  $z$ :  $z_1, \dots, z_k$ , для которых  $|z_i - a| < r$  и  $f(z_i) = w$ .

Утверждение (20.31) демонстрирует истинный смысл нуля порядка  $k$ . Это ситуация, при которой  $k$  различных корней совпадают.

Рассмотрим частный случай  $k = 1$  утверждения (20.31), который весьма важен для нас.

Пусть  $f$  — голоморфная функция в окрестности точки  $a$ , где  $f(a) = b$  и  $f'(a) \neq 0$ . Тогда можно найти такие  $r > 0$  и  $\varepsilon > 0$ , что для всех  $w$ , удовлетворяющих условию  $|w - b| < \varepsilon$ , существует единственное значение  $z$ , для которого  $|z - a| < r$  и  $f(z) = w$ . Другими словами, существует единственная функция  $g$ , определенная при  $|w - b| < \varepsilon$ , удовлетворяющая условиям  $|g(w) - b| < r$  и  $f(g(w)) = w$ . (20.32)

При этом мы утверждаем, что

функция  $g$ , определенная в (20.32), голоморфна и (20.33)  $g'(w) = 1/f'(g(w))$ .

В утверждении (20.32) мы можем выбрать  $r$  так, чтобы  $f'(z) \neq 0$  для  $|z - a| < r$ . Поэтому, чтобы доказать (20.33), достаточно доказать, что функция  $g$  дифференцируема в комплексном смысле в точке  $b$  и  $g'(b) = 1/f'(a)$ . Аналогичные аргументы

потом приведем для произвольной точки  $z = g(w)$ , и формула для  $g'$  покажет, что она непрерывна. Итак, предположим, что  $f'(a) = c \neq 0$ . Пусть  $\eta < \frac{1}{2}|c|$ . Из определения производной функции  $f$  можно найти такое  $\delta > 0$ , что

$$\left| \frac{f(z) - f(a)}{z - a} - f'(a) \right| < \eta \quad \text{для} \quad |z - a| < \delta. \quad (20.34)$$

Выбирая  $\delta$  в качестве нового  $r$  (если это необходимо) в (20.32), мы видим, что для всех  $w$ , достаточно близких к  $b$ , имеем

$$\left| \frac{w - b}{z - a} - c \right| < \frac{1}{2}|c|, \quad \text{поэтому} \quad |w - b| > \frac{1}{2}|c(z - a)|,$$

где  $z = g(w)$  или

$$\left| \frac{z - a}{(w - b)c} \right| < \frac{2}{|c|^2}.$$

Если теперь умножить обе части равенства (20.34) на соответствующие части предыдущего неравенства, то получим

$$\left| \frac{1}{f'(a)} - \frac{g(w) - a}{w - b} \right| < 2|c|^{-2}\eta.$$

Мы можем сделать  $\eta$  сколь угодно малым, выбирая  $\delta$  и, следовательно,  $\varepsilon$  достаточно малыми. Это и доказывает, что  $g$  дифференцируема в комплексном смысле в точке  $b$  и  $g'(b) = 1/f'(a)$ , тем самым доказано утверждение (20.33). Утверждения (20.32) и (20.33) составляют теорему о неявной (обратной) функции для случая голоморфных функций. Следует, однако, подчеркнуть, что эта теорема локальная. Давайте еще раз вернемся к функции  $f(z) = z^2$ . Известно, что для любого значения  $w \neq 0$  есть два корня уравнения  $f(z) = w$ . Предположим, что мы выбрали квадратный корень положительного числа положительным. Например, для  $f(a) = 1$  возьмем  $a = 1$ . Утверждение (20.32) означает, что для  $w$  в окрестности единицы, мы будем выбирать единственный квадратный корень — тот, который близок к единице, а утверждение (20.33) означает, что функция  $g(w) = w^{1/2}$  будет голоморфной в окрестности  $w = 1$ . Действительно, легко

видеть, что в (20.33) можно взять в качестве  $\varepsilon$  любое число, меньшее единицы. Конечно, мы можно взять любую точку  $w$  внутри диска радиуса  $\varepsilon$  с центром в  $w = 1$  в качестве нового значения  $b$ , а  $g(w)$  в качестве нового значения  $a$ , и еще раз применить теорему о неявной функции. Предположим, что мы взяли последовательность таких точек с условием  $|w| \approx 1$ , как это показано на рис. 20.10. Когда, пройдя таким образом по всему кругу, мы вернемся на положительную ось  $w$ , то получим отрицательное (противоположное) значение квадратного корня, что не удивительно. Итак, хотя (20.32) гарантирует существование локальной функции, обратной к  $f$ , последовательное применение (20.32) может привести к глобальной несогласованности. В случае функции  $f(z) = z^2$  не существует глобально определенной функции  $w^{1/2}$  на плоскости  $w$ .

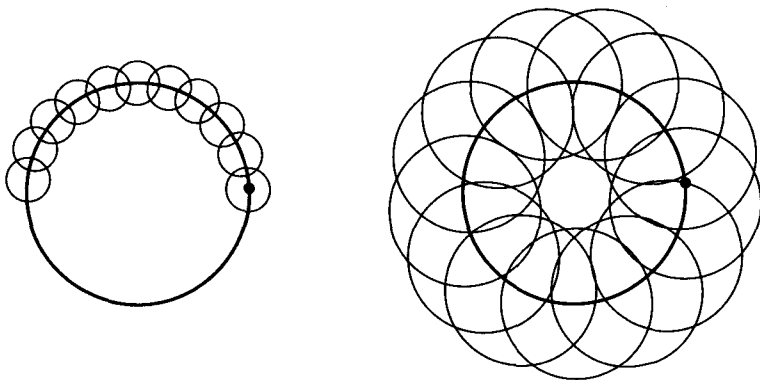


Рис. 20.10

Существует два способа работать с этой проблемой неопределенности, и каждый из них мы продемонстрируем на примере квадратного корня. Один способ — просто жить с этой неопределенностью. Плоскость  $z$  рассматривается как *двулистное накрытие* плоскости  $w$  с *точкой ветвления* в нуле, где значению  $w$  соответствует только одно значение  $z$ . Любую заданную функцию типа  $\cos(w^{3/2} + 1)$  мы рассматриваем не как функцию, определенную в плоскости  $w$ , а как функцию от  $z$ , т. е.  $\cos(z^3 + 1)$ . Плоскость  $z$  — это *риманова поверхность*, связанная со всеми функциями

$w^{1/2}$  в том смысле, что все они определены как функции от  $z$ . Такой точки зрения придерживался сам Риман. Подробное исследование структуры римановых поверхностей различных голоморфных функций способствовало развитию геометрии, топологии и алгебры вплоть до настоящего времени. Однако, здесь мы не будем останавливаться на этом.

Другой, более приземленный подход к проблеме определения  $w^{1/2}$  состоит в задании некоторой кривой, выходящей из начала координат и идущей на бесконечность вдоль пути, где мы считаем, что  $w^{1/2}$  не определен, чтобы сделать определение корня единственным везде. Например, предположим, что мы *разрезали* плоскость  $w$  вдоль отрицательной оси  $w$ . Таким образом, каждое значение  $w$  вне этого *разреза* может быть записано в виде  $w = re^{i\theta}$ , где  $-\pi < \theta < \pi$ . Тогда, выбирая  $w^{1/2} = r^{1/2}$  для  $\theta = 0$  (т.е. положительный квадратный корень для положительных  $w$ ), и двигаясь далее (через первый и четвертый квадрант) ко всем значениям  $\theta$  указанного интервала, имеем

$$w^{1/2} = r^{1/2} e^{i\theta/2} \quad \text{для} \quad -\pi < \theta < \pi.$$

Поэтому, например, при таком выборе квадратного корня

$$i^{1/2} = e^{i\pi/4} \quad \text{и} \quad (-i)^{1/2} = e^{-i\pi/4} \quad (\text{рис. 20.11}).$$

Вопрос, где провести разрез, решается каждый раз в зависимости от задачи. Вышеуказанный выбор разреза применяется довольно часто, и иногда его называют *главной ветвью* квадратного корня.

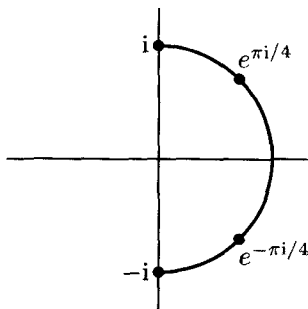


Рис. 20.11

Покажем, как выбор разреза применяется для вычисления некоторых интегралов, которые потом будут очень важны для нас. Рассмотрим интеграл

$$\int_{-\infty}^{\infty} e^{-\lambda x^2/2} dx.$$

Когда  $\operatorname{Re} \lambda > 0$ , функция  $e^{-\lambda x^2/2}$  быстро стремится к нулю на  $\pm\infty$ , и интеграл сходится абсолютно. Когда же  $\operatorname{Re} \lambda < 0$ , функция  $e^{-\lambda x^2/2}$  стремится к бесконечности, и этот интеграл не имеет смысла. Когда  $\operatorname{Re} \lambda = 0$ , абсолютное значение функции  $e^{-\lambda x^2/2}$  равно единице для всех  $x$ , поэтому заведомо ясно, что интеграл не сходится абсолютно. И тем не менее, утверждается, что интеграл сходится. По существу, утверждается следующее.

Интеграл  $\int_{-\infty}^{\infty} e^{-\lambda x^2/2} dx$  сходится равномерно, когда для  $\lambda$  выполняются условия  $\operatorname{Re} \lambda \geq 0$ ,  $|\lambda| \geq 1$ , и поэтому он является непрерывной функцией  $\lambda$  в этой области.

Чтобы доказать равномерную сходимость, надо показать, что для любого  $\varepsilon > 0$  можно найти  $R_\varepsilon$  (зависящее от  $\varepsilon$ , но не от  $\lambda$ ) такое, что

$$\left| \int_R^S e^{-\lambda x^2/2} dx \right| < \varepsilon \quad \text{и} \quad \left| \int_{-S}^{-R} e^{-\lambda x^2/2} dx \right| < \varepsilon$$

для всех  $S > R > R_\varepsilon$ . Далее,

$$\frac{d}{dx} e^{-\lambda x^2/2} = -\lambda x e^{-\lambda x^2/2},$$

поэтому интегрирование по частям дает

$$\begin{aligned} \int_R^S e^{-\lambda x^2/2} dx &= - \int_R^S \frac{1}{\lambda x} \frac{d}{dx} e^{-\lambda x^2/2} dx \\ &= \frac{1}{\lambda R} e^{-\lambda R^2/2} - \frac{1}{\lambda S} e^{-\lambda S^2/2} - \frac{1}{\lambda} \int_R^S \frac{1}{x^2} e^{-\lambda x^2/2} dx. \end{aligned}$$

Поскольку  $|e^{-\lambda x^2/2}| \leq 1$  для всех  $x$  и рассматриваемых значений  $\lambda$ , интеграл справа можно оценить как  $\int_R^S x^{-2} dx$ , и, таким образом, (поскольку  $S > R$ ) получаем оценку

$$\left| \int_R^S e^{-\lambda x^2/2} dx \right| \leq \frac{4}{|\lambda|R},$$

что меньше  $\varepsilon$ , если  $R > 4|\lambda|^{-1}\varepsilon^{-1}$ . Точно так же оценивается и  $\int_{-S}^{-R}$ . Поскольку  $|\lambda| \geq 1$ , то можно взять  $R_\varepsilon = 4/\varepsilon$ . (Заметим, что вместо условия  $|\lambda| \geq 1$  можно взять  $|\lambda| \geq c$ , где стоит любое фиксированное  $c > 0$ , и получить равномерную сходимость.) Поскольку функция  $e^{-\lambda x^2/2}$  непрерывно зависит от  $\lambda$  и  $x$ , равномерная сходимость интеграла обеспечивает то, что интеграл является непрерывной функцией  $\lambda$ . Вычислим этот интеграл для  $\operatorname{Re} \lambda > 0$ , где он сходится абсолютно. При этом поступим весьма хитро. А именно, вычислим квадрат этого интеграла, перейдя к полярным координатам:

$$\begin{aligned} \left[ \int_{-\infty}^{\infty} e^{-\lambda x^2/2} dx \right]^2 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\lambda x^2/2} \cdot e^{-\lambda y^2/2} dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\lambda(x^2+y^2)/2} dx dy = \int_0^{2\pi} \int_0^{\infty} e^{-\lambda r^2/2} r dr d\theta \\ &= 2\pi \int_0^{\infty} e^{-\lambda r^2/2} r dr = 2\pi/\lambda. \end{aligned}$$

Таким образом,

$$\int_{-\infty}^{\infty} e^{-\lambda x^2/2} dx = (2\pi/\lambda)^{1/2}. \quad (20.35)$$

Какой корень взять? Когда  $\lambda$  — вещественное положительное число, подынтегральная функция  $e^{-\lambda x^2/2}$  — тоже положительная, поэтому интеграл должен быть положительным. Следовательно, для вещественных и положительных  $\lambda$  мы должны выбрать положительный квадратный корень. Поскольку интеграл — непрерывная функция  $\lambda$ , мы должны взять главную ветвь квадратного корня. Это типичная иллюстрация того, как сама задача говорит нам, как определять квадратный корень.

В силу непрерывности, формула (20.35) справедлива для всех  $\operatorname{Re} \lambda \geq 0$ . В частности, положив  $\lambda = -it$  и  $\lambda = it$ , получаем важные формулы

$$\int_{-\infty}^{\infty} e^{itx^2/2} dx = e^{i\pi/4} (2\pi/t)^{1/2}, \quad \int_{-\infty}^{\infty} e^{-itx^2/2} dx = e^{-i\pi/4} (2\pi/t)^{1/2},$$

справедливые для всех вещественных  $t > 0$ , где подразумевается положительный квадратный корень. Эти интегралы, чрезвычайно важные для оптики, называются *интегралами Френеля*. Позднее мы к ним вернемся.<sup>11</sup>

Все, что здесь говорилось о квадратном корне, справедливо и для корня степени  $n$  (с небольшими изменениями), где  $n$  — любое положительное число. Это же на самом деле относится к экспоненциальной функции. Функция  $e^z$  отображает плоскость  $z$  на плоскость  $w$  бесконечное число раз. Каждая полоса ширины  $2\pi$ , параллельная оси  $x$ , отображается на всю<sup>12</sup> плоскость  $w$ . Логарифмическая функция  $\log w$  не однозначна. Она определена с точностью до произвольного кратного  $2\pi i$ . И на этот раз есть два подхода к неопределенности. Один подход: полагаем, что плоскость  $z$  является *бесконечнолистным накрытием* всей плоскости  $w$ . Любая функция, содержащая  $\log w$ , рассматривается как функция  $z$ . Другой подход: разрезаем плоскость  $w$ , например, вдоль отрицательной вещественной оси. Тогда можно написать

$$w = re^{i\theta}, \quad -\pi < \theta < \pi.$$

*Главная ветвь* логарифмической функции определяется при этом условием

$$\log w = \log r + i\theta.$$

В силу (20.32) и (20.33), мы знаем, что  $\log w$  — голоморфная функция, производная которой равна  $1/w$ . Определив логарифм,

<sup>11</sup> Не менее важен и случай вещественного  $\lambda > 0$ . Подынтегральное выражение  $e^{-\lambda x^2/2}$ , деленное на значение интеграла  $\sqrt{2\pi/\lambda}$ , представляет собой плотность *нормального*, или *гауссовского* распределения случайной величины, которое играет основополагающую роль в теории вероятностей и статистике. — *Прим. ред.*

<sup>12</sup> Исключая значение  $w = 0$ , см. выше. — *Прим. ред.*

можно определить степень комплексного числа  $w$  с произвольным комплексным показателем  $c$ , а именно

$$w^c = e^{c \log w}.$$

Как уже упоминалось выше, есть две возможные интерпретации этого выражения. Согласно первой интерпретации, рассматриваем  $w^c$  как *многозначную функцию*  $w$ , т.е. рассматриваем ее не как функцию  $w$ , а как функцию  $z$ . Согласно второй интерпретации, надо выбрать ветвь, например, главную ветвь  $\log w$ . Тогда  $w^c$  становится определенной для всех точек, за исключением  $w$ , лежащих на отрицательной вещественной оси.

## 20.7. Контурные интегралы

Теорема Коши о вычетах применяется для вычисления определенных интегралов. Типичные применения мы показали в параграфе 20.4. В параграфе 20.5 изложен удобный прием вычисления вычетов, в частности, формула (20.25). В этом параграфе мы дадим еще несколько показательных примеров, причем некоторые из них мы используем позднее. Во всех случаях требуется определенная изобретательность при выборе контура интегрирования. Итак, мы должны вычислить определенный интеграл. Обычно это значит, что надо вычислить интеграл вдоль кривой в комплексной плоскости. К этой кривой добавляются другие, чтобы образовать границу некой области. А затем применяется теорема Коши о вычетах. Куски контура выбираются так, чтобы их вклад в интеграл был известен, пропорционален искомому интегралу, или же бесконечно мал при переходе к соответствующему пределу.

(а) Пусть  $R(x) = P(x)/Q(x)$  — рациональная функция, т.е.  $P$  и  $Q$  — полиномы. Предположим, что  $Q$  не имеет вещественных нулей. Если степень  $Q$  больше степени  $P$  по крайней мере на два, то интеграл  $\int_{-\infty}^{\infty} R(x) dx$  сходится и равен пределу  $\int_{-r}^r R(x) dx$  при  $r \rightarrow \infty$ . Мы считаем, что берется интеграл от комплексной функции  $R(z)$  вдоль отрезка прямой от точки  $-r$  до точки  $r$ , расположенных на вещественной оси. Соединим эти точки

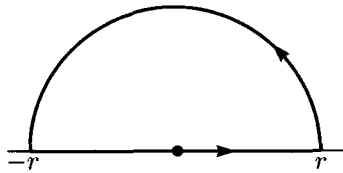


Рис. 20.12

полуокружностью радиуса  $r$ , лежащей в верхней полуплоскости (рис. 20.12).

Теперь у нас имеется замкнутая траектория, ограничивающая полукруг в верхней полуплоскости, и можно применить теорему Коши. С другой стороны, для точек  $z$  на полуокружности  $|R(z)| \leq Kr^{-2}$ , где  $K$  — некоторая константа. Поскольку полуокружность имеет длину  $\pi r$ , то интеграл по ней стремится к нулю, когда  $r \rightarrow \infty$ . Таким образом, получаем

$$\int_{-\infty}^{\infty} R(x) dx = 2\pi i \sum_{\text{Im } z > 0} \text{res } R(z),$$

где сумма берется по всем полюсам в верхней полуплоскости.

(b) Рассмотрим интеграл  $\int_{-\infty}^{\infty} e^{ix} R(x) dx$ , где  $R = P/Q$  — рациональная функция, не имеющая полюсов на вещественной оси. Если  $\deg Q \geq \deg P + 2$ , то можно действовать, как и в предыдущем случае. Если  $\deg Q \geq \deg P + 1$ , то нам известно только, что  $R(x)$  ведет себя на бесконечности как  $1/x$ . Однако в этом случае еще можно доказать, что интеграл сходится. Как и в параграфе 20.6, можно интегрировать по частям. Поскольку  $e^{ix} = -i(d/dx)e^{ix}$ , мы имеем

$$\int_r^s e^{ix} R(x) dx = -ie^{is} R(s) + ie^{ir} R(r) + \int_r^s e^{ix} R'(x) dx,$$

причем функция  $R'$  ведет себя на бесконечности как  $1/x^2$ . Чтобы вычислить этот интеграл, мы опять возьмем предел интеграла от  $-r$  до  $r$ , соединив эти точки полуокружностью в верхней полуплоскости. Но здесь, оценивая интеграл по полуокружности, надо быть более осторожными. Если  $z = re^{i\theta} = r(\cos \theta + i \sin \theta)$ , то

$$e^{iz} = e^{ir \cos \theta} e^{-r \sin \theta},$$

и получается оценка интеграла по полуокружности

$$\begin{aligned} \left| \int_{\text{по дуге}} e^{iz} R(z) dz \right| &= \left| \int_0^\pi e^{ir \cos \theta} e^{-r \sin \theta} R(re^{i\theta}) r d\theta \right| \\ &\leq \int_0^\pi e^{-r \sin \theta} r |R(re^{i\theta})| d\theta. \end{aligned}$$

Теперь функция  $|rR(re^{i\theta})|$  ограничена константой, скажем,  $K$ . Поэтому нам надо оценить интеграл

$$\int_0^\pi e^{-r \sin \theta} d\theta = 2 \int_0^{\pi/2} e^{-r \sin \theta} d\theta.$$

Здесь мы воспользуемся неравенством  $\sin \theta \geq 2\theta/\pi$  для  $0 \leq \theta \leq \pi/2$ . (Действительно, обе части равны в конечных точках, и их разность имеет один максимум внутри.) Поэтому

$$\int_0^{\pi/2} e^{-r \sin \theta} d\theta \leq \int_0^{\pi/2} e^{-2r\theta/\pi} d\theta = (\pi/2r)(1 - e^{-r}) \rightarrow 0.$$

Таким образом, интеграл по полуокружности стремится к нулю, и мы получаем

$$\int_{-\infty}^{\infty} e^{ix} R(x) dx = 2\pi i \sum_{\text{Im } z > 0} \text{res } R(z),$$

где сумма берется по всем полюсам в верхней полуплоскости. Если  $Q$  имеет нуль на вещественной оси, то рассмотренный выше интеграл, в принципе, не имеет смысла. Но тем не менее, иногда он может иметь смысл, если рассматривать его как предел особого рода, называемый *главным значением Коши*. Покажем это на примере. Предположим, что функция  $R$  имеет простой полюс при  $z = 0$  с вычетом  $A$ , так что  $R(z) = A/z + R_0(z)$ , где функция  $R_0$  голоморфна в окрестности нуля. Рассмотрим тот же интеграл, что и раньше, но теперь в контур вставим полуокружность радиуса  $\varepsilon$  в нижней полуплоскости, чтобы обойти начало координат (рис. 20.13). Теперь охватываемая область содержит все полюса в верхней полуплоскости плюс полюс в начале координат. Интеграл от  $A/z$  по малой полуокружности равен  $\pi i A$ , сокращая

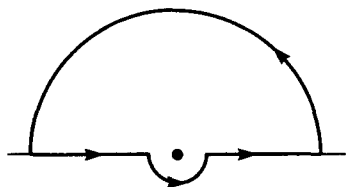


Рис. 20.13

половину вычета от точки  $z = 0$ . Тогда

$$P\left(\int_{-\infty}^{\infty} e^{ix} R(x) dx\right) = \lim_{\varepsilon \rightarrow 0} \left[ \int_{-\infty}^{-\varepsilon} + \int_{\varepsilon}^{\infty} \right] e^{ix} R(x) dx \\ = i\pi A + 2\pi i \sum_{\operatorname{Im} z > 0} \operatorname{res}(e^{iz} R(z)).$$

Здесь  $P(\cdot)$  означает главное значение Коши. Например, пусть  $R(x) = 1/x$ , тогда

$$P\left(\int_{-\infty}^{\infty} \frac{e^{ix}}{x} dx\right) = \pi i.$$

Возьмем вещественную часть этого равенства и получим, что  $P(\int_{-\infty}^{\infty} (\cos x/x) dx) = 0$ . Результат очевиден, поскольку  $\cos x/x$  — нечетная функция. Если взять мнимую часть равенства, то получаем  $P(\int_{-\infty}^{\infty} (\sin x/x) dx) = \pi$ . Функция  $\sin x/x$  не имеет особенностей в начале координат, поэтому нет нужды брать главное значение этого интеграла. Функция  $\sin x/x$  четная, поэтому

$$\int_0^{\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}.$$

Мы еще используем этот интеграл в теории рядов Фурье.

(с) Рассмотрим интегралы вида  $\int_0^{\infty} x^{\alpha} R(x) dx$ , где  $\alpha$  — вещественное число, причем  $0 < \alpha < 1$ ,  $R$  — рациональная функция, которая на бесконечности убывает как  $1/x^2$  и в худшем случае имеет полюс первого порядка в нуле. Трюк состоит в том, что делается замена переменной:  $x = t^2$ , так что теперь интеграл имеет вид  $2 \int_0^{\infty} t^{2\alpha+1} R(t^2) dt$ . Для его вычисления удобно выбрать ветвь функции  $z^{2\alpha}$ , которая получается, если плоскость  $z$  разрезать

вдоль отрицательной мнимой оси. Тогда можно написать  $z = re^{i\theta}$ , где  $-\pi/2 < \theta < 3\pi/2$ , а  $z^{2\alpha} = r^{2\alpha}e^{i2\alpha\theta}$ , где  $-\pi\alpha < 2\alpha\theta < 3\pi\alpha$ . Теперь мы должны выбрать контур так, чтобы он не проходил через отрицательную мнимую ось. Пусть контур состоит из двух отрезков прямой вдоль отрицательной и положительной вещественной оси, а также большой и малой полуокружностей, лежащих в верхней полуплоскости (рис. 20.14). Мы считаем, что интегралы по полуокружностям стремятся к нулю, когда радиус большой дуги стремится к бесконечности, а радиус малой — к нулю. Тогда

$$2\pi i \sum \text{res} = \int_{-\infty}^{\infty} z^{2\alpha+1} R(z^2) dz = \int_0^{\infty} (z^{2\alpha+1} + (-z)^{2\alpha+1}) R(z^2) dz.$$

Мы выбрали ветвь  $z^{2\alpha}$  так, что  $(-z)^{2\alpha} = e^{2\pi i\alpha} z^{2\alpha}$  для положительных значений  $z$ . Поэтому последний интеграл имеет вид

$$(1 - e^{2\pi i\alpha}) \int_0^{\infty} t^{2\alpha+1} R(t^2) dt.$$

Поскольку  $1 - e^{2\pi i\alpha} \neq 0$ , на этот коэффициент можно разделить и получить значение нашего интеграла. Слева мы должны просуммировать вычеты функции  $z^{2\alpha+1} R(z^2)$  в верхней полуплоскости. Они совпадают с вычетами функции  $z^\alpha R(z)$  во всей плоскости, разрезанной вдоль положительной вещественной оси. Поэтому другой способ решить нашу задачу состоит в том, что не надо делать замену  $x = t^2$ , а просто разрезать плоскость  $z$  вдоль положительной вещественной оси и использовать ветвь функции  $z^\alpha$ , определенную условием  $z = re^{i\theta}$ , где  $0 < \theta < 2\pi$ . Заметим, что

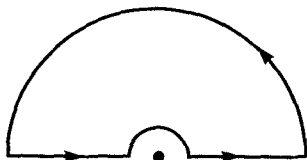


Рис. 20.14

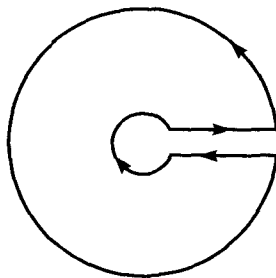


Рис. 20.15

мы разрежали плоскость точно вдоль пути исходного интеграла! Поэтому мы не можем получить наш интеграл как часть интеграла по замкнутому контуру. Вместо этого мы интегрируем вдоль пути, параллельного положительной вещественной оси, но слегка сдвинутого в верхнюю полуплоскость, имея в виду переход к пределу для получения нашего интеграла. Итак, рассмотрим замкнутый контур, состоящий из двух прямолинейных отрезков, один чуть выше положительной вещественной оси, а другой — чуть ниже, а также из большой и малой окружностей, взятых почти целиком (рис. 20.15). Вклад обеих окружностей в пределе обращается в нуль. Вклады от двух прямолинейных отрезков не сокращают друг друга, потому что определение функции  $z^\alpha$  выше и ниже вещественной положительной оси отличаются на множитель  $e^{2\pi i\alpha}$ . Таким образом, именно неопределенность функции  $z^\alpha$  в нуле позволяет нам вычислить интеграл.

(d) Интегрирование по контуру можно использовать для вычисления сумм бесконечных рядов. Идея состоит в том, чтобы применять теорему Коши о вычетах ко все увеличивающимся областям, чтобы интегралы по границе области стремились к нулю. Тогда останется (бесконечная) сумма вычетов, которая тоже должна давать нуль. Перенесем одно из слагаемых в другую часть уравнения, получим сумму бесконечного ряда. Проиллюстрируем этот метод на примере сумм, полезных при изучении тригонометрических функций. С этой целью сначала определим тригонометрические функции комплексной переменной:

$$\cos z = \frac{1}{2}(e^{iz} + e^{-iz}), \quad \sin z = -\frac{i}{2}(e^{iz} - e^{-iz}).$$

Очевидно, что эти функции голоморфны во всей комплексной плоскости и для вещественных  $z$  совпадают с тригонометрическими функциями вещественной переменной. Заметим, что  $\sin z = 0$ , если и только если  $e^{2iz} = 1$ . Это выполняется только для вещественных значений  $z = n\pi$ . Аналогично, комплексные нули  $\cos z$  — только в вещественных точках  $(n + 1/2)\pi$ . Определим

$$\tan z = \frac{\sin z}{\cos z}, \quad \cot z = \frac{\cos z}{\sin z}.$$

Эти функции мероморфны и имеют полюса, расположенные в нулях их знаменателей. Можно воспользоваться (20.25) и вычислить вычеты в этих полюсах, потому что все нули простые. Например,  $\cot z$  имеет полюса в точках  $n\pi$ , и все вычеты в них равны 1. Заметим, что функция

$$\cot z = i \frac{e^{iz} + e^{-iz}}{e^{iz} - e^{-iz}}$$

будет ограничена по абсолютному значению до тех пор, пока  $z$  находится на конечном расстоянии от точек  $n\pi$ . В частности  $|\cot z|$  будет равномерно ограничен на всех окружностях радиуса  $(n + \frac{1}{2})\pi$  с центром в начале координат. Рассмотрим интеграл по этим окружностям

$$\int_C \frac{\cot \xi}{\xi^2 - z^2} d\xi,$$

где  $z \neq n\pi$  — фиксированное комплексное число. Если радиус окружности устремить к бесконечности, то интеграл будет стремиться к нулю, потому что  $|\cot \xi|$  ограничен, а в знаменателе стоит полином второй степени по  $\xi$ . Полюса подынтегрального выражения находятся в точках  $\pm z$  и  $n\pi$ . Вычеты в точках  $\pm z$  равны  $\frac{1}{2} \cot z/z$ , а в точках  $n\pi$  они составляют  $(n^2\pi^2 - z^2)^{-1}$ . Поэтому

$$\frac{\cot z}{z} + \sum_1^{\infty} (n^2\pi^2 - z^2)^{-1} = 0.$$

Первый член перенесем вправо, умножим это равенство на  $z$  и из суммы выделим член с  $n = 0$ . В результате получим формулу

$$\cot z = \frac{1}{z} + 2z \sum_1^{\infty} \frac{1}{z^2 - n^2\pi^2}. \quad (20.36)$$

Ряд справа сходится, причем сходится равномерно при всех  $z$ , находящихся на конечном расстоянии от всех полюсов. Если  $f(z) = \sin z$ , то  $f'(z)/f(z) = \cot z$ . Можно также написать  $f'(z)/f(z) = (d/dz)(\log f(z))$ . После этого (20.36) переписываем в форме

$$\frac{d}{dz} \left( \log \frac{\sin z}{z} \right) = \sum_1^{\infty} \frac{d}{dz} \log(z^2 - n^2\pi^2).$$

Функция  $\sin z/z$  не равна нулю в пределе при  $z = 0$ . Проинтегрируем это уравнение по контуру, соединяющему 0 с точкой  $z$  и проходящему на конечном расстоянии от всех точек  $n\pi$ . Получим уравнение

$$\log(\sin z/z) - \log 1 = \sum_1^{\infty} (\log(z^2 - n^2\pi^2) - \log(-n^2\pi^2)).$$

Это уравнение содержит неоднозначности, потому что значения логарифма, определенные правилом, что  $\log g$  равняется интегралу от  $g'/g$ , зависит от выбранного контура интегрирования. Однако если вычислить экспоненту обеих частей, то все неопределенности исчезают, и мы получаем известную формулу

$$\sin z = z \prod_1^{\infty} \left(1 - \frac{z^2}{n^2\pi^2}\right),$$

представляющую  $\sin z$  как бесконечное произведение.

## 20.8. Пределы и ряды

Начнем с применения интегральной формулы Коши.

Пусть  $f_n$  — последовательность функций, голоморфных в области  $D$ . Предположим, что  $f_n \rightarrow f(z)$  равномерно в  $D$ . Тогда  $f$  — голоморфная функция в области  $D$ , и  $f'_n(z)$  сходится к  $f'(z)$  равномерно на любом подмножестве  $D$ , расположенном на конечном расстоянии от границы  $\partial D$ . (20.37)

*Доказательство.* Для любой точки  $z \in D$  выберем окружность  $C$ , диск которой лежит внутри  $D$  (а точка  $z$  внутри  $C$ ). Тогда

$$f_n(z) = \frac{1}{2\pi i} \int_C \frac{f_n(\xi)}{\xi - z} d\xi.$$

Переход к пределу дает

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(\xi)}{\xi - z} d\xi.$$

Из обсуждения, следующего после (20.18), мы знаем, что это означает голоморфность функции  $f$ . Из формулы (20.19) имеем

$$f'_n(z) = \frac{1}{2\pi i} \int_C \frac{f_n(\xi)}{(\xi - z)^2} d\xi,$$

которое, очевидно, равномерно сходится к

$$f'(z) = \frac{1}{2\pi i} \int_C \frac{f(\xi)}{(\xi - z)^2} d\xi.$$

Часто возникает желание применить утверждение (20.37) в случае, когда область  $D$  является ограниченной подобластью некоторой большей области, где последовательность  $f_n$  определена и сходится, т.е. функции  $f_n$  определены и голоморфны в большой области  $E$  и равномерно сходятся на каждой из подобластей  $D_k \subset E$ , образующих последовательность с  $\bigcup D_k = E$ . Применяя утверждение (20.37) к каждой области  $D_k$ , приходим к выводу, что  $f$  голоморфна во всей области  $E$ .

Применим утверждение (20.37) к частичным суммам и рядам.

Если ряд из голоморфных слагаемых  $f(z) = f_1(z) + f_2(z) + \dots$  равномерно сходится в области  $D$ , то сумма  $f$  — голоморфная функция, и ряд можно дифференцировать почленно. (20.38)

Наиболее важный ряд, связанный с голоморфными функциями, — ряд Тейлора. Предположим, что  $f$  голоморфна в области  $D$ , и диск радиуса  $R$  с центром в точке  $a$  целиком лежит в области  $D$ . Пусть  $z$  — точка внутри этого диска, причем  $|z - a| < R$ . Тогда можно воспользоваться (20.22) и оценить последний член в формуле (20.20). Его абсолютное значение не превышает

$$\frac{M|z - a|}{R - |z - a|} \frac{|z - a|^{n-1}}{R^{n-1}}$$

и равномерно стремится к нулю на любом диске меньшего радиуса. Таким образом, из утверждения (20.38) получаем следующее.

Если  $f$  голоморфна в области  $D$  и  $a$  — произвольная точка в  $D$ , то ряд

$$f(a) + f'(a)(z-a) + \frac{f''(a)}{2!}(z-a)^2 + \dots + \frac{f^{(n)}(a)}{n!}(z-a)^n + \dots \quad (20.39)$$

сходится к  $f(z)$  на самом большом открытом диске с центром в точке  $a$ , целиком лежащем в области  $D$ .

Этот ряд называется *рядом Тейлора* функции  $f$  в окрестности точки  $a$ . Многие операции с голоморфными функциями соответствуют простым операциям с их рядом Тейлора. Для упрощения записи положим  $a = 0$ . Тогда если

$$f(z) = a_0 + a_1 z + a_2 z^2 + \dots$$

и

$$g(z) = b_0 + b_1 z + b_2 z^2 + \dots,$$

то получаем

$$f'(z) = a_1 + 2a_2 z + 3a_3 z^2 + \dots,$$

$$f(z) + g(z) = (a_0 + b_0) + (a_1 + b_1)z + (a_2 + b_2)z^2 + \dots$$

$$f(z)g(z) = a_0 b_0 + (a_1 b_0 + a_0 b_1)z + (a_2 b_0 + a_1 b_1 + a_0 b_2)z^2 + \dots.$$

Если  $a_0 = b_0 = 0$ , то

$$\begin{aligned} f(g(z)) &= a_1(b_1 z + b_2 z^2 + \dots) + a_2(b_1 z + b_2 z^2 + \dots)^2 + \dots \\ &= a_1 b_1 z + (a_1 b_2 + a_2 b_1^2)z^2 + (a_1 b_3 + 2a_2 b_1 b_2 + a_3 b_1^3)z^3 + \dots. \end{aligned}$$

Если мы знаем степенной ряд функции  $f$ , то из этого последнего уравнения можно рекурсивно определить степенной ряд функции, обратной к  $f$ . Пусть мы знаем значения  $a_i$  и хотим найти такие  $b_i$ , чтобы  $f(g(z)) = z$ . (Чтобы обратить функцию, необходимо предположить, что  $f'(0) \neq 0$ .) Тогда

$$\begin{aligned} a_1 b_1 &= 1, & b_1 &= a_1^{-1}, \\ a_1 b_2 + a_2 b_1^2 &= 0, & b_2 &= -a_1^{-1}(a_2 b_1^2), \\ a_1 b_3 + 2a_2 b_1 b_2 + a_3 b_1^3 &= 0, & b_3 &= -a_1^{-1}(2a_2 b_1 b_2 + a_3 b_1^3) \end{aligned}$$

и т. д. Вообще коэффициенты при  $z^n$  будут выражениями типа  $a_1 b_n +$  (члены, содержащие  $b_i$  порядка, меньше, чем  $n$ ). Поэтому можно последовательно находить каждое  $b_n$ .

Известен элементарный факт, что любой степенной ряд имеет радиус сходимости  $R$  (возможно, и 0); при этом ряд сходится для  $|z| < R$  и не сходится для  $|z| > R$ . Из (20.38) следует, что ряд представляет голоморфную функцию внутри радиуса сходимости. И обратно, из (20.39) следует, что ряд Тейлора для голоморфной функции в качестве радиуса сходимости имеет радиус самого большого диска, содержащегося в истинной области определения<sup>13</sup> функции  $f$ .

Ряд вида  $a_0 + a_{-1}z^{-1} + a_{-2}z^{-2} + \dots$  можно рассматривать как степенной ряд по переменной  $1/z$ . Поэтому он будет сходиться вне окружности радиуса  $R$  (возможно, что  $R = \infty$ ). Говорят, что ряд вида  $\sum_{-\infty}^{\infty} a_n z^n$  сходится, если и только если сходятся его положительные и отрицательные части. Положительная часть будет сходиться при  $|z| < R_2$  для некоторого  $R_2$ , и отрицательная часть будет сходиться при  $|z| > R_1$  для некоторого  $R_1$ . Поэтому для двойного ряда будет непустая область сходимости тогда и только тогда, когда  $R_1 < R_2$ . В этом случае, в силу (20.38), он будет представлять голоморфную функцию в кольце  $R_1 < |z| < R_2$ . И обратно, предположим, что  $f$  голоморфна в таком кольце. Покажем, что в этом случае функция  $f$  представима в виде двойного ряда. Для доказательства мы должны только показать, что можно разложить  $f$  в сумму  $f = f_1 + f_2$ , где  $f_1$  голоморфна при  $|z| < R_2$  и  $f_2$  голоморфна при  $|z| > R_1$ . Причем  $f_2(1/z)$  должна иметь устранимую особенность в нуле, так что  $f_2(1/z)$  голоморфна для  $|1/z| < 1/R_1$ . Чтобы это сделать, определим

$$f_1(z) = \frac{1}{2\pi i} \int_{|\xi|=r_2} \frac{f(\xi)}{\xi - z} d\xi,$$

где интеграл берется по любой окружности радиуса  $r_2$  с условием  $|z| < r_2 < R_2$ ; и

$$f_2(z) = -\frac{1}{2\pi i} \int_{|\xi|=r_1} \frac{f(\xi)}{\xi - z} d\xi,$$

<sup>13</sup>т. е. области, за пределы которой функцию невозможно продолжить из-за полюсов или существенных особенностей на границе. — Прим. ред.

где интеграл берется по любой окружности радиуса  $r_1$ , причем  $R_1 < r_1 < |z|$ . Очевидно, что  $f_1$  голоморфна при  $|z| < R_2$  и  $f_2$  голоморфна при  $|z| > R_1$ . Из интегральной формулы Коши следует, что  $f_1(z) + f_2(z) = f(z)$ . Если положить  $z' = 1/z$  и подставить  $\xi' = 1/\xi$  в интеграл для  $f_2$ , то получим

$$f_2(1/z') = \frac{1}{2\pi i} \int_{|\xi'|=1/r_1} \frac{f(1/\xi')}{\xi' - z'} \frac{z'}{\xi'} d\xi'.$$

Это показывает, что  $f_2(1/z')$  голоморфна по  $z'$  для  $|z'| < 1/R_1$ . Функцию  $f_1$  можно разложить в степенной ряд

$$f_1(z) = \sum_0^{\infty} a_n z^n, \quad \text{где} \quad a_n = \frac{1}{2\pi i} \int_{|\xi|=r} \frac{f(\xi)}{\xi^{n+1}} d\xi,$$

а функцию  $f_2(1/z')$  — в ряд

$$f_2(1/z') = \sum_1^{\infty} b_n z'^n,$$

$$\text{где} \quad b_n = \frac{1}{2\pi i} \int_{|\xi'|=1/r} \frac{f(1/\xi')}{\xi'^{n+1}} d\xi' = \frac{1}{2\pi i} \int_{|\xi|=r} f(\xi) \xi^{n-1} d\xi;$$

здесь  $r$  — любое значение между  $R_1$  и  $R_2$ . Полагая  $a_{-n} = b_n$ , получаем искомое разложение, называемое *рядом Лорана*

$$f(z) = \sum_{-\infty}^{\infty} a_n z^n, \quad \text{где} \quad a_n = \frac{1}{2\pi i} \int_{|\xi|=r} \frac{f(\xi)}{\xi^{n+1}} d\xi. \quad (20.40)$$

В дальнейшем нам понадобится явный вид этой формулы в случае  $R_1 < 1 < R_2$ . Итак, мы предполагаем, что функция  $f$  определена и голоморфна в некоторой окрестности единичного круга, т. е.  $|z| = 1$ . Пусть  $F(\theta) = f(e^{i\theta})$  и в формуле (20.40)  $r = 1$ . Тогда получаем

$$f(z) = \sum a_n z^n, \quad \text{где} \quad a_n = \frac{1}{2\pi} \int_0^{2\pi} F(\theta) e^{in\theta} d\theta. \quad (20.41)$$

В частности, взяв  $z = e^{i\theta}$  (на единичной окружности) и подставив это в ряд Лорана для  $f$ , имеем

$$F(\theta) = \sum_{-\infty}^{\infty} a_n e^{in\theta}, \quad (20.42)$$

где коэффициенты  $a_n$  определены формулой (20.41). Ряд (20.42) называется *рядом Фурье* функции  $F$ . Если  $f$  голоморфна в области  $R < |z| < 1/R$ , где  $R$  — некоторое число, причем  $R < 1$ , то, поскольку  $a_n$  — коэффициенты сходящегося степенного ряда, из критерия сходимости степенных рядов получаем, что

$$|a_n| \leq cr^{-|n|} \quad \text{для любых } r > R$$

и соответствующей константы  $c = c_r$ .

## Добавление: связь между голоморфными и гармоническими функциями

Пусть  $f = u + iv$  — функция, голоморфная в области  $D$ . Тогда, как доказано в параграфе 20.4, она бесконечно дифференцируема, т. е. имеет в  $D$  производные любого порядка. Для первых частных производных выполняются уравнения Коши–Римана

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \quad (20.11)$$

Продифференцируем первое уравнение по  $x$ , второе по  $y$  и сложим. В правой части смешанные частные производные второго порядка от  $v$  сокращаются, и мы имеем

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

Итак, функция  $u$  удовлетворяет двумерному уравнению Лапласа

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

и является гармонической. Аналогично, гармонической функцией является мнимая часть  $v$  (чтобы убедиться в этом, можно снова продифференцировать уравнения Коши–Римана или заменить функцию  $f$  на  $-if$ ).

Таким образом,

вещественная и мнимая часть всякой голоморфной функции, рассматриваемые как функции двух вещественных переменных, являются гармоническими.

И обратно, каждая гармоническая функция  $u$  в односвязной (не имеющей «дырок») плоской области может быть реализована как вещественная часть голоморфной функции, причем последняя определена однозначно с точностью до аддитивной (чисто мнимой) постоянной. Иначе говоря, существует единственная по модулю константы функция  $v$ , которая связана с  $u$  уравнениями Коши–Римана. Функция  $v$  — также гармоническая и называется *сопряженной* к  $u$ .

*Доказательство.* Применим звездный оператор  $\star$  к дифференциалу функции  $u(x, y)$ :

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy, \quad \star du = -\frac{\partial u}{\partial y} dx + \frac{\partial u}{\partial x} dy.$$

Полученная 1-форма  $\star du$  является замкнутой, так как

$$d\star du = d\left(-\frac{\partial u}{\partial y} dx + \frac{\partial u}{\partial x} dy\right) = \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) dx \wedge dy = 0$$

в силу гармоничности  $u$ . Но тогда, как показано в параграфе 7.3, эта форма является точной, если рассматривать ее в звездной области. На самом деле область достаточно предполагать односвязной. Следовательно, форма  $\star du$  является дифференциалом некоторой функции  $v$ :

$$\star du = dv = \frac{\partial v}{\partial x} dx + \frac{\partial v}{\partial y} dy.$$

При этом функция  $v$  определена однозначно с точностью до аддитивной константы. Равенства коэффициентов при  $dx$  и  $dy$  в

выражениях для  $\star du$  суть уравнения Коши–Римана. Тем самым  $f = u + iv$  является искомой голоморфной функцией.

Отсюда легко вывести, что: 1) если  $u$  — гармоническая функция в области  $D$ , а функция  $g$  голоморфно отображает область  $G$  в область  $D$ , то  $u \circ g$  является гармонической функцией в  $G$ ; 2) всякая гармоническая функция является бесконечно дифференцируемой.

Кстати, последнее утверждение справедливо для функций любого числа переменных. Всякая гармоническая функция имеет внутри области определения частные производные всех порядков, которые, в свою очередь, также являются гармоническими функциями.

Отметим, что предположение об односвязности области существенно. Гармоническая функция  $u = \log r = \frac{1}{2} \log(x^2 + y^2)$  определена на всей плоскости, за исключением начала координат 0. Она является вещественной частью комплексной логарифмической функции  $\log z$ , которую можно определить однозначно только в односвязной подобласти области определения.

**Вывод теоремы о среднем.** Пусть функция  $f$  голоморфна в круге  $D$  радиуса  $R$  с центром в точке  $z = c$  и непрерывна на  $\overline{D}$ . Согласно интегральной формуле Коши, имеем

$$f(c) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(\xi) d\xi}{\xi - c}.$$

Замена переменной  $\xi = c + Re^{i\theta}$  на окружности  $\partial D$  дает

$$\frac{d\xi}{\xi - c} = i d\theta = \frac{1}{R} ds,$$

где  $ds$  — элемент длины дуги. Таким образом,

$$f(c) = \frac{1}{2\pi} \int_{\partial D} f(z) d\theta = \frac{1}{2\pi R} \int_{\partial D} f(z) ds.$$

Итак, значение  $f$  в центре круга совпадает со средним значением по окружности. Переходя к вещественной части, имеем двумерную теорему о среднем для гармонических функций.

**Решение задачи Дирихле для круга.** Пусть  $D$  — круг радиуса  $R$  с центром в точке  $c$ , и  $a = a(z)$  — непрерывная функция на окружности  $\partial D$ . Тогда решение задачи Дирихле

$$\Delta u = 0 \quad \text{внутри } D, \quad u = a \quad \text{на } \partial D$$

существует и дается *интегральной формулой Пуассона*

$$u(z) = \frac{1}{2\pi} \int_{\partial D} \frac{a(z)(R^2 - r^2)}{R^2 + r^2 - 2Rr \cos(\theta - \phi)} d\theta, \quad \text{где } z = c + re^{i\phi}.$$

Мы только кратко опишем идею вывода формулы в предположении существования решения. Чтобы вычислить  $u(z_0)$ , где  $z_0 = c + re^{i\phi}$ ,  $r < R$ , применим дробно-линейное преобразование

$$w = g(z) = \frac{R(z - z_0)}{R^2 - (\bar{z}_0 - \bar{c})(z - c)},$$

которое голоморфно и взаимно-однозначно отображает круг  $\bar{D}$  на единичный круг  $\bar{D}_1 = \{|w| \leq 1\}$ , причем точка  $z_0$  переходит в 0. Далее, пусть  $v$  — решение задачи Дирихле на единичном круге:

$$\Delta v(w) = 0 \quad \text{при } |w| < 1, \quad v(w) = a(g^{-1}(w)) \quad \text{при } |w| = 1.$$

Легко видеть, что  $u(z) = v(g(z))$  является решением исходной задачи Дирихле. При этом  $u(z_0) = v(g(z_0)) = v(0)$  вычисляется как среднее значение  $a(g^{-1}(w))$  по окружности  $\partial D_1 = \{|w| = 1\}$ . Детали доказательства оставляем читателю.

Описанный метод решения задачи Дирихле переносится на достаточно широкий класс областей, являющихся конформными образами круга.

## Резюме

### А. Голоморфные функции

Вы должны знать определение голоморфной функции через 2-формы и уметь показывать, что ее матрица Якоби представляет конформное преобразование.

Вы должны уметь выводить интегральную теорему Коши и формулу для вычетов из теоремы Стокса.

Вы должны уметь вычислять вычет функции в простом или в кратном полюсе.

## В. Интегрирование по контуру

Вы должны уметь вычислять интегралы с помощью теоремы о вычетах для выбранного контура. Для некоторых типичных интегралов нужно уметь находить соответствующий контур.

## С. Степенные ряды

Вы должны уметь раскладывать голоморфную функцию в ряд Тейлора или Лорана и уметь вычислять радиус сходимости этих рядов.

## Задачи

- 20.1. (а) Предположим, что  $f(z)$  — голоморфная функция. Пусть  $z = x + iy$ ,  $f(z) = u + iv$ . Покажите, что семейства кривых  $u(x, y) = \text{const}$  и  $v(x, y) = \text{const}$  ортогональны, т. е. что они пересекаются под прямым углом.<sup>14</sup>
- (б) Проверьте это свойство для функции  $f(z) = z^2$ . Нарисуйте несколько кривых. Сделайте то же самое для функции  $f(z) = 1/z$ . Что в этом случае происходит в нуле?
- 20.2. (а) Предположим, что  $f(z)$  — голоморфная функция, вещественная часть которой равна  $3x^2y - y^3$ . Получите мнимую часть  $f(z)$ .
- (б) Может ли функция быть голоморфной, если ее вещественная часть равна  $xy^2$ ? Если да, то получите ее, а если нет, то объясните, почему.
- 20.3. Голоморфные функции  $\cos z$  и  $\sin z$  можно определить тождествами

$$\cos z = \frac{1}{2}(e^{iz} + e^{-iz}), \quad \sin z = \frac{1}{2i}(e^{iz} - e^{-iz}),$$

используя определение комплексной экспоненциальной функции.

<sup>14</sup>Это означает, что ортогональны векторы касательных (и нормалей) к кривым  $u = \text{const}$  и  $v = \text{const}$  в каждой точке  $(x_0, y_0)$ . Если об этом помнить, то доказательство данного пункта не составит труда. — *Прим. ред.*

- (a) Выразите  $\cos z$  и  $\sin z$  через тригонометрические и гиперболические функции  $x$  и  $y$ . Используя эти выражения, вычислите  $d(\cos z)$  и  $d(\sin z)$ .
- (b) Докажите, что формулы сложения  $\sin(z+w) = \sin z \cos w + \cos z \sin w$  и  $\cos(z+w) = \cos z \cos w - \sin z \sin w$  выполняются для комплексных  $z$  и  $w$ .
- (c) Используя формулы сложения, покажите явно, что функции  $\sin z$  и  $\cos z$  непрерывно дифференцируемы в комплексном смысле. Получите производные этих функций, вычисляя соответствующие пределы.

20.4. С помощью метода, изложенного в параграфе 20.7, получите значения определенных интегралов:

(a)  $\int_0^\pi \sin^6 \theta \, d\theta = \frac{5\pi}{16};$

(b)  $\int_0^\pi \frac{\cos \theta \, d\theta}{1 - 2a \cos \theta + a^2} = \frac{\pi a}{1 - a^2}, \quad a^2 < 1.$

20.5. С помощью метода, изложенного в параграфе 20.7, пункт (a), проверьте формулы для определенных интегралов:

(a)  $\int_{-\infty}^{\infty} \frac{dx}{(x+b)^2 + a^2} = \frac{\pi}{a}, \quad a > 0;$

(b)  $\int_0^\infty \frac{dx}{(x^4 + 4a^4)} = \frac{\pi}{8a^3}, \quad a > 0;$

(c)  $\int_0^\infty \frac{x^2 \, dx}{x^6 + 1} = \frac{\pi}{6}.$

20.6. (a) Предположим, что  $f(z)$  голоморфна в некоторой области за исключением точки  $z = a$ , где она имеет полюс *второго* порядка. Докажите, что вычет в этом полюсе задается формулой

$$\text{res} = \lim_{z \rightarrow a} \frac{d}{dz} ((z-a)^2 f(z)).$$

(b) С помощью этой формулы вычислите интеграл

$$\int_{-\infty}^{\infty} \frac{dx}{(x^2 + a^2)^2}.$$

(с) Обобщив этот метод, вычислите интеграл

$$\int_{-\infty}^{\infty} \frac{dx}{(x^2 + a^2)^k}$$

для произвольного целого положительного числа  $k$ .

- 20.7. (а) Здесь излагается метод, альтернативный методу, изложенному в параграфе 20.7, пункт (b). Пусть  $R(z) = P(z)/Q(z)$ , где  $P$  и  $Q$  — полиномы, степени которых связаны условием  $\deg Q \geq \deg P + 1$ , так что

$$|R(z)| \leq A/|z|$$

для некоторой константы  $A$  и достаточно больших  $|z|$ . Предположим, что  $Q$  не имеет вещественных нулей. Следуя изложенному ниже плану, докажите, что интеграл  $\int_{-\infty}^{\infty} R(x)e^{ix} dx$  сходится, т. е. предел

$$\lim_{\substack{X_2 \rightarrow \infty \\ X_1 \rightarrow \infty}} \int_{-X_1}^{X_2} R(x)e^{ix} dx$$

существует, и выполняется равенство

$$\int_{-\infty}^{\infty} R(x)e^{ix} dx = 2\pi i \sum \operatorname{res}(R(z)e^{iz}),$$

где сумма берется по всем полюсам, лежащим в верхней полуплоскости.

*Шаг 1.* Выберем достаточно большие  $X_1$ ,  $X_2$  и  $Y$ , такие, чтобы прямоугольник, изображенный на рис. 20.16, содержал все полюса в верхней полуплоскости.

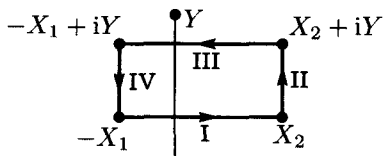


Рис. 20.16

*Шаг 2.* Покажите, что

$$\left| \int_{\Pi} R(z)e^{iz} dz \right| \leq A \int_0^Y \frac{e^{-y}}{|z|} dy \leq \frac{A}{X_2} \int_0^Y e^{-y} dy \leq \frac{A}{X_2}$$

для достаточно больших  $X_2$ .

Шаг 2'. Покажите, что

$$\left| \int_{\text{IV}} R(z) e^{iz} dz \right| \leq \frac{A}{X_1}.$$

Шаг 3. Покажите, что

$$\left| \int_{\text{III}} R(z) e^{iz} dz \right| \leq A e^{-Y} (X_1 + X_2) / Y.$$

Шаг 4. Пусть  $Y \rightarrow \infty$ . Докажите, что

$$\left| \int_{-X_1}^{X_2} R(x) e^{ix} dx - 2\pi i \sum_{z \rightarrow 0} \text{res}(R(z) e^{iz}) \right| < A \left( \frac{1}{X_1} + \frac{1}{X_2} \right)$$

и завершите доказательство.

(b) Вычислите интеграл

$$\int_0^\infty \frac{\cos x}{x^2 + a^2} dx.$$

20.8. Предположим, что  $Q(0) = 0$ , а  $Q'(0) \neq 0$ , так что функция  $R(z)e^{iz}$  имеет простой полюс в 0. Пусть  $B$  — вычет этой функции в 0. Рассмотрите тот же прямоугольный контур, но с обходом начала координат в нижней полуплоскости по малой полуокружности радиуса  $\delta$  (рис. 20.17).

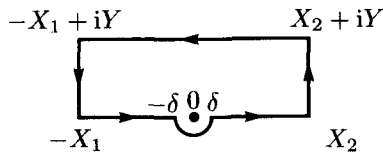


Рис. 20.17

- (a) Покажите, что предел интеграла по полуокружности при  $\delta \rightarrow 0$  равен  $\pi i B$ .
- (b) Докажите, что главное значение Коши определяется формулой

$$\lim_{\delta \rightarrow 0} \left( \left[ \int_{-\infty}^{-\delta} + \int_{\delta}^{\infty} \right] R(x) e^{ix} dx \right) = 2\pi i \sum_{z \rightarrow 0} \text{res}(R(z) e^{iz}) + \frac{1}{2} B.$$

(с) Покажите, что

$$\int_0^{\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}.$$

20.9. Пусть  $f(z) = (f_{ij}(z))$  — матричная функция. Мы говорим, что  $f$  голоморфна в области  $D$ , если каждый матричный элемент  $f_{ij}$  является голоморфной функцией в области  $D$ . Пусть  $\gamma$  — произвольная кривая. Интеграл  $\int_{\gamma} f_{ij}(z) dz$  должен быть матрицей,  $ij$ -й элемент которой равен  $\int_{\gamma} f_{ij}(z) dz$ . Заметим, что если  $B$  и  $C$  — постоянные матрицы и  $f$  — голоморфная матричная функция, то  $BfC$  — тоже голоморфная матричная функция, и для произвольной кривой  $\gamma$  мы имеем

$$\int_{\gamma} Bf(z)C dz = B \left( \int_{\gamma} f(z) dz \right) C.$$

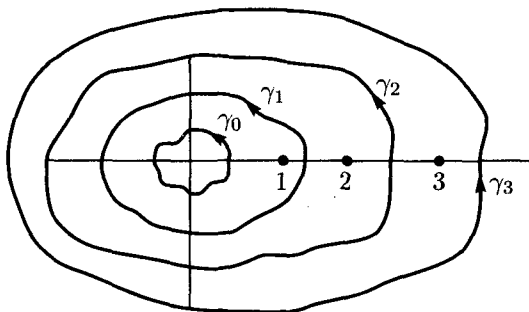


Рис. 20.18

(а) Пусть  $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$  и  $f(z) = (z\mathbb{I} - A)^{-1}$ , где

$$\mathbb{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Покажите, что  $f$  голоморфна везде, кроме точек  $z = 1$ ,  $z = 2$  и  $z = 3$ . Пусть  $P_{\gamma_i} = (1/2\pi i) \int_{\gamma_i} f(z) dz$  (контуры  $\gamma_i$  показаны на рис. 20.18). Вычислите  $P_{\gamma_i}$  для  $i = 0, 1, 2, 3$ .

- (b) Пусть  $D$  — матрица  $3 \times 3$  с собственными значениями 1, 2 и 3. Пусть  $g(z) = (z\mathbb{I} - D)^{-1}$ . Покажите, что  $g$  голоморфна везде, кроме точек  $z = 1, 2, 3$ . Пусть

$$Q_{\gamma_i} = \frac{1}{2\pi i} \int_{\gamma_i} g(z) dz.$$

Опишите  $\text{Im } Q_{\gamma_1}$  и  $\text{Ker } Q_{\gamma_1}$ . Что такое  $Q_{\gamma_1}^2$ ? Опишите  $\text{Im } Q_{\gamma_2}$  и  $\text{Ker } Q_{\gamma_2}$ . Что такое  $Q_{\gamma_2}^2$ ?

- (c) Сформулируйте общую теорему. Пусть  $A$  — матрица  $n \times n$  с различными собственными значениями  $\lambda_1, \dots, \lambda_n$ . Пусть  $f(z) = (z\mathbb{I} - A)^{-1}$ . Где голоморфна функция  $f$ ? Пусть  $\gamma$  — кривая, которая не проходит через какие-либо собственные значения и ограничивает область, содержащую точки  $\lambda_1, \dots, \lambda_k$  и не содержащую ни одно из остальных собственных значений. Положим

$$P_\gamma = \frac{1}{2\pi i} \int_\gamma f(z) dz = \frac{1}{2\pi i} \int (z\mathbb{I} - A)^{-1} dz.$$

Что такое  $\text{Im } P_\gamma$ ? Что такое  $\text{Ker } P_\gamma$ ? Что такое  $P_\gamma^2$ ?

- 20.10. (a) Пусть  $f(z)$  обозначает ветвь функции  $z^{1/4}$ , определенную везде, кроме положительной вещественной оси, и обладающую свойством

$$\lim_{\varepsilon \rightarrow 0} f(x + i\varepsilon) = |x|^{1/4} \quad (\text{вещественное число}).$$

Вычислите  $f(-x)$  и предел  $\lim_{\varepsilon \rightarrow 0} f(x - i\varepsilon)$  для вещественных  $x$ .

- (b) Используя контур, изображенный на рис. 20.19, вычислите определенный интеграл

$$I = \int_0^\infty \frac{t^{1/4} dt}{(t+a)^2}.$$

Обсудите вклад от каждой окружности.

- 20.11. Используя метод, изложенный в параграфе 20.7, пункт (c), вычислите интегралы по контуру, изображенному на рис. 20.14.

$$(a) \int_0^\infty \frac{x^{\alpha-1}}{x+1} dx = \frac{\pi}{\sin \alpha\pi}, \quad 0 < \alpha < 1,$$

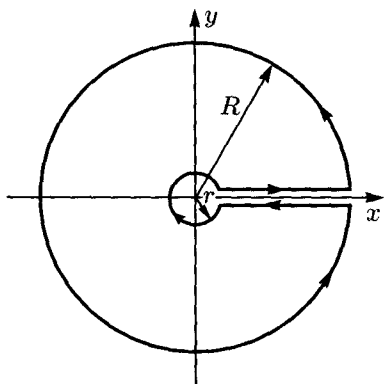


Рис. 20.19

$$(b) \int_0^\infty \frac{x^{\alpha-1}}{(x+1)^2} dx = \frac{(1-\alpha)\pi}{\sin \alpha\pi}, \quad 0 < \alpha < 2.$$

- 20.12. (a) Вычислите  $\int f(z) \log z \, dz$  по контуру, изображенному на рис. 20.15. Используя ветвь  $\log z$ , определенную условием

$$\log z = \log r + i\theta, \quad 0 < \theta < 2\pi,$$

докажите формулу

$$\int_0^\infty f(x) \, dx = - \sum \operatorname{res}(f(z) \log z),$$

которая верна всегда, когда  $f(z)$  стремится к нулю при  $|z| \rightarrow \infty$  достаточно быстро, чтобы интеграл по большому кругу стремился к нулю.

- (b) Покажите, что

$$\int_0^\infty \frac{dx}{(x+a)(x^2+b^2)} = \frac{\pi a + 2b \log(b/a)}{2b(a^2+b^2)}, \quad a, b > 0.$$

- (c) Рассмотрев интеграл  $\int f(z)(\log z)^2 \, dz$ , придумайте способ вычисления интеграла

$$\int_0^\infty f(x) \log x \, dx.$$

С его помощью докажите формулу

$$\int_0^\infty \frac{\log x}{x^2+a^2} \, dx = \frac{\pi \log a}{2a}.$$

20.13. (a) Получите ряд Лорана для функции

$$f(z) = \frac{5}{z^2 - 5z/2 + 1}$$

с помощью интегрирования по единичной окружности. Где сходится этот ряд?

(b) Этот же ряд получите с помощью разложения функции  $f(z)$  на простые дроби. Затем разложите каждый член по степеням  $z$  или  $1/z$  так, чтобы ряд сходился при  $|z| = 1$ .

(c) Получите ряд Фурье для функции

$$F(\theta) = (5/2 + 2 \cos \theta)^{-1}.$$

20.14. Получите ряд Фурье для функции

$$F(\theta) = \theta, \quad -\pi < \theta < \pi.$$

Можно ли это получить как разложение голоморфной функции в ряд Лорана?

## Глава 21

---

# Асимптотическое вычисление интегралов

---

### Введение

Асимптотическое поведение интегралов позволяет понять многие физические законы. В этой главе мы дадим элементарные примеры математических приемов, используемых в настоящее время во многих приложениях физики и математики.

*Асимптотическая формула Стирлинга* — одна из первых и до сих пор важнейших:

$$n! \sim (2\pi)^{1/2} n^{n+1/2} e^{-n}.$$

Она получена Джеймсом Стирлингом в первой трети XVIII века. Знак  $\sim$  следует понимать в том смысле, что отношение обеих частей формулы (они обе стремятся к  $\infty$  с ростом  $n$ ) стремится к 1 при  $n \rightarrow \infty$ . Мы приведем и докажем здесь более точную версию этой формулы вместе с оценкой погрешности.

Хотя формула Стирлинга должна применяться только при больших значениях  $n$ , она дает удивительно точное приближение (в смысле относительной погрешности) даже при малых  $n$ . Так, для  $n = 1$  мы имеем  $n! = 1$ , а правая часть формулы Стирлинга дает 0.92. Для  $n = 10$  мы имеем  $n! = 3\,628\,800$ , а правая сторона дает 3\,598\,600, т. е. ошибка меньше одного процента. Это систематическое явление. В этой главе мы дадим оценки, справедливые для больших значений параметра, но оказывается, что эта формула хорошо работает и при малых значениях.

## 21.1. Метод Лапласа

Мы применим этот метод для получения более точной версии формулы Стирлинга. Интегрирование по частям показывает, что *гамма-функция*, определенная формулой

$$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt, \quad x > 0,$$

удовлетворяет условию

$$\Gamma(x+1) = x\Gamma(x), \quad \Gamma(1) = 1,$$

так что

$$\Gamma(n+1) = n!.$$

Таким образом, можно переписать формулу Стирлинга в виде

$$\Gamma(x) \sim e^{-x} x^x \left( \frac{2\pi}{x} \right)^{1/2}.$$

На самом деле мы докажем более точный результат. Покажем, что

$$\Gamma(x) \sim e^{-x} x^x \left( \frac{2\pi}{x} \right)^{1/2} \left( 1 + \frac{1}{12x} + \frac{1}{288x^2} + \dots \right).$$

Асимптотический ряд в этой формуле означает, например, что

$$\begin{aligned} \left| \Gamma(x) - e^{-x} x^x \left( \frac{2\pi}{x} \right)^{1/2} \right| &< C_0 \left( e^{-x} x^x \left( \frac{2\pi}{x} \right) x^{-1} \right), \\ \left| \Gamma(x) - e^{-x} x^x \left( \frac{2\pi}{x} \right)^{1/2} \left( 1 + \frac{1}{12x} \right) \right| &< C_1 \left( e^{-x} x^x \left( \frac{2\pi}{x} \right)^{1/2} x^{-2} \right), \\ \left| \Gamma(x) - e^{-x} x^x \left( \frac{2\pi}{x} \right)^{1/2} \left( 1 + \frac{1}{12x} - \frac{1}{288x^2} \right) \right| &< C_2 \left( e^{-x} x^x \left( \frac{2\pi}{x} \right)^{1/2} x^{-3} \right), \end{aligned}$$

и т. д., где  $C_i$  — константы. Другими словами, обрывая ряд в любом месте, можно оценить погрешность через следующую, более

высокую степень  $x^{-1}$ . Начнем с формулы  $\Gamma(x) = \Gamma(x+1)/x$  и получим

$$\begin{aligned}\Gamma(x) &= \frac{1}{x} \int_0^\infty e^{-t} t^x dt \\ &= e^{-x} x^x \int_{-1}^\infty e^{-x[w - \log(1+w)]} dw \\ &= e^{-x} x^x \int_{-1}^\infty e^{-xp(w)} dw,\end{aligned}$$

где  $t = (1+w)x$ , так что  $dt = x dw$  и  $t = 0 \Leftrightarrow w = -1$ ,  $t = \infty \Leftrightarrow w = \infty$ , где функция

$$p(w) = w - \log(1+w) = \frac{1}{2}w^2 - \frac{1}{3}w^3 + \dots$$

раскладывается указанным образом в ряд Тейлора в нуле. Заметим, что  $p'(w) = 0$  только при  $w = 0$ , и  $p$  имеет абсолютный минимум в нуле. Множитель  $e^{-x}x^x$  присутствует в правой части формулы Стирлинга. Поэтому нам надо исследовать асимптотическое поведение интеграла

$$\int_{-1}^\infty e^{-xp(w)} dw.$$

Прежде всего отметим, что для больших  $x$  единственный большой вклад в интеграл получается от  $w$  в окрестности нуля — минимума функции  $p$ . Например, если  $\varepsilon$  — любое фиксированное положительное число, то

$$\int_{-1}^{-\varepsilon} e^{-xp(w)} dw < \int_{-1}^{-\varepsilon} e^{-x \cdot w^2/2} dw < \int_{-\infty}^{-\varepsilon} e^{-xw^2/2} dw.$$

Можно получить грубую оценку последнего интеграла, написав

$$e^{-xw^2/2} = e^{-xw^2/4} \cdot e^{-xw^2/4} \leq e^{-\varepsilon^2 x/4} e^{-xw^2/4}$$

для  $-\infty < w < -\varepsilon$ . Поэтому

$$\int_{-\infty}^{-\varepsilon} e^{-xw^2/2} dw \leq C(x) e^{-\varepsilon^2 x/4},$$

где  $C(x) = \int_{-\infty}^0 e^{-xw^2/4} dw$  — сходящийся интеграл

и, более того,  $C(x) \rightarrow 0$  при  $x \rightarrow \infty$ . Следовательно, интеграл  $\int_{-1}^{-\varepsilon}$  экспоненциально убывает с ростом  $x$ . Аналогично, при  $w > 0$  функция  $p(w)$  строго возрастает, так что

$$\int_{\varepsilon}^{\infty} e^{-xp(w)} dw < e^{-p(\varepsilon)x/2} \int_{\varepsilon}^{\infty} e^{-xp(w)/2} dw.$$

Этот второй интеграл сходится абсолютно и равномерно в области  $x > 0$  (причем все быстрее, когда  $x \rightarrow \infty$ ). Таким образом, мы можем контролировать оба интеграла  $\int_{-1}^{-\varepsilon}$  и  $\int_{\varepsilon}^{\infty}$  и заняться интегралом  $\int_{-\varepsilon}^{\varepsilon} e^{-xp(w)} dw$  для  $\varepsilon > 0$ . Теперь  $p(0) = 0$ ,  $p'(0) = 0$  и  $p''(0) = 1$ . Поэтому можно сделать замену переменной  $w = w(s)$  в окрестности нуля, так чтобы  $p$  приняла вид  $s^2/2$ , т. е.

$$p(w(s)) = s^2/2.$$

Напомним читателю, как это делается. Поскольку  $p(0) = 0$ , мы имеем

$$p(w) = \int_0^w p'(r) dr = w \int_0^1 p'(wu) du,$$

где мы положили  $r = wu$  при переходе от первого интеграла ко второму. Аналогично,

$$p'(w) = w \int_0^1 p''(wu) du$$

так что

$$p(w) = \frac{1}{2} w^2 q(w),$$

где

$$q(w) = 2 \int_0^1 \int_0^1 up''(uvw) du dv.$$

Заметим, что  $q$  — дифференцируемая функция от  $w$  и  $q(0) = 1$ . Значит,  $q(w) > 0$  для достаточно малых  $w$ . Теперь можно сделать замену переменных

$$s = w\sqrt{q(w)}$$

и

$$p(w(s)) = \frac{1}{2} s^2.$$

Выберем  $\varepsilon > 0$  настолько малым, чтобы такая замена переменной была допустима при  $-\varepsilon < w < \varepsilon$   $q(w)$  (т.е. чтобы функция  $g(w)$  была положительна в этом интервале). Теперь сделаем замену переменной в интеграле  $\int_{-\varepsilon}^{\varepsilon} e^{-xp(w)} dw$  и получим

$$\int_{-\varepsilon}^{\varepsilon} e^{-xp(w)} dw = \int_{-\varepsilon_1}^{\varepsilon_2} e^{-xs^2/2} \left( \frac{dw}{ds} \right) ds$$

где  $\varepsilon_2 = \varepsilon \sqrt{q(\varepsilon)}$  и  $-\varepsilon_1 = -\varepsilon \sqrt{q(-\varepsilon)}$  (см. график на рис. 21.1).

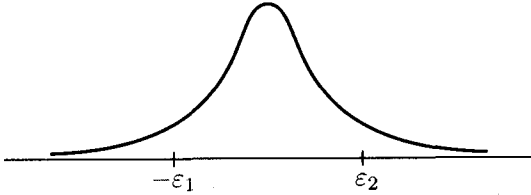


Рис. 21.1

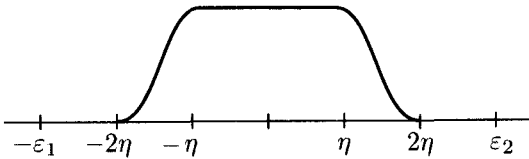


Рис. 21.2

При оценке интеграла играет роль только область в окрестности  $s = 0$ . Например, предположим, что у нас есть бесконечно дифференцируемая функция  $\rho$  такая, что  $\rho \equiv 1$  при  $|s| < \eta$  и  $\rho \equiv 0$  при  $s > 2\eta$ , где  $-\varepsilon_1 < -2\eta$  и  $2\eta < \varepsilon$  (рис. 21.2). Тогда обозначим

$$b(s) = \rho(s) \frac{dw}{ds}, \quad \text{т.е.} \quad b(s) \equiv \frac{dw}{ds}(s) \quad \text{для } |s| < \eta.$$

Напишем

$$\int_{-\varepsilon_1}^{\varepsilon_2} e^{-xs^2/2} w'(s) ds = \int_{-\varepsilon_1}^{\varepsilon_2} e^{-xs^2/2} b(s) ds + \text{погрешность},$$

где погрешность экспоненциально стремится к нулю с ростом  $x$ , в частности, быстрее по порядку величины, чем  $e^{-\eta^2 x}$ . (Это является следствием того, что погрешность дается интегралами вида  $\int_{-\varepsilon_1}^{\eta} e^{-xs^2/2}(b(s) - w'(s)) ds$  и  $\int_{\eta}^{\varepsilon_2} e^{-xs^2/2}(b(s) - w'(s)) ds$ . Тогда подынтегральное выражение стремится к нулю быстрее, чем константа на  $e^{-\eta^2 x}$ .)

Интеграл  $\int_{-\varepsilon_1}^{\varepsilon_2} e^{-xs^2/2} b(s) ds$  можно заменить интегралом

$$\int_{-\infty}^{\infty} e^{-xs^2/2} b(s) ds$$

поскольку  $b \equiv 0$  для  $s < \varepsilon_1$  и  $s > \varepsilon_2$ . Таким образом, мы все свели к оценке этого последнего интеграла.

Воспользуемся формулой Тейлора с остатком и напишем

$$b(s) = b_0 + b_1 s + b_2 s^2 + \dots + b_{N-1} s^{N-1} + b_N(s) s^N,$$

где  $b_N(s)$  — гладкая ограниченная функция. Тогда

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-xs^2/2} b(s) ds &= b_0 \int_{-\infty}^{\infty} e^{-xs^2/2} ds + b_1 \int_{-\infty}^{\infty} e^{-xs^2/2} s ds \\ &\quad + b_2 \int_{-\infty}^{\infty} e^{-xs^2/2} s^2 ds + \dots + R_N(s), \end{aligned}$$

где

$$R_N(s) = \int_{-\infty}^{\infty} e^{-xs^2/2} b_N(s) s^N ds.$$

Следовательно,

$$|R_N(s)| \leq D_N \int_{-\infty}^{\infty} e^{-xs^2/2} |s|^N ds,$$

где

$$D_N = \sup |b_N(s)|.$$

В силу симметрии

$$\int_{-\infty}^{\infty} e^{-xs^2/2} s^k ds = 0, \quad \text{если } k \text{ нечетно.}$$

Для четного числа  $k$  сделаем замену переменной

$$x^{1/2}s = y, \quad \text{так что} \quad ds = \frac{dy}{x^{1/2}}.$$

Тогда имеем

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-xs^2/2} s^k ds &= x^{-1/2-k} \int_{-\infty}^{\infty} e^{-y^2/2} y^k dy \\ &= C_k x^{-1/2-k/2}, \end{aligned}$$

где

$$C_k = \int_{-\infty}^{\infty} e^{-y^2/2} y^k dy$$

не зависит от  $x$  (причем  $C_0 = \sqrt{2\pi}$ ). Заметим, что

$$\int_{-\infty}^{\infty} e^{-y^2/2} y^k dy = 2 \int_0^{\infty} e^{-u} (2u)^{(k-1)/2} du = 2^{(k+1)/2} \Gamma\left(\frac{k+1}{2}\right),$$

откуда

$$C_k = 2^{(k+1)/2} \Gamma\left(\frac{k+1}{2}\right).$$

Аналогично, остаточный член можно оценить произведением  $x^{-1/2-N/2}$  на константу. Возьмем  $N = 2n$ , тогда

$$\int_{-\infty}^{\infty} e^{-xs^2/2} b(s) ds = \left(\frac{2\pi}{x}\right)^{1/2} (a_0 + a_1 x^{-1} + \dots + a_{n-1} x^{-n+1} + \varepsilon_n(x)),$$

где  $\varepsilon_n(x) = o(x^{-n})$ . Здесь  $a_0 = b_0$ ,  $a_1 = b_2$ ,  $a_2 = b_4/\sqrt{2\pi}$  и т.д. Функции  $b$  и  $dw/ds$  имеют *одинаковое* разложение в ряд Тейлора в 0. Подстановка разложения Тейлора функции  $p(w) = w - \log(1+w)$  дает более точное приближение Стирлинга.

Предположим, что нам дана гладкая положительная функция  $p$  с единственным абсолютным минимумом в точке  $t_0$  внутри области определения этой функции (рис. 21.3), и при этом  $p''(t_0) > 0$ . Предположим, что  $a(t)$  — другая гладкая функция. Рассмотрим интеграл

$$\int a(t) e^{-kp(t)} dt.$$

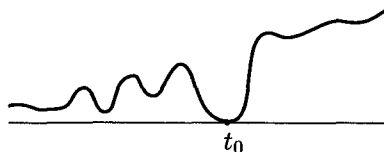


Рис. 21.3

Это можно переписать в виде

$$e^{-kp(t_0)} \int a(t) e^{-k(p(t)-p(t_0))} dt,$$

где функция  $p(t) - p(t_0)$  имеет минимальное значение 0 при  $t = t_0$ . Изложенный выше метод позволяет доказать *принцип Лапласа* (открыт в 1820 году):

$$\int a(t) e^{-kp(t)} dt \sim \frac{e^{-kp(t_0)}}{\sqrt{p''(t_0)}} \left( \frac{2\pi}{k} \right) \left( a_0 + \frac{a_1}{k} + \frac{a_2}{k^2} + \dots \right),$$

где  $a_0 = a(t_0)$ , а последующие коэффициенты выражаются через функции  $a$  и  $p$  и их производные в точке  $t_0$ .

## 21.2. Метод стационарной фазы

Давайте разберемся, что произойдет с методом Лапласа, если в экспоненте  $-kp$  заменить на  $ikp$ . Окажется, что справедлива аналогичная формула, но в этом случае в ней учитываются все критические точки функции  $p$ , а не только абсолютный минимум. В методе Лапласа множитель  $e^{-p(w)}$  в подынтегральном выражении экспоненциально убывает по сравнению со значением вблизи минимума  $p$ . Поэтому вся информация об асимптотике функции была локализована в этом минимуме. Когда мы рассматриваем интеграл

$$\int e^{ikp(y)} a(y) dy,$$

множитель  $e^{ikp(y)}$  уже не стремится к нулю при  $k \rightarrow \infty$ . Вместо этого он быстро осциллирует как функция от  $y$  везде, где  $p'(y) \neq 0$ . Быстрые осцилляции приводят к взаимным сокращениям, поэтому существенный вклад в интеграл вносят области в

окрестности критических точек. Чтобы это увидеть, давайте рассмотрим случай, когда  $p'(y)$  не обращается в нуль во всей области, где  $a(y) \neq 0$ . Точнее, предположим, что функция  $a(y)$  имеет производные всех порядков и компактный носитель, т. е. равна нулю вне конечного интервала. Кроме того, предположим, что

$$|p'(y)| > C > 0 \quad \text{для всех } y, \text{ где } a(y) \neq 0.$$

Тогда можно утверждать, что для любого числа  $N$  существует постоянная  $C_N$  такая, что

$$\left| \int e^{ikp(y)} dy \right| < \frac{C_N}{k^N}.$$

Другими словами, мы утверждаем, что интеграл стремится к нулю быстрее, чем любая отрицательная степень  $k$ . Действительно, рассмотрим дифференциальный оператор

$$D = \frac{1}{p'(y)} \frac{\partial}{\partial y},$$

который (в силу наших допущений) имеет смысл в области, где  $a(y) \neq 0$ , и

$$De^{ikp(y)} = ike^{ikp(y)}$$

или

$$e^{ikp(y)} = \frac{1}{ik}(De^{ikp(y)}).$$

Тогда

$$\begin{aligned} \int e^{ikp(y)} a(y) dy &= \int \frac{1}{ik} (De^{ikp(y)}) a(y) dy \\ &= \int \frac{1}{ik} \left( \frac{\partial}{\partial y} e^{ikp(y)} \right) \frac{a(y)}{p'(y)} dy \\ &= \frac{1}{k} \int e^{ikp(y)} c(y) dy, \end{aligned}$$

где

$$c(y) = -\frac{1}{i} \frac{\partial}{\partial y} \left( \frac{a(y)}{p'(y)} \right).$$

Здесь мы проинтегрировали по частям и воспользовались тем, что функция  $a$  вместе со своими производными равна нулю на границах области интегрирования. Функция  $c$  обладает теми же свойствами, что и  $a$ , поэтому эту процедуру можно повторить и получить

$$\int b(y)e^{ikp(y)} dy = \frac{1}{k^N} \int c_N(y)e^{ikp(y)} dy$$

для соответствующей функции  $c_N(y)$ . Если теперь обозначить

$$c_N = \left| \int c_N(y)e^{ikp(y)} dy \right|,$$

то наше утверждение доказано.

Рассмотрим теперь общий случай. По-прежнему считаем, что  $a$  — гладкая функция с компактным носителем. Но теперь мы не исключаем, что  $p'(y)$  имеет нули. Однако предполагаем, что функция  $p$  имеет конечное число критических точек  $y_1, \dots, y_r$  (в которых  $p'(y) = 0$ ), и что  $p''(y) \neq 0$  в каждой из них. Тогда утверждается, что

$$\int a(y)e^{ikp(y)} dy \sim \left(\frac{2\pi}{k}\right)^{1/2} \sum_{j=1}^r \frac{e^{\pm\pi i/4} e^{ikp(y_j)}}{\sqrt{|p''(y_j)|}} \left(a_{0,j} + \frac{a_{1,j}}{k} + \dots\right),$$

где  $a_{0,j} = a(y_j) = a_j$ , а остальные коэффициенты выражаются через  $a$ ,  $p$  и их производные в точке  $y_j$ . Знак  $+$  в функции  $e^{\pm\pi i/4}$  берется в точке минимума, где  $p''(y_j) > 0$ , а знак  $-$  в точке максимума, где  $p''(y_j) < 0$ . В этой формуле функция  $p$  принимает вещественные значения, а функция  $a$  может быть комплексной функцией  $y$ . Это называется формулой стационарной фазы. Докажем ее. В окрестности каждой критической точки  $y_i$  можно найти такую координату  $s$ , для которой

$$p(y(s)) = \pm s^2/2$$

(знак выбирается в зависимости от знака  $p''(y_j)$ ). Сделаем замену переменной и получим

$$\frac{dy}{ds} = \frac{1}{\sqrt{|p''(y_j)|}} \quad \text{в точке } y_j.$$

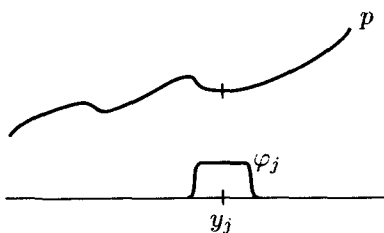


Рис. 21.4



Рис. 21.5

Этим объясняется множитель  $1/\sqrt{|p''(y_j)|}$  в формуле стационарной фазы.

В окрестности каждой точки  $y_j$  определим такую функцию  $\phi_j$ , что  $\phi_j \equiv 1$  вблизи  $y_j$  и  $\phi_j \equiv 0$  вне малой окрестности, где определена процедура замены переменной (рис. 21.4); причем окрестности различных точек не пересекаются (рис. 21.5). Тогда функция

$$b = a - (a_1\phi_1 + a_2\phi_2 + \dots + a_r\phi_r)$$

имеет компактный носитель и  $p'(y) \neq 0$  везде, где  $b(y) \neq 0$ . Таким образом, можно применить предыдущий результат к  $b$  и получить

$$\int e^{ikp} a \, dy = \sum \int e^{ikp} (a_j \phi_j) \, dy + O(k^{-N})$$

для любого значения  $N$ . Каждый член суммы можно рассмотреть отдельно, сделав в нем соответствующую замену переменной. Для  $j$ -го слагаемого получаем

$$\int e^{\pm iks^2/2} \left( a_j \phi_j \frac{dy}{ds} \right) ds.$$

Определим функцию  $\psi$  равенством

$$\psi = a_j \phi_j \frac{dy}{ds}.$$

Тогда функция  $\psi$  имеет компактный носитель, и мы свели задачу к вычислению интеграла

$$\int_{-a}^{\infty} e^{\pm iks^2/2} \psi(s) \, ds.$$

Чтобы это сделать, сначала рассмотрим гауссовы интегралы. Мы уже слегка коснулись этого вопроса в предыдущей главе.

### 21.3. Гауссовы интегралы

Материал, излагаемый в этом параграфе, основывается на равенстве

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx = 1.$$

Его можно доказать, взяв квадрат от левой части

$$\begin{aligned} \left[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx \right]^2 &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)/2} dx dy \\ &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} e^{-r^2/2} r dr d\theta \\ &= \int_0^{\infty} e^{-r^2/2} r dr = 1. \end{aligned}$$

Функция

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2 - \eta x} dx$$

сходится равномерно для всех комплексных значений  $\eta$  в произвольной компактной области. Следовательно, этот интеграл определяет аналитическую функцию, которую можно вычислить для вещественных  $\eta$  и потом аналитически продолжить. Для вещественного числа  $\eta$  показатель экспоненты дополняем до полного квадрата и делаем замену переменной

$$\begin{aligned} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2 - \eta x} dx &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x+\eta)^2/2} dx \\ &= e^{\eta^2/2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-(x+\eta)^2/2} dx \\ &= e^{\eta^2/2}. \end{aligned}$$

Эта формула верна для всех комплексных  $\eta$ . Если взять  $\eta = i\zeta$ , то получим

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2 - i\zeta x} dx = e^{-\zeta^2/2}.$$

Другими словами, преобразованием Фурье функции  $e^{-x^2/2}$  является  $e^{-\zeta^2/2}$ . Теперь положим  $x = \sqrt{\lambda} u$ , где  $\sqrt{\lambda} > 0$  — положительное вещественное число. Тогда  $dx = \sqrt{\lambda} du$ . Положив  $\xi = \sqrt{\lambda} \zeta$ ,

получаем

$$\frac{1}{\sqrt{2\pi}} \int e^{-\lambda u^2/2} e^{-i\xi u} du = \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2}.$$

Этот интеграл сходится для всех  $\lambda$  при условии, что  $\operatorname{Re} \lambda > 0$  и  $\lambda \neq 0$ . Действительно, для  $S > R > 0$  имеем

$$\begin{aligned} \int_R^S e^{-\lambda x^2/2} e^{-i\xi x} dx &= - \int_R^S \frac{1}{\lambda x} \frac{d}{dx} (e^{-\lambda x^2/2}) e^{-i\xi x} dx \\ &= \frac{e^{-\lambda R^2/2} e^{-i\xi R}}{\lambda R} - \frac{e^{-\lambda S^2/2} e^{-i\xi S}}{\lambda S} + \int_R^S e^{-\lambda x^2/2} \frac{d}{dx} \left( \frac{e^{i\xi x}}{\lambda x} \right) dx. \end{aligned}$$

В последнем интеграле можно написать

$$e^{-\lambda x^2/2} = - \frac{1}{\lambda x} \frac{d}{dx} e^{-\lambda x^2/2}$$

и еще раз проинтегрировать по частям. Тогда получим

$$\left| \int_R^S e^{-\lambda x^2/2} e^{-i\xi x} dx \right| = O \left( \frac{1}{|\lambda|R} \right).$$

Отсюда видно, что функция

$$\frac{1}{\sqrt{2\pi}} e^{-\lambda x^2/2} dx$$

равномерно (но не абсолютно) сходится по  $\lambda$  в произвольной области вида

$$\operatorname{Re} \lambda > 0, \quad |\lambda| > \delta, \quad \text{при любом } \delta > 0.$$

Сделаем аналитическое продолжение. Значения функции задаются той же формулой, что и при вещественных положительных  $\lambda$ , но теперь квадратный корень должен быть аналитическим продолжением положительного квадратного корня. В частности, если взять  $\lambda = -ir$ ,  $r > 0$ , то

$$\lambda^{1/2} = |\lambda|^{1/2} e^{-\pi i/4}$$

и

$$\frac{1}{\sqrt{2\pi}} \int e^{irx^2/2} e^{-i\xi x} dx = e^{\pi i/4} r^{-1/2} e^{-i\eta^2/2r}.$$

Если взять  $\lambda = ir$ , то аналогично

$$\frac{1}{\sqrt{2\pi}} \int e^{-irx^2/2} e^{-i\xi x} dx = e^{-\pi i/4} r^{-1/2} e^{i\eta^2/2r}.$$

Теперь можно закончить доказательство формулы стационарной фазы. Для этого нам надо получить асимптотическую оценку интеграла вида

$$\int_{-\infty}^{\infty} e^{\pm iks^2/2} \psi(s) ds$$

где  $\psi$  — функция с компактным носителем, которая имеет вид

$$\psi(s) = \psi(0) + sx(s).$$

Умножим это равенство на функцию  $\rho$ , которая тождественно равна 1 там, где  $\psi \neq 0$  и равна нулю для больших значений  $|s|$ . Тогда

$$\psi(s) = \psi_0(s) + s\rho(s)x(s),$$

где

$$\psi_0(s) = \rho(s)\psi(0) \equiv \psi(0) \quad \text{вблизи } s = 0.$$

Таким образом,

$$\int \psi_0(s) e^{\pm iks^2/2} ds = \frac{2\pi}{k} e^{\pm i/4} + \text{погрешность},$$

где погрешность убывает быстрее любой степени  $k$ . Однако,

$$\begin{aligned} \int s[\rho(s)x(s)] e^{iks^2/2} ds &= \frac{1}{\pm ik} \int [\rho(s)x(s)] \frac{d}{ds} e^{\pm iks^2/2} ds \\ &= -\frac{1}{ik} \int \frac{d}{ds} [\rho(s)x(s)] \cdot e^{\pm iks^2/2} ds, \end{aligned}$$

и  $(d/ds)[\rho(s)x(s)]$  равняется нулю для больших  $|s|$ . Эту процедуру можно повторять, пока не получим асимптотический ряд, что и завершает доказательство нашей формулы.

Давайте получим  $n$ -мерную версию гауссова интеграла. Пусть

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ — векторная переменная и } Q = \begin{pmatrix} \pm r_1 & & 0 \\ & \ddots & \\ 0 & & \pm r_n \end{pmatrix}$$

является матрицей  $n \times n$  для всех  $r_i > 0$  и  $\text{sgn } Q$  равняется числу плюсов минус число минусов. Тогда  $|r_1| \cdots |r_n| = |\text{Det}(Q)|$ . Умножая формулы для одномерных интегралов, получаем

$$\frac{1}{(2\pi)^{n/2}} \int \exp\left(\frac{i}{2} Q \mathbf{x} \cdot \mathbf{x}\right) \exp(-i \boldsymbol{\xi} \cdot \mathbf{x}) d\mathbf{x} = |\text{Det}(Q)|^{-1/2} \\ - \exp\left(\frac{\pi i}{4} \text{sgn } Q\right) \times \exp\left(-\frac{i}{2} Q^{-1} \boldsymbol{\xi} \cdot \boldsymbol{\xi}\right),$$

где теперь  $\boldsymbol{\xi}$  — тоже вектор, а  $Q^{-1}$  — обратная матрица. Итак, мы доказали формулу для диагональной матрицы  $Q$ . Однако, для любой несингулярной симметричной матрицы  $Q$  можно найти ортогональную матрицу  $O$  такую, чтобы матрица  $OQO^{-1}$  была диагональной. Это значит, что полученная выше формула верна для всех таких матриц  $Q$ .

Мы можем использовать  $n$ -мерную версию гауссова интеграла, чтобы получить  $n$ -мерную версию формулы стационарной фазы. Пусть  $p$  — функция с конечным числом критических точек. Предположим, что все критические точки функции  $p$  невырождены. Тогда гессиан

$$H(y_j) = \left( \frac{\partial^2 p}{\partial y_k \partial y_l} \right) (y_j).$$

является невырожденной матрицей во всех критических точках  $y_j$ . Пусть  $\sigma = \sigma_j$  обозначает сигнатуру  $j$ -й критической точки. Тогда

$$\int a(y) e^{ikp(y)} dy \\ \sim \left( \frac{2\pi}{k} \right)^{n/2} \sum \frac{e^{\pi i \sigma / 4}}{\sqrt{|\text{Det}(H(y_j))|}} \left( a(y_j) + \frac{a_1(y_j)}{k} + \cdots \right),$$

где сумма берется по всем критическим точкам. Здесь величины  $a_1(y_j)$ ,  $a_2(y_j)$  и т.д. выражаются через функции  $a$  и  $p$  и их производные различных порядков в точках  $y_j$ .

## 21.4. Групповая скорость

В качестве применения метода стационарной фазы в одномерном пространстве рассмотрим следующую ситуацию. Пусть у нас есть семейство бегущих волн

$$e^{-(i/h)(E(p)t - px)},$$

где  $h$  — малое число, так что  $1/h$  играет роль нашего большого параметра  $k$ . Волновое число для *пространственной вариации* равно  $p/h$  в каждый момент времени  $t$ . Мы полагаем, что  $E$  зависит от  $p$ , т.е. каждая волна имеет свою скорость. Предположим, что мы сложили семейство бегущих волн, что выражается интегралом вида

$$\int a(p) e^{-(i/h)(E(p)t - px)} dp.$$

Далее, предположим, что функция  $a(p)$  сосредоточена вокруг некоторого фиксированного значения  $p_0$ . Другими словами, мы предполагаем, что функция  $a$  отлична от нуля только для значений  $p$  в небольшой окрестности  $p_0$ . Согласно методу стационарной фазы, существенный вклад в интеграл возникает только от тех значений  $p$ , где производная по  $p$  показателя экспоненты равняется нулю, т.е. только при условии, что

$$E'(p)t - x = 0.$$

Поскольку  $a(p)$  равняется нулю, если  $p$  не лежит в малой окрестности  $p_0$ , то это уравнение дает нам связь между  $x$  и  $t$ . Из уравнения следует, что интеграл практически равен нулю везде, кроме тех значений  $x$  и  $t$ , для которых равенство

$$x = E'(p_0)t$$

выполняется приближенно. Другими словами, интеграл выглядит как небольшой сгусток, если рассматривать его как функцию  $x$ , причем этот сгусток движется со временем со скоростью  $E'(p)$ . (Этот сгусток называется *волновым пакетом*, а его скорость  $E'(p)$  называется *групповой скоростью*.)

Давайте исследуем, какой должна быть функциональная зависимость  $E$  от  $p$ , если потребовать, чтобы выражение  $E \cdot t - p \cdot x$  было инвариантно относительно преобразований Лоренца. Согласно преобразованию Лоренца, точка  $(t, x)$  может перейти в точку  $(t', x')$  при условии, что выполняется равенство

$$c^2 t^2 - x^2 = c^2 t'^2 - x'^2.$$

Тогда  $(E, p)$  преобразуется в  $(E', p')$ , причем

$$E^2 - c^2 p^2 = E'^2 - c^2 p'^2.$$

Таким образом, единственная инвариантная связь между  $E$  и  $p$  имеет вид

$$E^2 - c^2 p^2 = \text{const.}$$

Обозначим эту константу  $m^2 c^4$ , так что

$$E^2 - c^2 p^2 = m^2 c^4$$

или

$$E(p) = ((pc)^2 + m^2 c^4)^{1/2}.$$

Тогда

$$E'(p) = \frac{pc^2}{E(p)} = \frac{p}{M},$$

где

$$E(p) = Mc^2 \quad \text{и} \quad M = (m^2 + (pc)^2)^{1/2},$$

так что, если  $p/c$  мало по сравнению с  $m$ , то  $M \doteq m$ . Если считать, что  $M$  — это масса, то связь между групповой скоростью  $E'(p)$  и  $p$  — это в точности соотношение между скоростью и импульсом. Таким образом, мы связали волновое число  $k = p/h$  с импульсом  $p$ , а соответствующая длина волны

$$\lambda = \frac{1}{k} = \frac{h}{p} \quad (\text{формула де Бройля}).$$

Если считать, что  $E$  — энергия, то связь между энергией и частотой  $\nu$  временной вариации дается формулой

$$E = h\nu \quad (\text{формула Эйнштейна}).$$

В этих формулах стоит малый параметр  $h$ , который мы устремляли к нулю. Великое открытие квантовой механики состоит в том, что параметр  $h$  не должен стремиться к нулю. Это фундаментальная константа природы, известная как постоянная Планка. В формуле Эйнштейна она возникает как множитель, связывающий энергию и обратное время, следовательно, ее размерность энергия  $\times$  время. Значение постоянной Планка равно

$$h = 6.626 \times 10^{-34} \text{ Дж} \cdot \text{с}.$$

## 21.5. Формула обратного преобразования Фурье

Давайте рассмотрим очень важное применение формулы стационарной фазы. Пусть  $n = 2$ , координаты обозначим  $x$  и  $\xi$ . Возьмем функцию

$$p(x, \xi) = p_\eta(x, \xi) = x(\xi - \eta),$$

где  $\eta$  — фиксированное число. Эта функция имеет только одну (невырожденную) критическую точку, а именно

$$x = 0, \quad \xi = \eta$$

с сигнатурой  $\sigma = 0$ .

Тогда

$$\frac{1}{2\pi} \int e^{ikx(\xi-\eta)} a(x, \xi) dx d\xi = \frac{1}{k} \left( a(0, \eta) + \frac{a_1}{k} + \dots \right).$$

Эта формула была доказана для  $a(x, \xi)$  с компактным носителем. Дальнейшее обобщение доказательства показывает, что она годится для функций, обращающихся на бесконечности в нуль вместе со своими производными любого порядка. Причем они стремятся к нулю быстрее, чем любая отрицательная степень  $k$ . В частности, допустим, что

$$a(x, \xi) = f(x)g(\xi),$$

где  $f$  и  $g$  — гладкие функции одной переменной, быстро стремящиеся к нулю на бесконечности вместе со своими производными. Тогда можно написать

$$\frac{1}{2\pi} \iint f(x)g(\xi)e^{-ikx(\xi-\eta)} dx d\xi = \frac{1}{k}f(0)g(\eta) + O\left(\frac{1}{k^2}\right).$$

В этом интеграле сделаем замену переменной  $u = kx$ . Тогда

$$\begin{aligned} \frac{1}{k}f(0)g(\eta) + O\left(\frac{1}{k^2}\right) &= \frac{1}{2\pi k} \int f\left(\frac{u}{k}\right) g(\xi)e^{-iu(\xi-\eta)} du d\xi \\ &= \frac{1}{k} \frac{1}{\sqrt{2\pi}} \int f\left(\frac{u}{k}\right) e^{iu\eta} \left( \frac{1}{\sqrt{2\pi}} \int g(\xi)e^{-iu\xi} d\xi \right) du. \end{aligned}$$

Пусть  $\hat{g}$  — Фурье-образ функции  $g$ , определенный формулой

$$\hat{g}(u) = \frac{1}{\sqrt{2\pi}} \int g(\xi)e^{-iu\xi} d\xi.$$

Подставим это выражение в предыдущую формулу и получим

$$\frac{1}{k} \frac{1}{\sqrt{2\pi}} \int f\left(\frac{u}{k}\right) \hat{g}(u)e^{iu\eta} du = \frac{1}{k}f(0)g(\eta) + O\left(\frac{1}{k^2}\right).$$

Предположим, что функция  $f$  удовлетворяет условию  $f(0) = 1$ . Тогда

$$f\left(\frac{u}{k}\right) = 1 + O\left(\frac{1}{k}\right)$$

и

$$\frac{1}{\sqrt{2\pi}} \int f\left(\frac{u}{k}\right) g(u)e^{iu\eta} du = \frac{1}{\sqrt{2\pi}} \int g(u)e^{iu\eta} du + O\left(\frac{1}{k}\right).$$

Переходя к пределу  $k \rightarrow \infty$ , получаем

$$g(\eta) = \frac{1}{\sqrt{2\pi}} \int \hat{g}(u)e^{-iu\eta} du.$$

Эта формула называется *формулой обратного преобразования Фурье*. Она показывает, как восстановить функцию по ее Фурье-образу.

Аналогично доказывается формула для пространства  $n$  измерений

$$g(\eta) = \frac{1}{(2\pi)^{n/2}} \int \hat{g}(u)e^{-iu\eta} du.$$

## 21.6. Асимптотическая оценка формулы Гельмгольца

Пусть  $u$  является решением приведенного волнового уравнения  $(\Delta + k^2)u = 0$  в трехмерном пространстве. В параграфе 19.6 мы показали, что если  $u$  удовлетворяет этому уравнению, то для точек  $P$  вне некоторой замкнутой поверхности  $S$ , работает формула Гельмгольца

$$u(P) = \frac{1}{4\pi} \iint_S \left( \frac{e^{ikr}}{r} \star du - u \star d \left( \frac{e^{ikr}}{r} \right) \right). \quad (19.2)$$

При этом для точек  $P$  внутри  $S$  тот же интеграл равен нулю. Там мы упоминали, что Гельмгольц таким образом объяснил принцип Гюйгенса — отсутствие обратных волн. Однако, обращение интеграла в нуль внутри  $S$  зависит от вкладов со всей поверхности. Френель полагал, что если бы внутри  $S$  были все волны, то интеграл от каждого малого элемента поверхности давал бы нулевой эффект внутри за счет интерференции. Используя метод стационарной фазы, мы сейчас покажем, что Френель был прав с точностью до членов порядка  $1/k$ . Когда мы применяем метод стационарной фазы, функция  $u$  в правой части формулы осциллирует. Прежде, чем двигаться дальше, необходимо сделать ряд предположений относительно нее. Предположим, что вблизи поверхности  $S$  эта функция имеет вид  $u = ae^{ik\phi}$ , где  $a$  и  $\phi$  — гладкие функции. Кроме того, пусть  $\|\text{grad } \phi\| = 1$ . Такая функция будет, например, описывать излучение от точки  $Q$ , лежащей внутри поверхности  $S$ , где  $\phi(y) = y - Q$ .

Далее, предположим, что мы находимся достаточно далеко от поверхности  $S$ , так что  $1/r^2$  пренебрежимо мало по сравнению с  $k$ , и что  $a$  и  $da$  тоже малы по сравнению с  $k$ . Подставим все в формулу Гельмгольца и получим, что член самого высокого порядка (относительно степеней  $k$ ) имеет вид

$$\frac{ik}{4\pi} \iint_S (a/r) e^{ik(\phi+r)} (\star d\phi - \star dr).$$

Здесь точками стационарной фазы являются точки  $y$  на поверхности  $S$ , где  $\phi(y) + \text{grad } r(y)$  перпендикулярен  $S$ . При этом возникают две возможных ситуации (рис. 21.6).

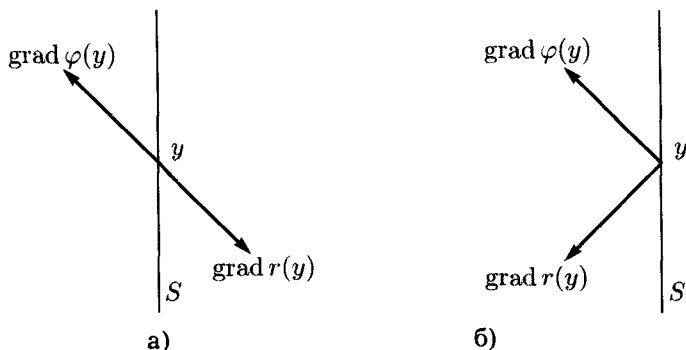


Рис. 21.6. Два типа: а)  $\text{grad } \phi(y) = -\text{grad } r(y)$ ,  $\star d\phi(y) = -\star dr(y)$ ,  $P = y + r \text{grad } \phi(y)$ ; б)  $\text{grad } \phi(y) = 2(\text{grad } \phi(y), n)n - \text{grad } r(y)$ ,  $\star d\phi(y) = \star dr(y)$ ,  $P = y - r(2(\text{grad } \phi(y), n)n - \text{grad } r(y))$

Сначала предположим, что  $y$  — невырожденная критическая точка типа (б). Член самого высокого порядка в формуле стационарной фазы равен нулю, и полный вклад в формулу Гельмгольца от такой точки  $y$  будет порядка  $1/k$ . (Заметим, что если бы поверхность  $S$  была выпуклой и  $\text{grad } \phi$  был направлен наружу, то для любой точки  $P$  внутри поверхности  $S$  все критические точки были бы типа (б). В некотором смысле это оправдывает точку зрения Френеля, что есть локальное сокращение обратных волн.) Для невырожденной критической точки типа (а), поскольку  $\star d\phi(y) = -\star dr(y)$ , при вычислении вклада самого высокого порядка в формулу стационарной фазы мы можем заменить прежний интеграл интегралом вида

$$\frac{ik}{2\pi} \iint_S (a/r) e^{ik(\phi+r)} d\star r.$$

Это показывает, что (с точностью до порядка  $1/k$ ) индуцированное *вторичное излучение* от поверхности  $S$  обладает следующими свойствами:

- (i) его амплитуда равна значению  $\lambda$ , умноженному на амплитуду первичной волны, где  $\lambda = 2\pi/k$  — длина волны;
- (ii) его фаза на одну четверть периода опережает фазу первичной волны. (Этим объясняется множитель  $i$ .)

Именно эти два свойства Френель постулировал в своей формулировке принципа Гюйгенса, и поэтому многие рассматривали его теорию как случайную. Мы здесь показали, что эти предположения являются следствием метода стационарной фазы и формулы Гельмгольца.

Теперь необходимо обсудить ситуацию, когда критические точки не вырождены. Исследуем точки типа (а). (Точки типа (б) рассматриваются аналогично.) Мы воспользуемся здесь теми же аргументами, что и при исследовании излучения. Определим экспоненциальное отображение  $E : S \times \mathbb{R} \rightarrow \mathbb{R}^3$  формулой

$$E(y, r) = y + r \operatorname{grad} \phi(y).$$

Тогда критические точки на поверхности  $S$ , связанные с точкой  $P$ , — это как раз те точки  $y$ , для которых  $E(y, r) = P$ , где  $r = \|y - P\|$ . Если  $\operatorname{grad} \phi(y)$  не касателен к поверхности  $S$ , то  $E$  является диффеоморфизмом в окрестности  $(y, 0)$ .

Нетрудно показать, что  $y$  — вырожденная критическая точка для  $P = E(y, r)$ , если и только если  $(y, r)$  — точка, в которой отображение  $E$  сингулярно. В этом случае назовем точку  $P$  фокальной точкой отображения  $E$  в  $y$ . Если же  $P$  — не фокальная точка, то индекс гессиана  $\phi + r$  в точке  $y$  — это число фокальных точек (с учетом их кратности) в сегменте от  $y$  до  $P$ . Доказать это мы оставляем читателю.

## Резюме

### А. Асимптотическое разложение интегралов

Для данной функции, определенной интегралом вида

$$\int e^{-xp(\omega)} d\omega,$$

надо уметь находить область значений  $\omega$ , дающую основной вклад в интеграл для больших  $x$ , и получать первые два члена асимптотического разложения интеграла.

### В. Стационарная фаза

Для интеграла вида

$$\int a(y) e^{ikp(y)} dy$$

надо уметь получать формулу для первых двух членов асимптотического разложения при больших  $k$ .

Методом стационарной фазы надо уметь получать формулу для обратного преобразования Фурье.

## Задачи

- 21.1. Модифицированная функция Бесселя третьего рода определяется интегралом

$$k_0(x) = \frac{1}{2} \int_{-\infty}^{\infty} e^{-x \cosh t} dt.$$

Получите первые два члена асимптотического разложения функции  $k_0(x)$ .

- 21.2. Функция Бесселя  $J_0(x)$  определяется интегралом

$$J_0(x) = \frac{1}{2\pi} \int_0^{2\pi} e^{ix \sin t} dt.$$

Методом стационарной фазы получите первые два члена асимптотического разложения  $J_0(x)$ .

- 21.3. Бета-функция  $B(x+1, x+1)$  представляется интегралом

$$I = \int_0^1 [t(1-t)]^x dt = \int_0^1 e^{x(\log t + \log(1-t))} dt.$$

- (a) Где набирается основной вклад в интеграл для очень больших  $x$ ?
  - (b) Получите главные члены асимптотического разложения  $B(x+1, x+1)$ .
  - (c) Опишите возможно подробнее, как получить следующие члены асимптотического разложения. Вычислите еще один член.
- 21.4. (a) Опишите процедуру получения асимптотического разложения интеграла вида

$$I = \int_{\theta_1}^{\theta_2} e^{-ixp(\theta)} q(\theta) d\theta.$$

Предположите, что  $p(\theta)$  имеет единственную критическую точку  $\theta = \alpha$  на отрезке  $[\theta_1, \theta_2]$ .

- (b) Примените эту процедуру для получения главного члена асимптотического разложения функции Бесселя

$$J_n(x) = \frac{1}{\pi} \left( \operatorname{Re} \int_0^\pi e^{ix \sin \theta} e^{-in\theta} d\theta \right),$$

где  $n$  — целое число.

## Глава 22

---

# Термодинамика

---

В этой главе показано, как внешнее дифференциальное исчисление используется в классической термодинамике, следуя идеям Борна и Каратеодори.

Тема этой главы — равновесная термодинамика, имеющая несколько ветвей. Первая из них — «классическая термодинамика» — имеет дело с понятиями тепла и работы, из которых следуют понятия энтропии, температуры и знаменитое «второе начало термодинамики». В классической термодинамике обсуждаются общие принципы, но не делается попыток описать поведение вещества или вычислять измеряемые величины в рамках микроскопической модели. Вторая ветвь — «равновесная статистическая механика» в классическом или квантовом варианте — дает явные формулы для всех «макроскопически наблюдаемых величин» равновесного состояния в рамках модели атомных или молекулярных взаимодействий. Эти формулы имеют общую природу, т. е. все они имеют общую форму и связаны с именами Максвелла, Больцмана и особенно Гиббса. Третья ветвь занимается объяснением, почему эти формулы работают, изучает проблемы с точки зрения теории вероятностей или динамики, или того и другого, занимается проблемами установления равновесия. В этой главе мы совсем немного будем говорить об этой третьей ветви. В первых двух параграфах рассматривается классическая термодинамика. Но при этом мы начнем с чисто математической теоремы Каратеодори, которая геометрически вполне наглядна, хотя ее доказательство во всех деталях мы завершим в приложении

---

к этой главе. Эта теорема является математическим инструментом, согласно которому в рамках подхода Борна и Каратеодори к термодинамике функция энтропии существует. После этого мы получим ряд физических следствий теории, а конец главы посвятим равновесной статистической механике.

## 22.1. Теорема Каратеодори

Пусть  $\alpha = A_1 dx^1 + \dots + A_n dx^n$  — линейная дифференциальная форма в пространстве  $\mathbb{R}^n$ , где все  $A_i$  — вещественные функции на этом пространстве. Возьмем некоторую фиксированную точку  $\mathbf{P}$  и будем рассматривать значение  $\alpha$  в точке  $\mathbf{P}$  как вектор-строку  $(A_1(\mathbf{P}), \dots, A_n(\mathbf{P}))$ . Если этот вектор не равен нулю, то его нулевое пространство — это  $(n-1)$ -мерное пространство всех

$\mathbf{X} = \begin{pmatrix} X^1 \\ \vdots \\ X^n \end{pmatrix}$  таких, что

$$A_1(\mathbf{P})X^1 + \dots + A_n(\mathbf{P})X^n = 0.$$

Пусть  $\gamma = x(t)$  — кусочно-дифференцируемая кривая. Тогда

$$\int_{\gamma} \alpha = \int [A_1(x(t))\dot{x}^1(t) + \dots + A_n(x(t))\dot{x}^n(t)] dt.$$

В частности, если для каждого  $t$  касательный вектор  $\dot{\mathbf{x}}(t) = \begin{pmatrix} \dot{x}^1(t) \\ \vdots \\ \dot{x}^n(t) \end{pmatrix}$  лежит в нулевом пространстве  $(A_1(x(t)), \dots, A_n(x(t)))$ ,

то интеграл  $\int_{\gamma} \alpha = 0$ . Кривая  $\gamma$  называется *нулевой кривой*  $\alpha$ , если  $\gamma$  непрерывна и кусочно-дифференцируема и если при каждом  $t$ , для которого определена  $\dot{\mathbf{x}}(t)$ , эта  $\dot{\mathbf{x}}(t)$  лежит в нулевом пространстве  $(A_1(x(t)), \dots, A_n(x(t)))$ . Сейчас мы хотим рассмотреть геометрическую задачу.

Предположим, что мы начали с некоторой точки  $\mathbf{P}$ . Где находятся те точки  $\mathbf{Q}$ , которые соединяются нулевыми кривыми с  $\mathbf{P}$ ?

Например, предположим, что для некоторой функции  $f$  выполняется равенство

$$\alpha = df.$$

Тогда если  $\gamma$  — нулевая кривая, соединяющая  $\mathbf{P}$  и  $\mathbf{Q}$ , мы имеем

$$0 = \int_{\gamma} \alpha = f(\mathbf{Q}) - f(\mathbf{P}).$$

Тогда  $f(\mathbf{Q}) = f(\mathbf{P})$ . Если  $df \neq 0$ , то уравнение  $f = f(\mathbf{P}) = \text{const}$  задает  $(n-1)$ -мерную поверхность<sup>1</sup> (рис. 22.1), проходящую через точку  $\mathbf{P}$ , и  $\mathbf{Q}$  должна быть на этой поверхности.

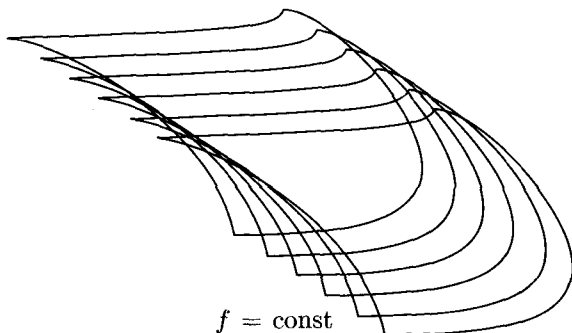


Рис. 22.1

При этом существуют точки, сколь угодно близкие к  $\mathbf{P}$ , которые не могут быть соединены с  $\mathbf{P}$  посредством нулевых кривых.

То же самое верно и в случае, когда  $\alpha = g df$ , где  $g$  — некоторая ненулевая функция, и  $df \neq 0$ . В самом деле,  $\gamma$  — нулевая кривая для  $\alpha$ , если и только если она является нулевой кривой для  $g^{-1}\alpha = df$ . И условия, чтобы быть нулевой кривой, те же. Поэтому, чтобы можно было соединить точки  $\mathbf{Q}$  и  $\mathbf{P}$  нулевой кривой,  $\mathbf{Q}$  должна лежать на  $(n-1)$ -мерной поверхности  $f = f(\mathbf{P})$ . Заметим, что если  $\alpha = g df$ , то  $d\alpha = dg \wedge df$  и поэтому

$$\alpha \wedge d\alpha = g df \wedge dg \wedge df = 0.$$

<sup>1</sup>В приложениях, где форма  $\alpha = df$  интерпретируется как силовое поле, а функция  $f$  — как его потенциал, такая поверхность называется эквипотенциальной. — Прим. ред.

А теперь рассмотрим форму

$$\alpha = dz + x dy,$$

определенную в пространстве  $\mathbb{R}^3$ . В этом случае  $d\alpha = dx \wedge dy$  и

$$\alpha \wedge d\alpha = dx \wedge dy \wedge dz$$

нигде не обращается в нуль. Покажем, что любые две точки можно соединить нулевой кривой для  $dz + x dy$ . Для этого достаточно показать, что можно соединить начало координат  $\mathbf{0}$  с любой другой точкой  $\mathbf{Q}$ . Напишем

$$\mathbf{Q} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}.$$

Сначала предположим, что  $b \neq 0$ . Теперь видно, что ось  $x$  является нулевой кривой для  $\alpha$ , потому что  $dz$  и  $dy$  равны нулю вдоль прямой  $z = y = 0$ . В силу этих же причин, любая прямая, параллельная оси  $x$ , является нулевой кривой для  $\alpha$ . Теперь пусть за время  $0 \leq t \leq 1$  мы сдвигаемся вдоль оси  $x$  от начала  $\mathbf{0}$  до точки,  $x$ -я координата которой равна  $-c/b$ . Далее двигаемся из точки

$$\begin{pmatrix} -c/b \\ 0 \\ 0 \end{pmatrix} \quad \text{в точку} \quad \begin{pmatrix} -c/b \\ b \\ c \end{pmatrix}$$

вдоль

$$\mathbf{x}(t) = \begin{pmatrix} -c/b \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ (t-1)b \\ (t-1)c \end{pmatrix}, \quad 1 \leq t \leq 2.$$

Поскольку

$$\dot{\mathbf{x}}(t) = \begin{pmatrix} 0 \\ b \\ c \end{pmatrix} \quad \text{и} \quad (A_1(\mathbf{x}(t)), A_2(\mathbf{x}(t)), A_3(\mathbf{x}(t))) = \left(0, -\frac{c}{b}, 1\right),$$

мы видим, что это тоже нулевая кривая. Наконец, перейдем из точки

$$\begin{pmatrix} -c/b \\ b \\ c \end{pmatrix} \quad \text{в точку} \quad \begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

вдоль

$$\mathbf{x}(t) = \begin{pmatrix} -c/b \\ b \\ c \end{pmatrix} + \begin{pmatrix} (t-2)(a-c/b) \\ 0 \\ 0 \end{pmatrix}, \quad 2 \leq t \leq 3.$$

Поскольку эта прямая параллельна оси  $x$ , она тоже является нулевой кривой. Итак, мы перешли из точки  $\mathbf{0}$  в точку  $\mathbf{Q}$  вдоль непрерывной траектории — ломаной, состоящей из трех прямолинейных отрезков. Эта траектория является кусочно-дифференцируемой нулевой кривой.

Если  $b = 0$  и  $c \neq 0$ , то сделаем немного по-другому. Сначала перейдем вдоль нулевой кривой в точку  $\begin{pmatrix} 1 \\ c \\ 0 \end{pmatrix}$ . Сначала идем в  $\begin{pmatrix} 0 \\ c \\ 0 \end{pmatrix}$  вдоль  $x = z = 0$ , а затем в  $\begin{pmatrix} 1 \\ c \\ 0 \end{pmatrix}$  вдоль  $y = c, z = 0$  (параллельно оси  $x$ ). Наконец, перейдем из точки

$$\begin{pmatrix} 1 \\ c \\ 0 \end{pmatrix} \quad \text{в точку} \quad \begin{pmatrix} 1 \\ 0 \\ c \end{pmatrix}$$

вдоль

$$\mathbf{x}(t) = \begin{pmatrix} 1 \\ c \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ -(t-3)c \\ (t-3) \end{pmatrix}, \quad 2 \leq t \leq 3.$$

Поскольку вдоль этой кривой  $dz = c dt$ ,  $dy = -c dt$  и  $x \equiv 1$ , мы видим, что  $sz + x dy$  тождественно равно нулю на этой кривой, поэтому она — нулевая кривая. Теперь можно перейти в точку  $\mathbf{Q}$  вдоль прямой, параллельной оси  $x$ , т.е. координаты  $y$  и  $z$  фиксированы, а координата  $x$  меняется от 1 до  $a$ .

Конечно, если  $b = c = 0$ , можно сразу попасть из  $\mathbf{0}$  в точку  $\mathbf{Q}$  вдоль оси  $x$ . Итак, мы доказали, что любые две точки в пространстве  $\mathbb{R}^3$  можно соединить нулевой кривой формы  $dz + x dy$ .

В пространстве  $\mathbb{R}^3$  можно рассмотреть три типа 1-форм.

Тип	Пример
(i) $\alpha$ , где $\alpha \neq 0$ , но $d\alpha \equiv 0$	$\alpha = dz$
(ii) $\alpha$ , где $d\alpha \neq 0$ , но $\alpha \wedge d\alpha \neq 0$	$\alpha = x dy$
(iii) $\alpha$ , где $\alpha \wedge d\alpha \neq 0$	$\alpha = dz + x dy$

В случае (i) мы знаем, что если  $d\alpha \equiv 0$ , то локально можно найти такую функцию  $f$ , что  $\alpha = df$ . (Это можно сделать в любой звездной области, содержащейся в области определения  $\alpha$ .) Если  $\alpha$  не обращается в нуль в некоторой точке  $\mathbf{P}$ , то из теоремы о неявной функции следует, что можно сделать такую замену переменной, что функция  $f$  станет одной из координат, скажем,  $z$ . Другими словами, с точностью до замены переменной пример из первой строки представляет ее общий случай. Математическая теорема, которую мы здесь приведем, говорит, что это справедливо и в двух других случаях, т.е. любая форма типа (ii) или (iii) сводится к соответствующему примеру. Позднее в этом параграфе мы дадим формулировку теоремы и наметим вкратце, как ее доказать, а все детали доказательства дадим в приложении. Теорема и ее доказательство верны не только в трехмерном пространстве, но и в пространстве  $n$  измерений. Мы уже видели, что в случае (iii) можно соединить любую точку с любой близлежащей с помощью нулевой кривой формы  $\alpha$ . Следовательно, если форма  $\alpha$  такова, что для любой точки  $\mathbf{P}$  существуют точки  $\mathbf{Q}$ , сколь угодно близкие к ней, которые нельзя соединить с  $\mathbf{P}$  посредством нулевой кривой формы  $\alpha$ , то имеет место случай (i) или (ii). Это утверждение, точнее его  $n$ -мерное обобщение, и есть содержание теоремы Каратеодори.

**Теорема Каратеодори.** Пусть  $\alpha$  — линейная дифференциальная форма, обладающая тем свойством, что для любой точки  $\mathbf{P}$  существуют сколь угодно близкие к ней точки  $\mathbf{Q}$ , которые не могут быть соединены с  $\mathbf{P}$  с помощью нулевой кривой формы  $\alpha$ . Тогда (локально) существуют такие функции  $f$  и  $g$ , что

$$\alpha = f dg.$$

Это выражение для  $\alpha$  не единственное. Действительно, если  $H$  — дифференцируемая функция одной переменной, производная которой нигде не равна нулю, и

$$G = H \circ g,$$

то, согласно цепному правилу,  $dG = H'(g) dg$ . Поэтому можно писать

$$\alpha = F dG, \quad \text{где} \quad G = H \circ g, \quad F = f/H'(g).$$

Прежде чем продолжать чисто математическое обсуждение теоремы Каратеодори, давайте кратко опишем, как она используется в термодинамике. Детальное обсуждение с точными определениями читатель найдет в следующем параграфе.

Можно измерять количество тепла, выделенное в химической реакции, которая *проводилась в известных условиях*, например, при постоянном *объеме* или при постоянном *давлении*. (Это можно сделать, например, в умном приборе, называемом ледяным калориметром. Он использует тот факт, что объем льда при таянии сокращается. При этом в обычных условиях, чтобы растаяло данное количество льда, требуется определенное количество тепла.) Количество тепла, необходимое для небольшого изменения какой-нибудь величины (например, повысить температуру одного грамма воды на один градус или увеличить его объем), когда эксперимент выполняется каким-то заданным способом, (приблизительно) линейно связано с этими малыми изменениями. Поэтому «количество тепла, сообщенное» системе в равновесном состоянии, является линейной дифференциальной формой  $\alpha$ . В течение довольно долгого времени было принято считать, что эта форма точная, т.е. что существует такая функция  $C$  (называемая калорической), представляющая «полное количество тепла в системе», для которой  $\alpha = dC$ . Другими словами, мы считаем, что «калорическая функция» в системе изменяется на величину сообщенного тепла. Только постепенно ученые осознали, что форма  $\alpha$  не замкнута и поэтому не может описывать бесконечно малое изменение какой-либо функции.

С другой стороны, довольно рано стало понятно, что работа, совершенная над системой, тоже задается линейной дифференциальной формой  $\omega$ , которая не замкнута. Действительно, пусть

в цилиндре находится газ, тогда при сжатии его поршнем совершается работа тем большая, чем больше давление газа и соответственно выше его температура. Вдавим поршень в цилиндр при высокой температуре, сокращая объем, и потом охладим систему в этом положении (поршень неподвижен, и работа не совершается), при низкой температуре вернем поршень в исходное положение и нагреем газ до той же температуры, что и вначале. Таким образом, мы завершили «цикл», т. е. мы прошли вдоль замкнутой кривой, вдоль которой интеграл от  $\omega$  не равен нулю. (Конечно, эффективный способ проверить, замкнута форма или нет, — это проинтегрировать вдоль замкнутой кривой. Вот почему «циклы» играют такую важную роль при изучении задач термодинамики.)

Согласно *первому закону термодинамики*<sup>2</sup>, несмотря на то, что формы  $\alpha$  и  $\omega$  не замкнуты, их сумма замкнута, т. е.  $d(\alpha + \omega) = 0$ . Другими словами, первый закон термодинамики говорит, что (локально)

$$\alpha + \omega = dU,$$

где  $U$  — функция системы (определена с точностью до аддитивной постоянной). Эта функция называется *внутренней энергией* системы. Существование функции  $U$  — это один из вариантов закона сохранения энергии.

Второй закон термодинамики следует из поведения системы, к которой не подводится тепла извне. Если мы поместим систему в оболочку (называемую адиабатической оболочкой), через которую невозможно обмениваться теплом, то состояние системы сможет изменяться только по нулевым кривым формы  $\alpha$ . В природе замечено, что в этих условиях для любого состояния системы существуют близкие состояния, в которые невозможно перейти. Согласно теореме Каратеодори, это значит, что существуют такие функции  $f$  и  $g$ , для которых

$$\alpha = f dg.$$

---

<sup>2</sup>К двум основным принципам термодинамики в русскоязычной литературе часто применяется название «начало» вместо «закон». — *Прим. ред.*

Конечно, это условие не определяет  $f$  и  $g$  полностью, о чем мы уже говорили выше. Однако, анализ обмена между системами, находящимися в тепловом контакте, показывает, что в качестве функции  $f$  можно взять универсальную функцию  $T$  для всех систем, называемую *абсолютной температурой*, которая определяется с точностью до постоянного коэффициента (масштабного множителя). Это значит, что при фиксированном масштабе абсолютной температуры для каждой системы функция  $g$  определяется с точностью до аддитивной постоянной. Обычно эту функцию обозначают буквой  $S$  и называют *энтропией* системы. Итак, согласно *второму закону термодинамики*, существует универсальная температура  $T$  (называемая абсолютной температурой и определенная с точностью до постоянного множителя), и для каждой системы существует функция  $S$  (называемая энтропией и определенная с точностью до аддитивной постоянной при фиксированном масштабе температур), для которых

$$\alpha = T dS.$$

Температурную функцию выбирают так, чтобы она всегда была положительной. При этом условии изменение энтропии всегда имеет тот же знак, что и количество тепла, сообщенное системе вдоль любого касательного вектора системы (т. е. бесконечно малое изменение). Поскольку изменение  $S$  вдоль любой замкнутой траектории должно быть равно нулю, это означает, что для любого цикла, если где-то добавляется тепло, в другом месте оно должно изыматься. Кельвин дал известную формулировку второго закона термодинамики: «не существует циклических процессов, в которых взятое у системы тепло полностью переходит в работу». Работа может производиться только с помощью тепловой машины, переводящей определенное количество тепла в холодный резервуар.

В следующем параграфе мы сюда вернемся и дадим точные формулировки всех этих законов. А сейчас обсудим чисто математически теорему Каратеодори. Давайте рассмотрим форму

$$\alpha = dz + x dy$$

в шестимерном пространстве  $\mathbb{R}^6$  с координатами  $x, y, z, u, v, w$ . Очевидно, что любая кривая, вдоль которой постоянны  $y$  и

$z$ , является нулевой кривой формы  $\alpha$ . Поэтому мы можем двигаться из  $\mathbf{0}$  в любую точку  $\mathbf{Q}$  пространства  $\mathbb{R}^6$  вдоль нулевой кривой формы  $\alpha$ , сначала двигаясь в трехмерном пространстве  $u = v = w = 0$  до точки, у которой первые три координаты совпадают с соответствующими координатам точки  $\mathbf{Q}$ , а затем уже вдоль прямой линии до точки  $\mathbf{Q}$ . (Вдоль этой прямой  $x$ ,  $y$  и  $z$  постоянные.) А теперь рассмотрим форму

$$\alpha = dz + x dy + u dv.$$

Как и раньше, в пространстве, где  $u = v = w = 0$ , мы двигаемся из начала координат до точки, первые три координаты которой совпадают с соответствующими координатами точки  $\mathbf{Q}$ . Затем, сохраняя значения  $x$  и  $y$ , мы можем изменить нужным образом координаты  $u$  и  $v$ , потому что вдоль любой кривой  $y = \text{const}$  форма  $\alpha$  принимает те же значения, что и форма  $dz + u dv$ , и мы, таким образом, сводим к случаю трехмерного пространства. Наконец, можно изменить координату  $w$ , фиксируя остальные пять координат.

Далее рассмотрим форму

$$\alpha = w dz + x dy + u dv.$$

Предположим, что у точки  $\mathbf{Q}$  координата  $w$  не равна нулю, т.е.  $w(\mathbf{Q}) = r \neq 0$ . Сначала двигаемся от точки  $\mathbf{0}$  вдоль прямой  $x = y = z = u = v = 0$  до точки, у которой координата  $w$  равна  $r$ . Теперь мы остаемся в гиперплоскости  $w = r$ . Но для любой кривой в этой гиперплоскости форма  $\alpha$  принимает те же значения, что и форма

$$\alpha' = r dz' + x dy + u dv.$$

Если мы сделаем замену переменной  $z' = rz$ , то

$$\alpha' = dz' + x dy + u dv.$$

Теперь, как и раньше, можно придать нужные значения остальным пяти переменным, используя трехмерный случай. Аналогично, если координата  $x$  или  $u$  точки  $\mathbf{Q}$  не равна нулю, то используются те же соображения, просто изменяется роль различных координат.

С другой стороны, если точка  $Q$  находится в трехмерном подпространстве, заданным условием  $x = u = w = 0$ , то любая кривая в этом подпространстве является нулевой кривой. Поэтому можно соединить начало координат с точкой  $Q$ . Итак, в пространстве  $\mathbb{R}^6$  можно рассмотреть шесть типов 1-форм.

Тип	Пример
(i) $\alpha \neq 0$ , но $d\alpha \equiv 0$	$dz$
(ii) $d\alpha \neq 0$ , но $\alpha \wedge d\alpha \equiv 0$	$x dy$
(iii) $\alpha \wedge d\alpha \neq 0$ , но $d\alpha \wedge d\alpha \equiv 0$	$dz + x dy$
(iv) $d\alpha \wedge d\alpha \neq 0$ , но $\alpha \wedge d\alpha \wedge d\alpha \equiv 0$	$x dy + u dv$
(v) $\alpha \wedge d\alpha \wedge d\alpha \neq 0$ , но $d\alpha \wedge d\alpha \wedge d\alpha \equiv 0$	$dz + x dy + u dv$
(vi) $d\alpha \wedge d\alpha \wedge d\alpha \neq 0$	$w dz + x dy + u dv$

Наши рассуждения показывают, что в случае формы типов (iii)–(vi) можно двигаться между двумя любыми точками вдоль нулевой кривой. Имеется математическая теорема *о нормальной форме*, согласно которой для любой из форм одного из типов, стоящих слева, можно сделать локальную замену координат, так что в новых координатах форма  $\alpha$  принимает соответствующий вид, указанный справа. (Набросок доказательства этой теоремы дается в приложении.) С учетом этого факта мы можем доказать теорему Каратеодори в шестимерном пространстве следующим образом.

Сначала покажем, что

$$d\alpha \wedge d\alpha \wedge d\alpha \equiv 0.$$

Предположим противное — выражение  $d\alpha \wedge d\alpha \wedge d\alpha$  не равно нулю тождественно. Тогда в некоторой точке  $P$  (и поэтому в окрестности  $P$ ) мы имеем дело со случаем (vi). Следовательно, в окрестности точки  $P$  можно сделать замену координат, так чтобы форма  $\alpha$  имела вид, стоящий справа. Но тогда точку  $P$  можно соединить с любой точкой с помощью нулевой кривой формы  $\alpha$ . Но ведь мы предположили, что *каждая* точка имеет в своей окрестности точки, в которые нельзя попасть с помощью нулевых кривых. Противоречие! Поэтому  $d\alpha \wedge d\alpha \wedge d\alpha \equiv 0$ . Но эти же рассуждения показывают, что

$$\alpha \wedge d\alpha \wedge d\alpha \equiv 0,$$

поскольку иначе мы будем иметь дело со случаем  $(v)$  вблизи некоторой точки и так далее. Таким образом, мы можем двигаться дальше и доказать, что  $\alpha \wedge d\alpha \equiv 0$ , а это значит, что теорема Каратеодори доказана. Аналогичные соображения работают в пространстве  $n$  измерений.

## 22.2. Классическая термодинамика по Борну и Каратеодори

В этом параграфе мы изложим основные положения термодинамики с точки зрения Борна и Каратеодори. Преимущество этого подхода состоит в том, что можно избежать довольно сложных соображений относительно цикла Карно и невозможности построить вечный двигатель. Этот подход основывается на абстрактных идеях, далеких от повседневного опыта, и ряде физических законов, которые просто формулировать. Здесь *тепло* — производное понятие. Теория формулируется в стандартных терминах элементарной механики. Однако, несмотря на то, что в теории содержатся понятия механики, ее законы должны быть модифицированы. Именно модификация этих законов приводит к понятиям температуры и тепла. После этого два простых физических закона вместе с теоремой Каратеодори приводят к понятиям абсолютной температуры и энтропии.

Основное предположение, общее для почти всех физических теорий, состоит в том, что мы можем выделить интересующую нас часть из остальной Вселенной. Эту часть Вселенной мы называем *системой*. Система может существовать в различных *состояниях*. Описание *всех* состояний системы может быть невероятно сложной задачей. Например, если наша система — газ в некотором объеме, то этот газ может совершать турбулентное движение. В этом случае для описания состояния газа мы должны будем определять локальную скорость каждой малой порции. Таким образом, следовало бы ожидать, что множество всех состояний газа может быть бесконечномерным в соответствующем смысле. В любом случае в нашей теории мы будем работать с неопределенными терминами — с «системой» и «состоянием». Позднее, когда мы

разработаем «модель» термодинамики, будут даны математические определения. Другой термин, также понятный интуитивно, но пока не определенный, — *взаимодействие*. Предполагается, что можно конкретизировать различные взаимодействия нашей системы с другими системами или со Вселенной.

Следующее фундаментальное предположение — среди всех возможных состояний системы существует выделенный класс состояний, называемых *равновесными*. Если все взаимодействия системы со Вселенной фиксированы, то система будет переходить из одного состояния в другое, стремясь при этом к определенному равновесному состоянию (оно определяется начальным состоянием системы в момент, когда были фиксированы все взаимодействия). Хотя произвольные состояния системы очень сложны, класс равновесных состояний относительно просто описывать, он может быть параметризован конечным числом переменных. Итак, множество равновесных состояний имеет структуру подмногообразия<sup>3</sup> конечномерного векторного пространства. Это *многообразие равновесных состояний* является подмножеством множества всех состояний системы.

Следующее предположение выделяет определенные типы взаимодействий. Существует особая форма взаимодействия между двумя системами, называемая *диатермальным контактом*. При взаимодействии такого типа не происходит наблюдаемого макроскопического движения и обмена веществом. Возникающие изменения можно описать следующим образом. Объединим две системы в одну новую систему (рис. 22.2). Эта составная система является прямым произведением двух первоначальных систем в том смысле, что состояния новой системы — это пара состояний  $(p_1, p_2)$ , где  $p_1$  — состояния первой системы, а  $p_2$  — состояния второй. При *диатермальном контакте* равновесные состояния объединенной системы — это подмножество множества пар равновесных состояний исходных систем (рис. 22.3). Другими словами, если  $p_1$  — равновесное состояние первой системы, а  $p_2$  — равновесное состояние второй системы, и обе системы находятся в диатермальном контакте, то не обязательно, что объединенная система будет в равновесном состоянии, но она будет стремиться

---

<sup>3</sup>См. обсуждение подмногообразия в параграфе 10.9.

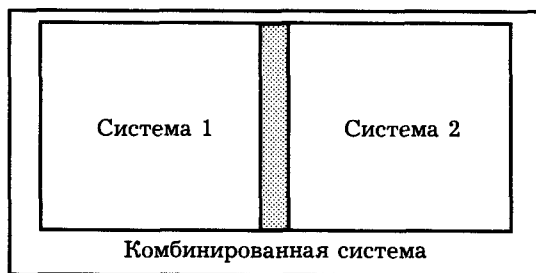


Рис. 22.2

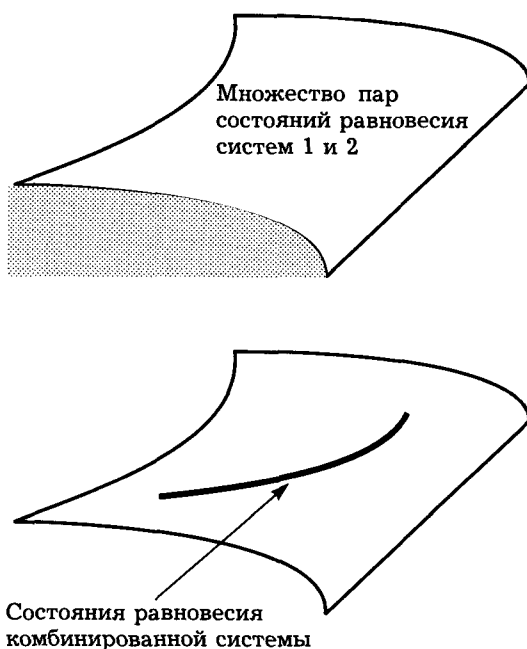


Рис. 22.3

к определенному равновесному состоянию  $(q_1, q_2)$ , где  $q_1$  — новое равновесное состояние первой системы, а  $q_2$  — новое равновесное состояние второй системы. Тогда мы говорим, что первая система в состоянии  $q_1$  находится в *тепловом равновесии* со второй системой в состоянии  $q_2$ .

Этот наблюдаемый закон природы иногда называется *нулевым законом термодинамики*. Согласно ему, *тепловое равновесие является отношением эквивалентности*. Это означает следующее. Пусть мы имеем три системы, причем  $p_1$ ,  $p_2$  и  $p_3$  — равновесные состояния этих систем. Предположим, что первые две системы в этих состояниях ( $p_1$  и  $p_2$ ), будучи приведены в диатермальный контакт, оказываются в тепловом равновесии, и то же самое справедливо для второй и третьей систем (в состояниях  $p_2$  и  $p_3$ ). Тогда первая и третья системы ( $p_1$  и  $p_3$ ) при диатермальном контакте также будут в тепловом равновесии.

Класс эквивалентности всех систем, находящихся в тепловом равновесии, называется *абстрактной температурой*. Предположим, что мы фиксируем одну определенную систему и некоторую численную функцию равновесных состояний этой системы, которая принимает различные значения в состояниях, принадлежащих различным абстрактным температурам<sup>4</sup>. Таким образом, у нас есть численная функция  $\theta$ , определенная для всех равновесных состояний всех систем, называемая *эмпирической температурой*. Выбранная система вместе с ее функцией называется *термометром*.

Кроме того, мы полагаем, что для любой системы  $\theta$  — дифференцируемая функция на множестве равновесных состояний. (Возможность выбрать дифференцируемую  $\theta$  — тоже часть «нулевого закона термодинамики».) Более того, можно найти некоторую систему и для нее выбрать такую функцию  $\theta$ , что  $d\theta \neq 0$  на всем множестве равновесных состояний этой системы. Только такие функции  $\theta$  мы будем использовать в качестве эмпирических температур. Конечно, если  $\theta'$  — другая эмпирическая температура, то она — строго монотонная дифференцируемая функция  $\theta$ . И обратно, взяв любую строго монотонную дифференцируемую функцию и применив ее к  $\theta$ , получим новую эмпирическую температуру  $\theta'$ . В любом случае два состояния имеют одинаковую абстрактную температуру, если и только если они имеют одинаковое значение функции  $\theta$  при любом выборе эмпирической температуры.

---

<sup>4</sup>Иначе говоря, множество абстрактных температур превращается в однопараметрическое семейство. Достаточность одного численного параметра — это еще одно нетривиальное допущение. — *Прим. ред.*

Существует ряд функций, которые играют важную роль при описании любой системы. Так например, если у нас есть газ, заключенный в некоторый сосуд, то такой функцией будет объем, а также массы различных химических составляющих газа. Эти функции называются *конфигурационными переменными*. Таким образом, каждая система имеет  $n$  функций, определенных на множестве ее состояний и называемых конфигурационными переменными. Число конфигурационных переменных  $n$  зависит от системы, но всегда  $n \geq 1$ . Если рассматривать только равновесные состояния, то конфигурационные переменные являются дифференцируемыми функциями. Конфигурационные переменные и функция  $\theta$  образуют систему локальных координат на множестве  $(n + 1)$ -мерных равновесных состояний. Конфигурационные переменные обозначим  $x_1, x_2, \dots, x_n$ . Тогда  $\theta, x_1, x_2, \dots, x_n$  — локальная система координат на множестве равновесных состояний. Во многих случаях одной из конфигурационных переменных является объем  $V$ . Обычно берут  $x_1 = V$ , так что  $\theta, V, x_2, \dots, x_n$  — локальные координаты, где  $x_2, \dots, x_n$  — остальные конфигурационные переменные.

Поскольку  $d\theta$  нигде не обращается в нуль, то в силу теоремы о неявной функции множества  $\theta = \text{const}$  образуют  $n$ -мерные подмногообразия. Подмногообразия  $\theta = \text{const}$  называются *изотермами*. Конфигурационные переменные можно использовать как локальные координаты для изотерм.

Существует другой важный класс взаимодействий системы с остальной Вселенной, которые называются *адиабатическими взаимодействиями*. Для таких взаимодействий равновесие нарушается только при изменении конфигурационных переменных. Если все взаимодействия системы в наблюдаемый период адиабатические, то такую систему называют *адиабатически замкнутой*. Примером адиабатически замкнутой системы (приблизительно) является термос.

Кривая  $\gamma$ , соединяющая состояния  $\rho$  и  $\rho'$  адиабатически замкнутой системы, называется *адиабатической кривой*, или *адиабатой*. Обобщение экспериментального опыта — *первый закон термодинамики* — говорит нам, что адиабатически замкнутая система переходит из одного состояния  $\rho$  в другое  $\rho'$  за счет совершения внешней работы, величина которой не зависит от спо-

соба ее приложения. Поэтому на пространстве состояний этой системы существует функция  $U$ , называемая *внутренней энергией*. Она обладает тем свойством, что количество работы, совершенной над системой при переходе из состояния  $\rho$  в состояние  $\rho'$  вдоль любой адиабатической кривой, равно  $W = U(\rho') - U(\rho)$ .

Если система не является адиабатически замкнутой и за счет взаимодействия она переходит из одного состояния  $\rho$  в другое  $\rho'$  вдоль некоторой кривой, то количество внешней работы уже не равняется  $U(\rho) - U(\rho')$ . Разность  $U(\rho) - U(\rho') - W = Q$  называется *теплом*, сообщенным системе в данном процессе. В этом случае количество совершенной работы и тепла, сообщенного системе, зависят от процесса, а не только от начального и конечного состояний  $\rho$  и  $\rho'$ .

Это — формулировка первого закона термодинамики в рамках подхода Борна–Каратеодори. Обратите внимание, что при таком подходе тепло *определяется* как разность между изменением внутренней энергии и работой, совершенной над системой. Идея рассматривать тепло не как фундаментальную, а как производную величину, является одновременно и сильной, и слабой стороной подхода Борна–Каратеодори. Сила заключается в том, что здесь отражается фундаментальный взгляд в физике, существующий более века, — основным объектом является энергия, которая сохраняется с учетом всех ее форм. А слабость — в том, что этот подход не отражает экспериментальные и исторические реалии. Обычно тепло, подводимое к системе, легче всего измерять. Джоуль провел в 1840 году основополагающие эксперименты с целью получить *механический эквивалент тепла*. Если система переходит в некоторое другое состояние за счет определенного количества тепла, подведенного к ней, то такое же изменение системы может произойти, если над ней совершить определенную работу. Но полное количество совершенной работы не зависит от того, как эта работа была выполнена. Поэтому в реальных экспериментах измеряется именно подведенное количество тепла и совершается разного рода работа. Из существования единственного механического эквивалента тепла мы делаем *вывод*, что полная работа для перевода системы из одного состояния в другое не зависит от способа выполнения этой работы, при условии, что она выполняется адиабатически.

Следующее фундаментальное понятие теории — обратимая кривая. Кривая  $\gamma$  называется (почти) *обратимой*, если в каждый момент времени  $t$  состояние  $\gamma(t)$  равновесное (достаточно близко к равновесному).

Интуитивно понятие обратимости состоит в том, что все взаимодействия в системе происходят очень медленно по сравнению со «временем релаксации» системы, т. е. со временем, необходимым для восстановления равновесия в системе<sup>5</sup>. Например, предположим, что в цилиндре с поршнем находится газ. Взаимодействие — это движение поршня, изменяющее объем газа. Если газ находился в равновесии, и мы внезапно сдвинули поршень, то новое состояние газа уже не будет равновесным. В нем будут разного рода ударные волны и вихри. Конечно, если поршень оставить в этом положении, то в цилиндре установится со временем новое равновесие. Можно полагать, что скорость возвращения в равновесное состояние достаточно велика, так что если поршень двигать достаточно медленно, то состояния системы будут все время (приблизительно) равновесными. Такая кривая состояний называется *обратимой*. В случае изменения объема из опыта известно, что существует функция  $p$ , называемая давлением, такая, что работа, совершенная над системой вдоль любой обратимой кривой  $\gamma$ , задается выражением  $-\int_{\gamma} p dV$ .

Заметим, что до сих пор мы измеряли работу через изменение, внешнее по отношению к системе. Давление — функция равновесных состояний, позволяющая нам вычислить совершенную работу посредством обратимого изменения объема. Замечено, что существуют аналогичные функции, соответствующие другим конфигурационным переменным, т. е. в пространстве равновесных состояний существуют такие функции  $\nu_1, \dots, \nu_n$ , что работа, совершенная над системой вдоль любой (почти) обратимой кривой  $\gamma$ , равна (приблизительно) интегралу вдоль  $\gamma$  от линейной дифференциальной формы

$$\omega = \nu_1 dx_1 + \dots + \nu_n dx_n,$$

<sup>5</sup>Такой процесс изменения состояния системы называют *квазистатическим*. — Прим. ред.

так что

$$W(\gamma) = \int_{\gamma} \omega.$$

Поэтому тепло, подведенное к системе вдоль пути  $\gamma$ , задается интегралом

$$Q(\gamma) = \int_{\gamma} \alpha,$$

где  $\alpha$  — линейная дифференциальная форма  $\alpha = dU - \omega$ . В случае, когда  $x_1 = V$  — объем, принято называть  $p = -\nu_1$  давлением. Следовательно,

$$\omega = -p dV + \nu_2 dx_2 + \cdots + \nu_n dx_n$$

и

$$\alpha = dU - \omega = dU + p dV - \nu_2 dx_2 - \cdots - \nu_n dx_n.$$

В частности, по определению, *адиабатическая обратимая кривая* — это кривая на множестве равновесных состояний, которая является нулевой кривой для линейной дифференциальной формы  $\alpha$ .

Итак, мы подошли к знаменитому «второму закону термодинамики». Каждодневный опыт говорит, что определенные типы работы, выполняемые над системой, не могут быть возвращены. Точнее, для адиабатически замкнутой системы существуют определенные типы работ, совершаемые над системой, такие как сильное встряхивание или размешивание, когда нельзя получить эту работу обратно вдоль какой-либо адиабатической кривой. Утверждается, что это может происходить вблизи любого равновесного состояния. Итак, формулировка *второго закона термодинамики* такова.

В окрестности любого равновесного состояния  $\rho$  системы существуют сколь угодно близкие равновесные состояния, которые нельзя соединить с  $\rho$  с помощью обратимых адиабатических кривых.

Как мы увидим в дальнейшем, это простое утверждение имеет далеко идущие следствия. Закон говорит, что в окрестности любой точки множества равновесных состояний существуют произвольно близкие точки, в которые нельзя попасть вдоль нулевых

кривых формы  $\alpha$ . Из теоремы Каратеодори следует, что для любой системы существуют такие функции  $\lambda$  и  $\phi$  на множестве равновесных состояний, что  $\alpha = \lambda d\phi$ .

Конечно, для любой заданной системы ни функция  $\lambda$ , ни функция  $\phi$  не определяются полностью равенством  $\alpha = \lambda d\phi$ . Однако, если выбрать функцию  $\lambda$ , то  $\phi$  определяется с точностью до аддитивной константы. Мы покажем в следующем параграфе, что из рассмотрения поведения формы  $\alpha$  при объединении систем можно сделать вывод, что существует предпочтительный выбор температурной функции  $T$  (определенной с точностью до постоянного множителя), при этом для всех систем можно положить  $\lambda = T$ . Эта функция  $T$  называется *абсолютной температурой*. В терминах некоторой выбранной эмпирической температуры  $\theta$  утверждается, что существует такая универсальная функция  $T = T(\theta)$ , что для всех систем можно выбрать  $\lambda = T(\theta)$ .

Если мы решили использовать для нашей температурной шкалы абсолютную температуру, то для любой системы можно написать

$$\alpha = T dS$$

где функция  $S$  определена с точностью до аддитивной константы. Эта функция  $S$  называется *энтропией*.

### 22.3. Энтропия и абсолютная температура

Предположим, что две системы находятся в диатермальном контакте. Можно также считать, что энергия, необходимая для объединения этих систем, пренебрежимо мала. Между системами нет сдвигаемых перегородок и нет прямого обмена веществом. Тогда полная внутренняя энергия  $U$  объединенной системы равна сумме энергий состояний отдельных систем

$$U(\rho_1, \rho_2) = U_1(\rho_1) + U_2(\rho_2)$$

для любого состояния объединенной системы  $(\rho_1, \rho_2)$ . Здесь  $U$  обозначает функцию внутренней энергии объединенной системы,  $U_1$  — внутреннюю энергию первой системы, а  $U_2$  — внутреннюю энергию второй системы.

Дифференциальная форма, описывающая работу, тоже аддитивна:

$$\omega = \omega_1 + \omega_2.$$

Здесь  $\omega_1$  — линейная дифференциальная форма, определенная на множестве равновесных состояний первой системы. Она может рассматриваться как линейная дифференциальная форма, определенная на произведении пространств, состоящем из всех пар равновесных состояний. При этом на  $\omega_1$  не влияют равновесные состояния второй системы. Аналогично на множестве таких пар равновесных состояний определена форма  $\omega_2$ , которая не зависит от состояния первой системы. Следовательно, форма  $\omega_1 + \omega_2$  определена на множестве всех пар равновесных состояний. Пространство равновесных состояний объединенной системы, находящейся в тепловом контакте, есть подмногообразие множества всех пар (определяемое условием  $\theta_1 = \theta_2$ , см. рис. 22.4). Это значит, что  $\omega$  является ограничением формы  $\omega_1 + \omega_2$  на это подмногообразие. Отсюда следует, что (в том же смысле)

$$\alpha = \alpha_1 + \alpha_2.$$

Написав

$$\alpha = \lambda d\phi, \quad \alpha_1 = \lambda_1 d\phi_1, \quad \alpha_2 = \lambda_2 d\phi_2,$$

получаем

$$\lambda d\phi = \lambda_1 d\phi_1 + \lambda_2 d\phi_2. \quad (*)$$

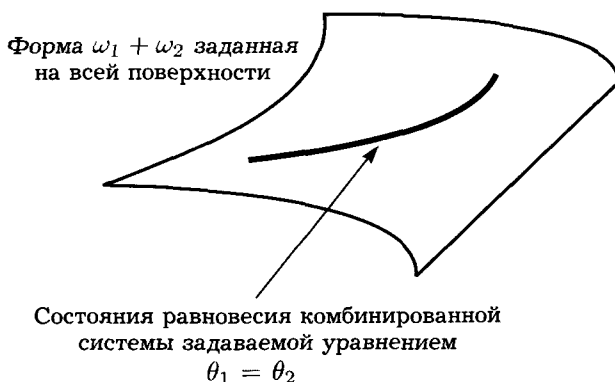


Рис. 22.4

Для любой системы температуру можно изменять адиабатически (например, изменяя  $V$  при постоянном давлении). Это означает, что для первой системы формы  $\alpha_1$  и  $d\theta_1$  линейно независимы; и аналогично для второй системы. Условие  $\alpha_1 = \lambda_1 d\phi_1$  означает, что дифференциальные формы  $d\phi_1$  и  $d\theta_1$  линейно независимы. В силу теоремы о неявной функции, это значит, что можно сделать замену переменных, так что  $\theta_1$  и  $\phi_1$  станут первыми координатами, т. е.  $(\theta_1, \phi_1, y_1, \dots)$  — локальные координаты на множестве равновесных состояний первой системы. Аналогично для второй системы. Поэтому в качестве локальных координат на множестве пар равновесных состояний можно взять

$$(\theta_1, \theta_2, \phi_1, \phi_2; y_1, \dots; z_2, \dots).$$

Равновесные состояния объединенной системы (в тепловом контакте) задаются условием  $\theta_1 = \theta_2$ . Поэтому можно взять переменные

$$(\theta, \phi_1, \phi_2; y_1, \dots; z_2, \dots)$$

в качестве координат на множестве равновесных состояний объединенной системы. Уравнение (\*) показывает, что в этих координатах  $d\phi$  является линейной комбинацией  $d\phi_1$  и  $d\phi_2$ , т. е.  $\phi = F(\phi_1, \phi_2)$ . Другими словами, функция  $\phi$  объединенной системы зависит только от значений функций  $\phi_1$  и  $\phi_2$  и не зависит от остальных координат. К любой системе в любом состоянии можно подвести тепло, поэтому форма  $\alpha$  и, следовательно,  $\lambda$  нигде не обращается в нуль. Поэтому уравнение (\*) можно разделить на  $\lambda$  и написать

$$d\phi = (\lambda_1/\lambda)d\phi_1 + (\lambda_2/\lambda)d\phi_2. \quad (**)$$

Это значит, что

$$\lambda_1/\lambda = f_1(\phi_1, \phi_2), \quad \text{где} \quad f_1 = \partial F / \partial \phi_1,$$

зависит только от  $\phi_1$  и  $\phi_2$ , и аналогично

$$\lambda_2/\lambda = f_2(\phi_1, \phi_2), \quad \text{где} \quad f_2 = \partial F / \partial \phi_2.$$

Первое из этих уравнений можно разделить на второе, поскольку ни одна из функций  $\lambda_1$ ,  $\lambda_2$  и  $\lambda$  не обращается в нуль. Получаем

$$\lambda_1/\lambda_2 = f_1/f_2$$

или

$$\log \lambda_1 - \log \lambda_2 = \log(f_1/f_2),$$

где правая часть зависит только от  $\phi_1$  и  $\phi_2$ . В частности, если вычислить частную производную этого равенства по  $\theta$ , то правая часть обратится в нуль, и мы получим

$$\frac{\partial \log \lambda_1}{\partial \theta} = \frac{\partial \log \lambda_2}{\partial \theta}.$$

Слева стоит функция от переменных, описывающих первую систему, т. е. от  $(\theta, \phi_1, y_1, \dots)$ , а справа стоит функция от переменных  $(\theta, \phi_2, z_2, \dots)$ . Единственная возможность совпадения этих функций — это когда обе части этого равенства зависят только от  $\theta$ , причем являются одной и той же функцией  $\theta$ . Другими словами, существует некоторая универсальная функция  $g$ , для которой

$$\partial(\log \lambda)/\partial \theta = g(\theta)$$

для всех систем. Исследование конкретных систем показывает, что функция  $g$  нигде не обращается в нуль.

Теперь предположим, что делается замена переменной в эмпирической температуре, т. е. заменим  $\theta$  на  $T(\theta)$ . Тогда будем иметь координаты  $(T, \phi, \dots)$  и равенство

$$\partial(\log \lambda)/\partial T = [\partial(\log \lambda)/\partial \theta](\partial \theta/dT) = [\partial(\log \lambda)/\partial \theta](\partial T/\partial \theta)^{-1}.$$

Если выбрать  $T$  так, что

$$dT/d\theta = Tg(\theta), \quad (***)$$

то

$$\partial(\log \lambda)/\partial T = 1/T = \partial(\log T)/\partial T.$$

Уравнение (\*\*\*) определяет  $T$  с точностью до постоянного множителя. Действительно,  $T(\theta) = ce^{G(\theta)}$ , где  $G$  — неопределенный интеграл от  $g$ , а  $c$  — константа. Такой выбор функции  $T$  называется *абсолютной температурой*. Давайте вернемся к началу этого параграфа, но вместо  $\theta$  будем использовать абсолютную температуру  $T$ . Тогда для каждой системы мы имеем

$$\log \lambda_i = \log T + \log G_i,$$

где  $G_i$  не зависит от  $T$ . Тогда

$$\lambda_1 = TG_1, \quad \lambda_2 = TG_2, \quad \lambda = TG,$$

где функции  $G_1$ ,  $G_2$  и  $G$  не зависят от  $T$ . Кроме того, мы знаем, что  $\lambda_1/\lambda_2$  зависит только от  $\phi_1$  и  $\phi_2$ . Но  $\lambda_1$  — функция первой системы, следовательно, она не зависит от  $\phi_2$ , и аналогично,  $\lambda_2$  не зависит от  $\phi_1$ . Следовательно, для любой системы каждая  $G_i$  зависит только от  $\phi_i$ .

Давайте опустим индексы. Итак, мы показали, что для любой системы выполняется равенство

$$\alpha = TG(\phi) d\phi,$$

где функция  $G$  нигде не обращается в нуль и зависит только от переменных системы. Но тогда можно решить уравнение

$$dS = G d\phi,$$

которое определяет  $S$  с точностью до аддитивной константы. Другими словами, функция  $S$  есть неопределенный интеграл от  $G$ . Тогда мы имеем  $\alpha = T dS$ . Итак, мы доказали следующее.

Существует универсальная шкала *абсолютной температуры*  $T$ , определенная с точностью до постоянно-го множителя. Фиксируя этот множитель и таким образом выбирая функцию  $T$ , мы определяем функцию  $S$  на множестве равновесных состояний любой системы. Эта функция  $S$  определена с точностью до аддитивной константы (одной для каждой системы) и называется *энтропией*. Дифференциальная форма  $\alpha$  для тепла имеет вид

$$\alpha = T dS.$$

Как мы уже отмечали, из существования функции  $S$  и уравнения  $\alpha = T dS$  следует кельвиновская формулировка второго закона термодинамики: для любой замкнутой кривой (циклического процесса) мы должны иметь  $\int dS = 0$ , функция  $T$  имеет постоянный знак (который обычно выбирается положительным). Если

вдоль одной части цикла к системе подводится тепло, то в другой части этого цикла тепло должно быть извлечено. Следовательно, невозможен цикл, в котором тепло полностью преобразуется в работу. Таким образом, «вечный двигатель второго рода» существовать не может.

Рассмотрим простой цикл  $C$  (рис. 22.5). Это кривая, состоящая из четырех частей, две из которых, скажем,  $C_3$  и  $C_4$  находятся на поверхностях, где функция  $S$  постоянна (и поэтому тепло не передается); вдоль одной из них, скажем,  $C_1$  система находится в тепловом контакте с большой системой (называемой тепловым резервуаром) при высокой температуре  $T_1$ . Вдоль четвертого участка  $C_2$  система находится в тепловом контакте с холодным резервуаром при низкой температуре  $T_2$ . Пусть  $Q_1 = \int_{C_1} \alpha$  — полное количество тепла, *поглощенного* системой на участке  $C_1$ , а  $Q_2 = -\int_{C_2} \alpha$  — полное количество тепла, *излученного* ею на участке  $C_2$ . Тогда  $Q_1 = \int_{C_1} T dS = T_1 \int_{C_1} dS$  и  $Q_2 = -T_2 \int_{C_2} dS$ . Но

$$0 = \int_C dS = \int_{C_1} dS + \int_{C_2} dS,$$

поэтому

$$\frac{Q_1}{Q_2} = \frac{T_1}{T_2}.$$

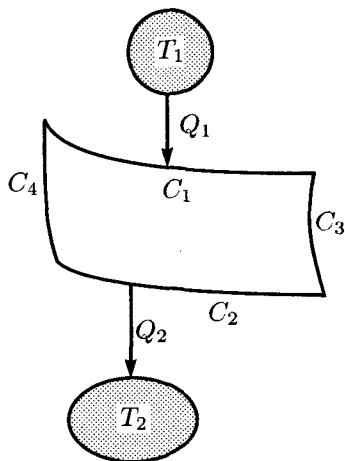


Рис. 22.5

Тогда можно написать

$$0 = \int_C dU = \int_C \omega + \int_C \alpha = \int_C \omega + Q_1 - Q_2.$$

Итак, если  $W$  — работа, совершенная *самой* системой, так что  $W = -\int_C \omega$ , то

$$W = Q_1(1 - T_2/T_1).$$

Это известная формула, которую можно найти в любом учебнике. Она определяет работу, совершенную в цикле Карно, между резервуарами с температурами  $T_2$  и  $T_1$ . На самом деле Карно утверждал, что это наиболее эффективный способ получить работу за счет этих температур, т. е. эта формула дает максимальное значение работы для любой машины, получающей тепло  $Q_1$  при температуре  $T_1$  и выделяющей тепло при температуре  $T_2$ . В основе этого факта лежит изменение функции  $S$  при необратимых адиабатических процессах. Давайте сформулируем общее утверждение, а определение эффективности работы<sup>6</sup> оставим читателю (его можно найти в стандартных учебниках).

Если два равновесных состояния  $\sigma_1$  и  $\sigma_2$  можно соединить обратимой адиабатической кривой, то (как мы уже знаем) они лежат на одной поверхности  $S = \text{const}$ , т. е.  $S(\sigma_1) = S(\sigma_2)$ . Предположим, что сначала система находится в равновесном состоянии  $\sigma_1$ . Рассмотрим адиабатический процесс, не обязательно обратимый, который переводит систему в равновесное состояние  $\sigma_2$ . В этом случае кривая, соединяющая состояния  $\sigma_1$  и  $\sigma_2$ , может и не лежать в подмножестве равновесных состояний, хотя конечные точки находятся именно там (рис. 22.6). Например, у нас имеется жидкость в термосе, и мы размешиваем ее миксером. Когда размешивание заканчивается, жидкость в термосе переходит в равновесное состояние с другим значением энтропии  $S$ . Опыт показывает, что мы не можем вернуться в исходное состояние, если не отвести тепло из такой системы. Тогда (если мы выбрали постоянный множитель такой, что температура  $T$  положительна) для такой необратимой адиабатической кривой мы

<sup>6</sup> Другое (не менее распространенное) название — коэффициент полезного действия (КПД). — *Прим. ред.*

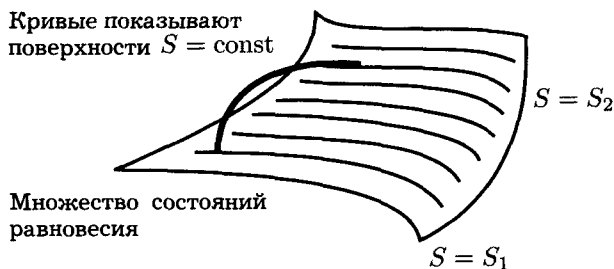


Рис. 22.6

имеем  $S(\sigma_1) < S(\sigma_2)$ . Дадим теперь более строгую формулировку второго закона термодинамики.

Если равновесные состояния  $\sigma_1$  и  $\sigma_2$  могут быть соединены адиабатической кривой (т. е. если в данном процессе замкнутая адиабатическая система может перевести систему из состояния  $\sigma_1$  в  $\sigma_2$ ), то

$$S(\sigma_1) \leq S(\sigma_2),$$

причем  $S(\sigma_1) = S(\sigma_2)$ , если и только если состояния  $\sigma_1$  и  $\sigma_2$  соединяются обратимой адиабатической кривой.

## 22.4. Системы с одной конфигурационной переменной

В этом параграфе мы рассмотрим важный частный случай, когда система имеет только одну конфигурационную переменную и, следовательно, множество равновесных состояний такой системы двумерно. Будем полагать, что такой переменной будет объем. Соответствующая функция, входящая в выражение формы для работы, — это давление, а работа, совершенная над системой, имеет вид

$$\omega = -p dV.$$

Тем не менее соображения этого параграфа применимы и ко многим другим физическим системам. Надо только заменить объем соответствующей конфигурационной переменной, а давление —

Таблица 22.1. Системы с одной конфигурационной переменной

Система	Конфигурационная переменная	Обобщенная сила	Форма для работы $\omega$ в Дж = Н · м
Газ	Объем $V$ в м <sup>3</sup>	Давление $p$ в Н/м <sup>2</sup>	$\omega = -p dV$
Струна	Длина $L$ в м	Натяжение $T$ в Н	$\omega = T dL$
Мембрана	Площадь $A$ в м <sup>2</sup>	Поверхностное натяжение $S$ в Н/м	$\omega = S dA$
Электрическая цепь	Заряд $Q$ в Кл	Напряжение $E$ в В	$\omega = E dZ$
Магнитный материал	Полный магнитный момент $M$ в А/м	Магнитное поле $H$ в А/м <sup>2</sup>	$\omega = \mu_0 H dM$

соответствующей «обобщенной силой». В таблице 22.1 приводятся некоторые из важных физических систем. В этом параграфе мы не будем больше говорить о них.

Прежде чем начать обсуждение, следует сделать несколько замечаний об истории вопроса и соответствующей литературе. Многие из основных математических понятий, которыми мы пользуемся здесь, такие как «множество», «функция», «многообразие», «координаты», «линейная дифференциальная форма», были введены (в их современной форме) или стали общепринятыми уже *после* фундаментальных открытий в термодинамике. Более того, большинство героев этой истории (за исключением Кельвина и Гельмгольца) не имели математического образования, или (как Карно) предпочитали не пользоваться математическим языком. В результате вплоть до наших дней в большинстве книг материал излагается языком первоначальных представлений, довольно туманным с математической точки зрения, в котором, однако, используются соотношения между частными производными. Приведем пример. Понятие линейной диффе-

ренциальной формы, криволинейный интеграл от нее и вариант теоремы Стокса (на плоскости) были введены Ампером в начале 1920-х годов в серии оригинальных статей, где сочетались новый важный математический аппарат и остроумные физические эксперименты. Более общая версия теоремы Стокса для криволинейного интеграла от 1-формы в трехмерном пространстве впервые появилась в письме, которое Кельвин написал Стоксу в 1850 году. (Через несколько лет Стокс опубликовал эту версию в качестве экзаменационного вопроса.) Тем не менее, в своих статьях по термодинамике Кельвин никогда не использовал теорему Стокса, хотя это значительно упростило бы его доказательства.

Итак, мы записали тепло, подведенное к системе, в виде линейной дифференциальной формы  $\alpha$ , определенной на многообразии равновесных состояний. Функции  $T$  и  $V$  имеют независимые дифференциалы. Поэтому, если указанное многообразие двумерное, можно использовать  $T$  и  $V$  как локальные координаты и написать

$$\alpha = \Lambda_V dV + C_V dT, \quad (22.1)$$

где  $\Lambda_V$  и  $C_V$  — функции на многообразии равновесных состояний. Аналогично, независимы функции  $p$  и  $T$ , и поэтому они тоже могут быть локальными координатами. Тогда

$$\alpha = \Lambda_p dp + C_p dT \quad (22.2)$$

где  $\Lambda_p$  и  $C_p$  — две другие функции. (Строго говоря, с исторической точки зрения мы должны использовать эмпирическую температуру  $\theta$  вместо  $T$ , поскольку мы обсуждаем ситуацию, сложившуюся до открытия второго закона термодинамики. Но тогда, вернувшись к современной трактовке, мы должны были бы переписать все формулы через  $T$ . Поэтому мы решили предоставить читателю самому сделать все эти замены.) Конечно, все четыре функции связаны между собой формулой замены переменных, связанной с переходом от координат  $V, T$  к координатам  $p, T$ . Скоро мы вернемся к этому вопросу. Соответствующие физические законы, выраженные двумя предыдущими утверждениями, были сформулированы как «доктрина скрытого и явного тепла». Функция  $\Lambda_V$  была известна как «скрытая теплота по отношению к объему», а функция  $C_V$  — как «теплоемкость при

постоянном объеме». Аналогично, функция  $\Lambda_p$  была известна как «скрытая теплота по отношению к давлению», а функция  $C_p$  — как «теплоемкость при постоянном давлении». Поясним смысл термина «скрытый». На протяжении длительного времени была путаница между понятиями теплоты и температуры. Ученые думали, что температура тела отражает «полное количество теплоты, содержащегося в нем». Это, конечно, бессмыслица, поскольку мы знаем, что форма  $\alpha$  не замкнута. Если мы подводим тепло, то температура тела повышается. Но, если тепло подводится при увеличении объема или давления, то оно становится «скрытым» или «латентным». Конечно, уравнения (22.1) и (22.2) дают четкую математическую формулировку этой загадочной физической «доктрине».

В пространстве двух измерений многие вычисления сильно упрощаются, потому что если  $\Omega_1$  и  $\Omega_2$  — 2-форма и  $\Omega_2 \neq 0$ , то  $\Omega_1 = f\Omega_2$ , где  $f$  — некоторая функция. Другими словами, можно взять «отношение» двух 2-форм и получить функцию, при условии, что знаменатель не равен нулю. Тогда мы пишем  $f = \Omega_1/\Omega_2$ .

В качестве примера рассмотрим функцию

$$\frac{\alpha \wedge dp}{\alpha \wedge dV} = f.$$

Что это такое? Для любой точки  $\mathbf{x}$  пусть  $\xi$  — вектор, касательный к адиабатической кривой, проходящей через точку  $\mathbf{x}$ , так что  $\alpha(\xi) = 0$ , а  $\eta$  — некоторый вектор, независимый от  $\xi$ , так что  $\alpha(\eta) \neq 0$ . Тогда, вычислив формы  $\alpha \wedge dp$  и  $\alpha \wedge dV$  на параллелограмме, построенном на векторах  $\xi$  и  $\eta$ , получаем

$$f(\mathbf{x}) = \frac{dp(\xi)}{dV(\xi)}.$$

Это уравнение можно переписать как

$$f = (dp/dV)_{\text{адиабат.}},$$

где правая часть обозначает отношение линейных дифференциальных форм  $dp$  и  $dV$ , вычисленных на векторах, касательных к

адиабатам. Аналогично, напомним

$$\frac{dT \wedge dp}{dT \wedge dV} = (dp/dV)_{\text{изотерм.}}$$

Если взять внешнее произведение уравнения (22.2) на  $dp$ , то получим

$$\alpha \wedge dp = C_p dT \wedge dp,$$

а если взять внешнее произведение (22.1) на  $dV$ , то получим

$$\alpha \wedge dV = C_V dT \wedge dV.$$

Разделим первое уравнение на второе и получим важный результат

$$(dp/dV)_{\text{адиабат.}} = \gamma(dp/dV)_{\text{изотерм.}}, \quad \text{где } \gamma = C_p/C_V. \quad (22.3)$$

Если ввести плотность  $\rho = m/V$ , то уравнение (22.3) можно записать в виде

$$(dp/d\rho)_{\text{адиабат.}} = \gamma(dp/d\rho)_{\text{изотерм.}}, \quad \gamma = C_p/C_V. \quad (22.4)$$

Приведем физические и исторические причины важности этого результата. Один из известных выводов, сделанных Ньютоном из его законов, определяет скорость звука в газе с формулой

$$c^2 = dp/d\rho, \quad (22.5)$$

где Ньютон предположил, что  $p$  и  $\rho$  связаны соотношением

$$p = N\rho, \quad \text{где } N = \text{const}, \quad (22.6)$$

т. е. что

$$pV = \text{const}. \quad (22.7)$$

Давление, плотность и скорость звука (например, в воздухе) могут быть измерены независимо. И утверждение Ньютона, что  $c^2 = p/\rho$ , противоречило экспериментам на протяжении почти ста лет. Было обнаружено, что скорость звука превышала предсказанное формулой Ньютона в 1.4 раза. С другой стороны, эксперименты Гей-Люссака и других показывали, что при постоянной

температуре формула (22.7) верна, но константа в правой части зависит от температуры. Итак, если мы пользуемся независимыми функциями  $p$  и  $V$  в качестве координат на многообразии равновесных состояний, уравнения (22.7) для различных значений постоянной справа являются уравнениями изотерм, и формулу Ньютона можно написать в виде

$$c^2 = (dp/d\rho)_{\text{изотерм.}}$$

В 1816 году Лаплас высказал идею, что формулу Ньютона следует заменить выражением

$$c^2 = (dp/d\rho)_{\text{адиабат.}} = \gamma(dp/d\rho)_{\text{изотерм.}} \quad (22.8)$$

Его аргументы были чисто физическими: скорость распространения возмущения гораздо больше скорости переноса тепла. Пропитируем самого Лапласа: «Истинная скорость звука равна произведению скорости, полученной из формулы Ньютона, на квадратный корень из отношения удельной теплоемкости воздуха при постоянном давлении и различных температурах к удельной теплоемкости при постоянном объеме». Кроме того, Лаплас подчеркивает, что мы можем использовать наблюдаемую скорость звука для экспериментального определения  $\gamma$ . Конечно, если мы знаем  $\gamma$  и уравнения изотерм, то уравнение (22.3) можно рассматривать как дифференциальное уравнение для адиабатических кривых. Например, для многих газов при низком давлении величина  $\gamma$  почти постоянна. (Водород является исключением из этого правила.) Кроме того, уравнение (22.7) верно при постоянных температурах. Если предположить, что (22.7) является уравнением изотерм и  $\gamma$  постоянна, то из уравнения (22.3) следует, что адиабатические кривые задаются формулой

$$pV^\gamma = \text{const.} \quad (22.9)$$

Заметим, что в этом параграфе мы до сих пор не пользовались ни первым, ни вторым законами термодинамики. Мы пользовались только предположением, что множество равновесных состояний двумерно, и что тепло определяется линейной дифференциальной формой, — т. е. «доктриной скрытого и явного тепла».

Теперь введем оба закона: первый закон —  $dU = \alpha - p dV$  и второй закон —  $\alpha = T dS$ . Соберем их вместе и получим

$$dU = T dS - p dV. \quad (22.10)$$

Возьмем внешнюю производную от этого уравнения и получим

$$dT \wedge dS = dp \wedge dV. \quad (22.11)$$

Как определить функции  $T$ ,  $S$  и  $U$  из эксперимента, скажем в координатах  $p$  и  $V$ ? Чтобы проиллюстрировать метод, предположим, что наблюдаемые изотермы и адиабаты задаются уравнениями (22.7) и (22.9) соответственно. Введем функции

$$t(p, V) = pV \quad (22.12)$$

и

$$a(p, V) = pV^\gamma. \quad (22.13)$$

Тогда

$$\text{изотермы — линии уровня } t \quad (22.14)$$

и

$$\text{адиабаты — линии уровня } a. \quad (22.15)$$

Поскольку  $t$  и  $T$  имеют одни и те же линии уровня и поскольку их дифференциалы нигде не обращаются в нуль, то функция  $T$  может быть выражена как функция  $t$ , т.е. что  $T = T(t)$  и  $T'(t)$  нигде не обращается в нуль. Точно так же заключаем, что  $S = S(a)$ . Наша задача — определить  $T(t)$  и  $S(a)$ . Далее, из уравнения (22.11) имеем

$$dT \wedge dS = T'(t)S'(a) dt \wedge da = dp \wedge dV.$$

Мы должны использовать явные выражения для  $t$  и  $a$  как функций от  $p$  и  $V$ , чтобы выразить  $dt \wedge da$  через  $dp \wedge dV$ . Тогда мы получим, что отношение

$$\frac{dp \wedge dV}{dt \wedge da}$$

является произведением функции от  $t$  и функции от  $a$ , и таким образом найдем  $T'$  и  $S'$ . Итак,

$$\begin{aligned} dt \wedge da &= (p dV + V dp) \wedge (\gamma p V^{\gamma-1} dV + V^\gamma dp) \\ &= (\gamma - 1)p V^\gamma dp \wedge dV = (\gamma - 1)a dp \wedge dV, \end{aligned}$$

откуда

$$\frac{dp \wedge dV}{dt \wedge da} = [(\gamma - 1)a]^{-1}.$$

Это значит, что

$$T'(t)S'(a) = [(\gamma - 1)a]^{-1}.$$

В этом уравнении умножим  $T'$  на любую константу  $k$ , а  $S'$  умножим на  $k^{-1}$ . Это отражает тот факт, что абсолютная температура определена с точностью до постоянного множителя. Поэтому без потери общности можно предположить, что  $T' \equiv 1$  и тогда  $S'(a) = [(\gamma - 1)a]^{-1}$ . Следовательно,

$$T = t + T_0, \quad S = (\gamma - 1)^{-1} \log a + S_0,$$

где  $T_0$  и  $S_0$  — константы. Вспомним, что функция  $S$  определена именно с точностью до аддитивной константы. Но нам еще надо определить  $T_0$  и  $U$ . С этой целью подставим их значения в (22.10) и после простых вычислений получим

$$dU = (\gamma - 1)^{-1} [T_0 d \log a + dt].$$

Вспомним еще один эксперимент — Джоуля–Томпсона, — который показывает, что если газ расширяется адиабатически в вакуум, то его температура при этом практически не изменяется. При таком расширении не совершается работа, и система не получает дополнительное тепло, поэтому внутренняя энергия  $U$  не изменяется. Но при этом расширение является необратимым процессом и поэтому увеличивается энтропия  $S$ . Это может произойти только при условии, что  $T_0 = 0$ . Тогда

$$U = (\gamma - 1)^{-1} t. \quad (22.16)$$

Конечно, количество внутренней энергии пропорционально полному количеству газа. Это значит, что существует константа, связывающая  $t$  с выбором универсальной шкалы температур. Поэтому можно написать

$$t = nRT, \quad (22.17)$$

где  $n$  — количество молей газа, и  $T$  — универсальная температура, измеряемая в кельвинах. (При таком выборе шкалы разность между температурой замерзания и кипения воды равна  $100^\circ$ . Точного соглашения относительно этой шкалы мы не приводим.) Тогда  $R$  — постоянный множитель, преобразующий единицы температуры в единицы энергии, называемый *газовой постоянной*<sup>7</sup>. Из (22.10) следует, что

$$dU \wedge dV = \alpha \wedge dV = C_V dT \wedge dV.$$

Тогда из (22.16) и (22.17) получаем, что

$$C_V = (\gamma - 1)^{-1} nR$$

является константой. Теперь  $C_p$  тоже должна быть постоянной и

$$C_p - C_V = nR.$$

Все перечисленные свойства получены на основании трех предположений, которые определяют *идеальный газ*:

(i) законы Гей-Люссака и Бойля (уравнение (22.7)):

$$pV = nRT;$$

(ii) внутренняя энергия  $U$  зависит только от  $T$  (эксперимент Джоуля);

(iii)  $\gamma$  — постоянная величина.

Давайте теперь соберем вместе наши результаты, но писать их будем для одного моля. Итак, считаем, что  $n = 1$ , и дальше пишем  $v$  вместо  $V$ ,  $c_p$  вместо  $C_p$  и т. д. Тогда получаем

### Законы идеального газа.

- Изотермы —  $pv = RT$ ;
- адиабаты —  $pv^\gamma = \text{const}$ ;

---

<sup>7</sup>Численное значение  $R \simeq 8.31$  Дж/К. — *Прим. ред.*

- $u = (\gamma - 1)^{-1}RT$ ;
- $c_v = (\gamma - 1)^{-1}R$ ;
- $c_p = [\gamma/(\gamma - 1)]R$ , откуда  $c_p - c_v = R$ ;
- $s = (\gamma - 1)^{-1}R \log pv^\gamma + s_0 = R \log v + c_v \log T + s'_0$ .

Для реальных газов можно измерять величины  $p$ ,  $v$ ,  $T$ ,  $c_v$ ,  $c_p$  и  $\gamma$ . Чтобы измерить  $c_v$ , газ помещают в стальной сосуд с тонкими стенками, вокруг которого намотана нагревательная спираль. Тепло, подводимое к газу, определяется по известному току, протекающему по обмотке, и измеряется изменение температуры газа. Чтобы измерить  $c_p$ , газ пропускают через аналогичный нагревательный элемент при постоянном давлении; измеряется температура втекающего и вытекающего газа. Ниже приводятся результаты этих экспериментов при низком давлении, когда газ ведет себя почти как идеальный.

*Для всех газов  $c_v$  и  $c_p$  зависят только от  $T$ , и*

$$c_p - c_v = R.$$

*Для одноатомных газов, таких как He, Ne, Ar, и паров большинства металлов  $c_v$  постоянна в широкой области температур  $T$  и очень близка к  $3R/2$ . Поэтому  $c_p$  очень близка к  $5R/2$ . Величина  $\gamma$  постоянна в широкой области температур  $T$  и близка к  $5/3$ .*

*Для двухатомных газов, таких как  $O_2$ ,  $N_2$ ,  $NO$ ,  $CO$  и воз-<sup>8</sup>дух, при обычных температурах  $c_v$  постоянна и примерно равна  $5R/2$ , Поэтому  $c_p$  примерно равна  $7R/2$ , а  $\gamma$  тоже постоянна при обычных температурах и примерно равна  $7/5$ . Эта константа уменьшается с ростом температуры.*

Именно это значение  $\gamma = 7/5$  участвует в рассуждениях Лапласа относительно скорости звука в воздухе. Истинное значение дробей  $3/2$ ,  $5/2$  и т. д. становится очевидным только в рамках статистической механики.

---

<sup>8</sup> который на 98% состоит из двухатомных компонент — азота  $N_2$  (75%) и кислорода  $O_2$  (23%). — *Прим. ред.*

Для *реальных* газов можно привести такие же рассуждения, начав с уравнения (22.14). Теперь мы уже не предполагаем, что изотермы и адиабаты задаются уравнениями (22.12) и (22.13), или что  $\gamma$  — постоянная величина. Но экспериментально можно определить изотермы и значение  $\gamma$ , следовательно, решить уравнение (22.3) (возможно, только численно) и получить адиабаты. После этого можно найти  $T'(t)$  и  $S'(a)$ , что даст возможность получить  $S$ , а затем  $U$  с помощью уравнения (22.10). Эта информация собиралась на протяжении 150 лет для различных веществ, и теперь она собрана в виде специальных диаграмм. Кроме этого, введены новые функции в дополнение к уже рассмотренным, весьма полезные для определенных задач. Например, *энтальпия*  $H$ , определяемая уравнением

$$H = U + pV.$$

Объясним, почему эта функция важна для нас. Из уравнения (22.10) следует, что тепло, подведенное к системе при постоянном объеме, равно изменению внутренней энергии  $U$ . Но в химических лабораториях гораздо удобнее проводить опыты при постоянном давлении (скажем, при атмосферном давлении) и при этом изменять объем. Дифференцируя определение энтальпии, получаем

$$dH = \alpha + V dp.$$

Поэтому изменение  $H$  определяет количество тепла, подведенного к системе при постоянном давлении. Тогда теплота химической реакции задается изменением энтальпии.

Среди различных эмпирических диаграмм особенно полезными являются диаграммы Моллье, на которых  $p$  и  $H$  — координатные оси, и в этой системе координат изображаются линии уровня  $S$ ,  $V$  и  $T$ . На следующей странице в качестве примера приводится диаграмма Моллье для кислорода.

Существуют другие комбинации  $U$ ,  $p$ ,  $V$  и  $S$ , из которых получаются функции, полезные при решении определенных задач. Их обозначения меняются, поэтому здесь мы не будем их описывать. Однако, есть функция, очень полезная для практических целей и имеющая важное теоретическое значение. В последующих параграфах мы будем неоднократно обращаться к ней. То, что будет

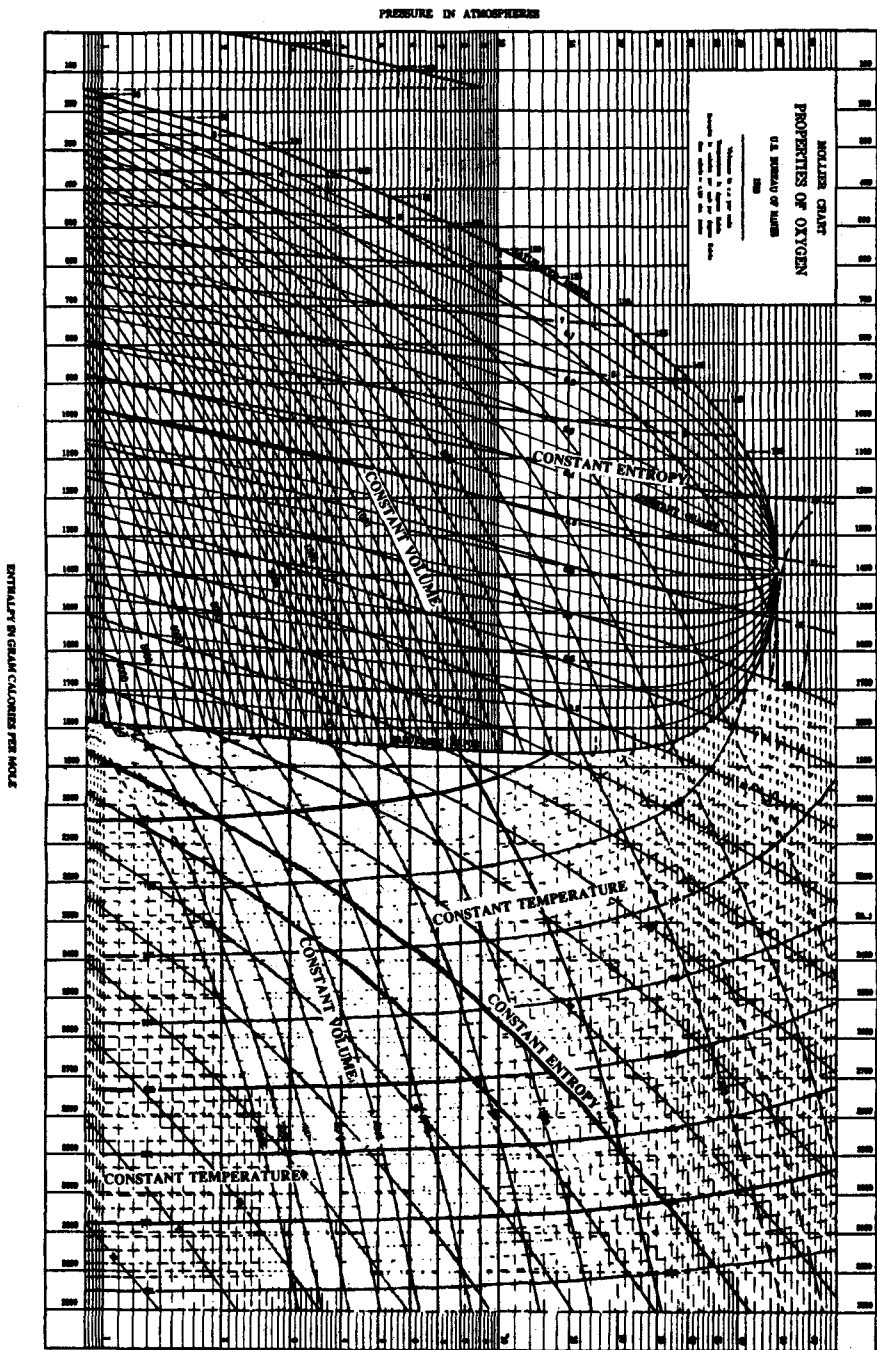


Диаграмма Моллье для кислорода. Если заданы значения любых двух из пяти функций  $p$ ,  $V$ ,  $H$ ,  $T$ ,  $S$ , то можно найти точку, в которой достигаются эти значения, как пересечение соответствующих линий уровня. Тогда можно определить значения остальных трех функций диаграммы и функции  $U$  из равенства  $U = H - pV$ .

сделано здесь, на первый взгляд, может показаться весьма странным, но в рамках статистической механики это будет вполне убедительным. Рассмотрим *функцию Планка*  $Y$ , определенную равенством

$$Y = S - (RT)^{-1}U - (RT)^{-1}pV.$$

Чтобы записать эту формулу в более удобной форме, введем

$$\beta = (RT)^{-1}, \quad \nu = \beta p. \quad (22.18)$$

Заметим (например, из уравнений (22.16) и (22.17)), что  $R$  — множитель, связывающий энергию и температуру, поэтому размерностью  $\beta$  является обратная энергия, скажем, в единицах  $(\text{Дж})^{-1}$ , а размерностью  $\nu$  — обратный объем, так что  $\beta dU$  и  $\nu dV$  — безразмерные величины. Предположим, что с самого начала мы включили  $R$  в определение температуры  $T$ , чтобы измерять температуру в единицах энергии. Тогда газовая постоянная исчезает. Поскольку размерностью  $\alpha = T dS$  является энергия, и мы измеряем температуру  $T$  тоже в единицах энергии, то в такой температурной шкале энтропия  $S$  становится безразмерной. Тогда  $S$  и функция Планка

$$Y = S - \beta U - \nu V \quad (22.19)$$

безразмерные. Комбинируя первый и второй законы термодинамики, уравнение (22.10) можно записать в виде

$$dS = \beta dU + \nu dV. \quad (22.20)$$

Тогда из формулы (22.19) следует, что

$$dY = -U d\beta - V d\nu. \quad (22.21)$$

Уравнения типа (22.20) и (22.21) будем интерпретировать следующим образом. Рассмотрим четырехмерное пространство  $\mathbb{R}^4$  с координатами  $(\beta, \nu, U, V)$ . В этом пространстве введем внешнюю 2-форму

$$\Omega = d\beta \wedge dU + d\nu \wedge dV. \quad (22.22)$$

Во всем пространстве  $\mathbb{R}^4$  мы имеем

$$\Omega = d(\beta dU + \nu dV) \quad \text{и} \quad \Omega = d(-U d\beta - V d\nu).$$

Двумерное подмногообразие  $\mathcal{L}$  пространства  $\mathbb{R}^4$  называется *лагранжевым*, если 2-форма  $\Omega$ , ограниченная на  $\mathcal{L}$ , тождественно обращается в нуль. Для любого двумерного подмногообразия пространства  $\mathbb{R}^4$  величины  $\beta$  и  $\nu$  можно рассматривать как функции. Если их дифференциалы линейно независимы при ограничении на подмногообразие, то функции  $\beta$  и  $\nu$  можно использовать в качестве локальных координат. Аналогично, функции  $U$  и  $V$  можно рассматривать в качестве локальных координат, если их дифференциалы линейно независимы. См. рис. 22.7.

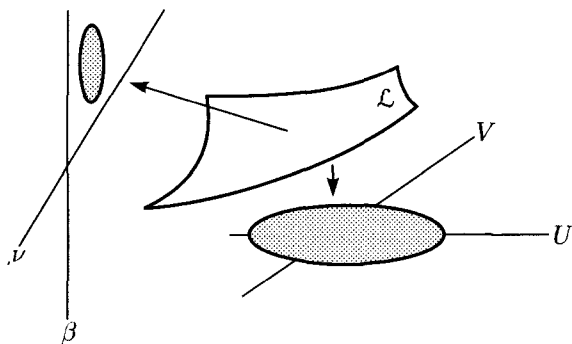


Рис. 22.7. Подмногообразие  $\mathcal{L}$  лагранжево. Проекция на плоскость  $\beta, \nu$  взаимно-однозначна, поэтому можно рассматривать  $\beta, \nu$  как локальные координаты на  $\mathcal{L}$ . По той же причине можно в качестве локальных координат на  $\mathcal{L}$  взять  $U, V$

Если подмногообразие  $\mathcal{L}$  является лагранжевым, то при ограничении на это подмногообразие имеем

$$d(\beta dU + \nu dV) = 0,$$

и, следовательно, (локально) существует функция  $S$ , определенная на  $\mathcal{L}$  так, что выполняется уравнение (22.20). Аналогично, поскольку при ограничении на подмногообразие  $\mathcal{L}$  мы имеем

$$d(-U d\beta - V d\nu) = 0,$$

то можно сделать вывод, что (локально) существует такая функция  $Y$ , что выполняется уравнение (22.21). Таким образом, можно дать общую формулировку для первого и второго законов:

многообразие равновесных состояний является лагранжевым подмногообразием в пространстве с координатами  $(\beta, \nu, U, V)$ .

Продолжим наше математическое рассуждение. Предположим, что мы ввели функции  $U$  и  $V$  в качестве локальных координат (см. рис. 22.7), и предположим, что мы знаем функцию  $S$ . Это означает, что нам задана функция  $S$ , явно зависящая от  $U$  и  $V$ . Если сопоставить равенство

$$dS = (\partial S / \partial U) dU + (\partial S / \partial V) dV$$

с равенством (22.20), то получим выражения для  $\beta$  и  $\nu$  через  $U$  и  $V$ , а именно

$$\beta = \partial S / \partial U \quad \text{и} \quad \nu = \partial S / \partial V. \quad (22.23)$$

Это, конечно, определяет подмногообразие  $\mathcal{L}$ . Другими словами, если у нас есть диаграмма Моллье в координатах  $U$  и  $V$  с линиями уровней и значениями только функции  $S$ , то можно восстановить  $\mathcal{L}$  и, следовательно, по крайней мере в принципе, все необходимые термодинамические функции. Вся информация относительно  $\mathcal{L}$  закодирована в  $S$ , если  $S$  является функцией  $U$  и  $V$ . Мы называем  $S$  *производящей функцией* лагранжева подмногообразия  $\mathcal{L}$ .

Аналогично, если в качестве локальных координат взять  $\beta$  и  $\nu$  и задать  $Y$  как функцию  $\beta$  и  $\nu$ , то из равенства (22.21) можно восстановить функции  $U$  и  $V$  в зависимости от  $\beta$  и  $\nu$  с помощью равенств

$$U = -\partial Y / \partial \beta \quad \text{и} \quad V = -\partial Y / \partial \nu.$$

Таким образом,  $Y$  — производящая функция  $\mathcal{L}$ . (Заметим, что это дает обратное отображение, выражая  $\beta$  и  $\nu$  как функции от  $U$  и  $V$ .)

Мы можем проделать то же самое, выбрав в качестве локальных координат другие переменные. Например, можно написать, что на всем пространстве  $\mathbb{R}^4$

$$\Omega = d(-U d\beta + \nu dV)$$

и соответственно на  $\mathcal{L}$

$$dZ = -U d\beta + \nu dV,$$

где функция Массье  $Z$  определяется выражением

$$Z = S - \beta U. \quad (22.24)$$

Таким образом функция Массье — производящая функция для  $\mathcal{L}$  в координатах  $\beta$  и  $V$ . В уравнении (22.24) закодированы первый и второй законы термодинамики вместе с полным описанием равновесных состояний нашей системы!

Когда мы перейдем к изучению равновесной статистической механики, то увидим, что функции  $S$  и  $Z$  являются базисными. Будет дана интерпретация функции  $S$  на языке теории вероятностей, из чего станет ясно, что ее значение распространяется далеко за пределы термодинамики. Будет показано, что функция  $Z$  обеспечивает связь между микроскопической теорией и наблюдаемыми макроскопическими явлениями, т. е. из общей конструкции станет понятно, как модель энергии на атомном или молекулярном уровне приводит к определенному выражению для  $Z$  как функции от  $\beta$  и  $V$ . (Многие технические аспекты станут чисто математической задачей — как вычислять или оценивать это выражение.) И, конечно, мы разовьем теорию так, чтобы она работала для систем с несколькими конфигурационными переменными.

Прежде чем переходить к статистической механике, следует еще обсудить ряд идей относительно равновесия. Это будет сделано в следующем параграфе. Но в заключение этого параграфа обратимся к замечанию, сделанному в самом начале. Мы говорили, что многие физически измеряемые величины задаются отношением 2-форм на поверхности равновесных состояний, и показали это на примере двух типов «теплоемкостей». Приведем еще два примера.

Коэффициент теплового расширения при постоянном давлении	$\frac{dV \wedge dp}{V dT \wedge dp}$
Коэффициент сжимаемости при постоянной температуре	$\frac{dV \wedge dT}{V dT \wedge dp}$

Другие примеры величин и способов обращения с ними мы дадим в задачах к этой главе.

## 22.5. Условия равновесия

Рассмотрим системы с произвольным числом конфигурационных переменных. До сих пор мы определяли функцию  $S$  только на многообразии равновесных состояний. Важно отметить, что некоторые функции, например, температура или давление, определяются только на многообразии равновесных состояний. Нет смысла говорить о температуре газа (как целого), если он не находится в равновесии. В свое время было сделано великое открытие — энтропия является функцией, определенной на множестве всех состояний, т. е. существует функция Ent такая, что

$$\text{Ent}(\rho) = S(\rho),$$

если  $\rho$  — равновесное состояние. Здесь мы не будем давать определение Ent, поскольку это потребует описания произвольных состояний системы. Такое определение будет дано, когда мы будем заниматься статистической механикой. Тогда мы напишем явную формулу для Ent, фактически полученную Больцманом. Однако, многие идеи, излагаемые в этом параграфе, принадлежат Клаузиусу и другим, кто работал до открытия статистической интерпретации.

Мы воспользуемся свойством аддитивности энтропии для двух систем, находящихся в контакте. (Сейчас это свойство будем рассматривать как аксиому, но потом оно будет доказано в рамках статистической механики.) Итак, под двумя системами, находящимися в контакте, мы подразумеваем большую систему, построенную из двух меньших, так что состояние большой системы  $\rho$  является парой состояний  $\rho = (\rho_1, \rho_2)$  отдельных компонент системы, и внутренняя энергия

$$U(\rho) = U_1(\rho_1) + U_2(\rho_2).$$

Другими словами, «энергия взаимодействия» пренебрежимо мала по сравнению с суммой индивидуальных энергий. Мы не предполагаем, что контакт только тепловой (рис. 22.8). Например, возможно движение поршня между резервуарами, так что возможен «обмен объемами». Или имеется проницаемая перегородка между системами, так что возможен обмен молекулами газа

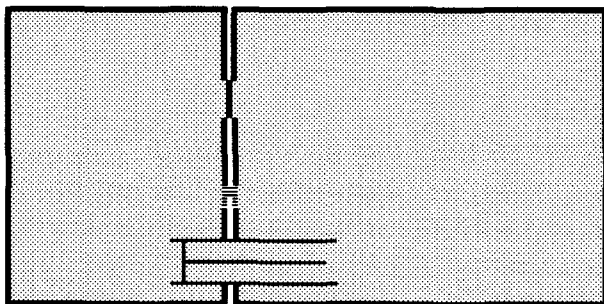


Рис. 22.8. Две системы, находящиеся в контакте трех видов. Они могут обмениваться теплом (сверху), веществом (в середине) и объемом (внизу)

и т. д. Энтропия составной системы является суммой энтропий составляющих систем, т. е.

$$\text{Ent}(\rho) = \text{Ent}(\rho_1) + \text{Ent}(\rho_2), \quad \text{если } \rho = (\rho_1, \rho_2).$$

Далее мы будем использовать это важное свойство. Если мы функция энтропии определена на множестве всех состояний, то разумно дать более сильную формулировку второго закона термодинамики.

Если система адиабатически замкнута, то энтропия любого состояния не может убывать со временем. Более того, энтропия всегда возрастает, если только система уже не находится в равновесном состоянии.

Поскольку энергия сохраняется, это приводит к важной характеристике равновесных состояний, наиболее важной в статистической механике.

**Принцип максимальной энтропии:** среди всех состояний с данной внутренней энергией равновесное состояние обладает максимальной энтропией.

**Принцип минимальной энергии Гиббса:** если  $\rho$  — равновесное состояние, а  $\sigma$  — другое состояние с той же энтропией, что и  $\rho$ , т. е.  $\text{Ent}(\sigma) = \text{Ent}(\rho)$ , то  $U(\rho) \leq U(\sigma)$ .

Чтобы получить правило Гиббса из принципа максимальной энтропии, воспользуемся физическим опытом. Мы всегда можем подводить тепло к системе и поэтому непрерывно увеличивать энергию и энтропию системы, причем это можно делать, не изменяя конфигурационных переменных. Предположим, что  $\rho$  — равновесное состояние, и предположим, что (вопреки опыту) есть другое состояние  $\sigma$  в окрестности  $\rho$ , причем  $\text{Ent}(\sigma) = \text{Ent}(\rho)$  и  $U(\sigma) < U(\rho)$ . Подводя тепло к системе, мы можем непрерывно увеличивать энергию и энтропию, пока не достигнем нового состояния с той же внутренней энергией, что и  $\rho$ , но большей энтропией, что противоречит принципу максимальной энтропии. Это доказывает принцип Гиббса.

Прежде чем пользоваться принципом Гиббса, сделаем ряд предварительных замечаний. Предположим, что  $f_1, \dots, f_k$  — функции от конфигурационных переменных системы. Фиксируя значения этих функций, можно получить *новую* систему. Например, предположим, что две системы, изображенные на рис. 22.8, сначала были разделены, и каждая имела свой поршень. До объединения мы имели систему, которая была просто прямым произведением двух подсистем. Поэтому у нас было четыре конфигурационные переменные: объемы подсистем —  $V_1$  и  $V_2$ , полная масса газа в каждой подсистеме —  $M_1$  и  $M_2$ . (Для упрощения обсуждения предположим, что в подсистемах был одинаковый газ.) Приводя подсистемы в контакт, как показано на рис. 22.8, мы имеем фиксированные полный объем и полную массу. Другими словами, мы наложили ограничения на формы  $f = \text{const}$  и  $g = \text{const}$ , где

$$f = V_1 + V_2 \quad \text{и} \quad g = M_1 + M_2,$$

сократив число независимых переменных с четырех до двух.

Возьмем некоторую систему и перейдем к новой системе, наложив некоторые ограничения на конфигурационные переменные. Такая система будет иметь собственные равновесные состояния, которые характеризуются максимумом энтропии и минимумом внутренней энергии, но, конечно, по отношению к более ограниченному классу состояний, т. е. тех, что удовлетворяют ограничениям. Предположим, что нам задано подмножество  $M$  состояний исходной системы (без ограничений), которое содержит все ее

равновесные состояния, и предположим, что  $\mathcal{M}$  — многообразие. Кроме того, предположим, что функция Ent и ограничивающие функции  $f_1, \dots, f_k$  дифференцируемы на  $\mathcal{M}$ . Пусть  $\rho$  — равновесное состояние системы с ограничениями, в которой функции  $f_s$  постоянны. Согласно принципу Гиббса, функция  $U$  должна иметь локальный минимум в состоянии  $\rho$  при условии, что

$$\text{Ent} = \text{const}, \quad f_1 = \text{const}, \quad \dots, \quad f_k = \text{const}.$$

Воспользуемся методом множителей Лагранжа. Существуют функции  $\lambda_0, \lambda_1, \dots, \lambda_k$ , для которых

$$dU = \lambda_0 d\text{Ent} + \lambda_1 df_1 + \dots + \lambda_k df_k. \quad (22.25)$$

Например, в случае системы, изображенной на рис. 22.8, многообразие  $\mathcal{M}$  должно быть парой равновесных состояний отдельных систем, когда они не контактируют друг с другом. Каждая подсистема в отдельности имеет конфигурационные переменные  $V$  и  $M$ . Тогда для двух подсистем имеем

$$dU_1 = T_1 dS_1 - p_1 dV_1 + \mu_1 dM_1 \quad \text{и} \quad dU_2 = T_2 dS_2 - p_2 dV_2 + \mu_2 dM_2.$$

Коэффициент  $\mu$  при  $dM$  называется «химическим потенциалом». Многообразие  $\mathcal{M}$  шестимерное, поскольку множество равновесных состояний каждой подсистемы трехмерное. Мы предполагаем, что

$$U = U_1 + U_2.$$

Свойство аддитивности энтропии означает, что

$$\text{Ent} = S_1 + S_2.$$

Тогда

$$dU = T_1 dS_1 + T_2 dS_2 - p_1 dV_1 - p_2 dV_2 + \mu_1 dM_1 + \mu_2 dM_2.$$

Из условия (22.25) следует, что

$$dU = \lambda_0(dS_1 + dS_2) + \lambda_1(dV_1 + dV_2) + \lambda_2(dM_1 + dM_2),$$

поскольку  $V_1 + V_2$  и  $M_1 + M_2$  — ограничивающие функции.

Дифференциалы, стоящие в правой части этих двух равенств, независимы на многообразии  $M$ , поэтому мы имеем

$$T_1 = T_2, \quad p_1 = p_2, \quad \mu_1 = \mu_2.$$

Таким образом, давление, температура и химические потенциалы должны быть равны. Очевидно, что это утверждение обобщается на случай нескольких систем, находящихся в контакте, когда каждая подсистема может содержать смесь разного количества разных веществ. Поэтому, если выражение для внутренней энергии на множестве равновесных состояний каждой отдельной системы имеет вид

$$dU_i = T_i dS_i - p_i dV_i + \mu_i^a M_i^a + \mu_i^b dM_i^b + \dots,$$

то получаем условия равновесия объединенной системы

$$\begin{aligned} T_1 &= T_2 = \dots, \\ p_1 &= p_2 = \dots, \\ \mu_1^a &= \mu_2^a = \dots, \\ \mu_1^b &= \mu_2^b = \dots, \\ &\text{и т. д.} \end{aligned} \tag{22.26}$$

Эти равенства являются общими условиями равновесия. Они дают нам возможность почувствовать, что такое «интенсивные переменные»:  $T$ ,  $p$ ,  $\mu$  и т. д.

В качестве примера рассмотрим газ, состоящий только из одного вещества. Каждый участок такого газа можно рассматривать как подсистему. Если каждая такая подсистема, взятая в отдельности, находится в состоянии равновесия, то мы имеем для нее определенные значения  $T$ ,  $p$  и  $\mu$ . Другими словами, мы считаем, что  $T$ ,  $p$  и  $\mu$  являются функциями в трехмерном пространстве, которые каждой точке приписывают значения температуры, давления и химического потенциала в небольшой окрестности этой точки. Условие того, что газ как целое находится в состоянии равновесия, состоит в том, что эти функции должны быть постоянными по всему объему газа. Тогда газ как целое, скажем,

с постоянной полной массой имеет только одну конфигурационную степень свободы  $V$ . Мы можем использовать  $p$  и  $T$  в качестве координат двумерного пространства равновесных состояний газа как целого. Это означает, что величина  $\mu$  выражается как функция координат  $p$  и  $T$ . Тогда для газа, состоящего из одного вещества, находящегося в состоянии равновесия, химический потенциал  $\mu$  — это некоторая определенная функция от  $p$  и  $T$ , т. е.  $\mu = \mu_{\text{газ}}(p, T)$ . Предположим, что это вещество может находиться и в жидкой фазе. Для жидкости в состоянии равновесия мы получим другую функцию  $\mu = \mu_{\text{жидк.}}(p, T)$ . А теперь рассмотрим комбинированную систему, состоящую из газа и жидкости одного вещества. По отдельности газ и жидкость находятся в состоянии равновесия. Предположим, что эта система, как целое, тоже находится в равновесии. Тогда из условий (22.26) следует, что

$$p_{\text{жидк.}} = p_{\text{газ}}, \quad T_{\text{жидк.}} = T_{\text{газ}}, \quad \mu_{\text{жидк.}} = \mu_{\text{газ}}.$$

Если  $p$  и  $T$  обозначают общие значения для первых двух условий, то третье условие принимает вид

$$\mu_{\text{жидк.}}(p, T) = \mu_{\text{газ}}(p, T).$$

Поскольку, как правило,  $\mu_{\text{жидк.}}$  и  $\mu_{\text{газ}}$  — независимые функции от  $p$  и  $T$ , то последнее равенство определяет кривую на плоскости  $p, T$ . Другими словами, жидкая и газообразная фаза одного и того же вещества могут находиться в равновесии только при определенном соотношении между температурой и давлением. Если имеются три фазы одного вещества, скажем, газ, жидкость и твердая фаза, то условия равновесия в такой системе имеют вид

$$\mu_{\text{тверд.}}(p, T) = \mu_{\text{жидк.}}(p, T) = \mu_{\text{газ}}(p, T).$$

Тогда решением этих уравнений будет так называемая тройная точка с определенным значением температуры и давления. И наконец, четыре фазы одного вещества не могут находиться в равновесии. Именно так Гиббс получил свое знаменитое *правило фаз*. Оно (как и его обобщения) содержится в условиях (22.26), которые являются прямым следствием равенства (22.25) — принципа минимальной энергии Гиббса.

## 22.6. Системы и состояния в статистической механике

В этом параграфе мы займемся микроскопическими теориями, которые могут послужить моделью для термодинамики. Начнем с классической статистической механики, в рамках которой «состояние» определяется как «мера вероятности» определенного рода, и энтропией состояния будет «количеством беспорядка этого состояния». Математические основы теории вероятностей были заложены в первой трети XX века, главным образом, Борелем, Лебегом и Колмогоровым. Это произошло гораздо позднее появления фундаментальных идей статистической механики, выдвинутых Больцманом и Гиббсом. Таким образом, и здесь возник коммуникационный барьер. Языком теории вероятностей является теория меры. Мы будем пользоваться этим языком, не вдаваясь глубоко в саму теорию. Для любого серьезного результата теории вероятностей, связанного с бесконечным числом возможностей, существенно использование всего аппарата теории меры. Поскольку мы будем рассматривать наиболее простые факты, то пользоваться будем только самыми простыми понятиями, смысл которых станет ясным на конкретных примерах.

*Измеримое пространство* — это множество  $M$  вместе с семейством  $\mathcal{A}$  подмножеств этого множества. Мера  $\mu$  задает правило, согласно которому каждому множеству  $A$  семейства  $\mathcal{A}$  ставится в соответствие неотрицательное число (или  $+\infty$ )  $\mu(A)$ . Для семейства  $\mathcal{A}$  и меры  $\mu$  существует ряд аксиом, на которые мы не будем здесь отвлекаться. (Для нас существенно, что  $\mathcal{A}$  — замкнутое семейство относительно счетного объединения, пересечения и дополнения<sup>9</sup>, и что  $\mu(\bigcup A_i) = \sum \mu(A_i)$  для любого счетного объединения непересекающихся множеств  $A_i$ .) Мы не рассматриваем набор *всех* подмножеств  $M$ , чтобы избежать противоречий, возникающих при рассмотрении сильно патологических множеств.

В большинстве случаев мы здесь рассматриваем очень простые множества  $M$ . Например, пусть  $M$  — конечное множество, и пусть семейство  $\mathcal{A}$  будет семейством всех подмножеств  $M$ . Тогда множества, состоящие из одной точки  $M$ , тоже будут входить

<sup>9</sup>Такое семейство множеств принято называть  $\sigma$ -алгеброй. — Прим. ред.

в наше семейство, и каждому такому множеству  $\{m\}$  приписывается неотрицательное число  $\mu(\{m\})$ . Зная эти числа, можно восстановить меру  $\mu$ , поскольку

$$\mu(A) = \sum_{m \in A} \mu(\{m\}).$$

Если  $f$  — произвольная (вещественная или векторная) функция на  $M$ , то можно определить «интеграл по мере  $\mu$ » в виде

$$\int f \mu = \sum_{m \in M} f(m) \mu(\{m\}).$$

(Это такая же договоренность, которая была сделана в теории электрических цепей.)

В качестве  $M$  можно было бы взять  $\mathbb{Z}$  — множество всех целых чисел, или  $\mathbb{N}$  — множество всех неотрицательных целых чисел. А в качестве  $\mathcal{A}$  можно было бы взять семейство, состоящее из всех подмножеств. Определение меры означает задание неотрицательного числа для каждого целого числа (в случае  $\mathbb{Z}$ ) или для каждого неотрицательного целого числа (в случае  $\mathbb{N}$ ). Тогда мы получили бы такое же выражение для  $\mu(A)$ , как и раньше, но сумма справа в этом случае будет бесконечной, которая может сходиться, а может и нет. Если она сходится, то  $\mu(A)$  равно пределу суммы; если же она расходится, то  $\mu(A) = +\infty$ .<sup>10</sup> То же самое происходит и для функции  $f$ , для которой мы определяем интеграл  $\int f \mu$  в виде суммы. Сумма теперь становится бесконечным рядом. Мы говорим, что функция  $f$  интегрируема, если ряд сходится абсолютно, и тогда  $\int f \mu$  определяется как сумма этого ряда. В противном случае  $\int f \mu$  не имеет смысла.

В качестве  $M$  можно взять вещественную прямую  $\mathbb{R}$ . Тогда семейство  $\mathcal{A}$  должно содержать все отрезки  $[a, b] = \{x \mid a \leq x \leq b\}$ . Обычная мера (Лебега) каждому отрезку приписывает его длину

$$\mu([a, b]) = b - a.$$

<sup>10</sup> Легко видеть, что если  $\sum_n \mu(\{n\})$  — сходящийся ряд (для  $n \in \mathbb{Z}$  или  $n \in \mathbb{N}$ ), то мера всякого множества конечна и не превосходит суммы ряда. — Прим. ред.

Если  $f$  — произвольная кусочно-непрерывная функция, достаточно быстро стремящаяся к нулю на бесконечности, то мы знаем, как вычислять интеграл<sup>11</sup>

$$\int f \mu = \int_{-\infty}^{\infty} f(x) dx.$$

Аппарат теории меры предназначен для построения достаточно богатого класса семейств  $\mathcal{A}$ , на которых определена мера  $\mu$ , и достаточно широкого класса функций, интегралы от которых определены. Мы будем рассматривать достаточно простые множества функций, и поэтому нам не придется вдаваться в детали теории меры. Как мы уже говорили, необходимо задействовать весь аппарат теории меры, если заниматься глубокими проблемами.

*Вероятностная мера* приписывает значение 1 самому множеству  $M$ , т.е.  $\mu(M) = 1$ . Это соответствует принятому в теории условию, что «вероятности» принимают значения от 0 до 1.

Фундаментальное понятие «системы» в статистической механике — это пространство с мерой. Множество  $M$ , семейство  $\mathcal{A}$  и мера  $\mu$  заданы нам физикой или геометрией задачи. В общем случае мера  $\mu$  не будет вероятностной мерой. Она играет роль «*априорного*» состояния знания системы. Покажем это на нескольких примерах.

*А. Пространство конечной выборки с равной априорной вероятностью.* Это простейший пример.  $M$  — конечное множество,  $\mathcal{A}$  — семейство всех подмножеств  $M$ . В этом случае  $\mu(\{m\}) = 1$

<sup>11</sup>в смысле Римана. Интеграл Лебега численно с ним совпадает, но определен для значительно более широкого класса функций (в том числе таких всюду разрывных, как функция Дирихле, равная 1 в рациональных точках и 0 в иррациональных), и процедура его вычисления существенно отличается от римановых сумм, поскольку разбивается не область определения функции, а область значений. Для вещественной функции  $f$  интеграл  $\int f \mu$  по мере Лебега  $\mu = dx$  (или другой мере  $\mu$ ) равен пределу при  $h \rightarrow 0$  сумм вида

$$\sum_{i \in \mathbb{Z}} y_i \mu(\{x \mid y_i \leq f(x) < y_{i+1}\}),$$

где  $\{y_i\}_{i \in \mathbb{Z}}$  — такое разбиение области значений  $f$ , что  $0 < y_{i+1} - y_i < h$ . — *Прим. ред.*

для каждого  $m \in M$ . Так описывается ситуация, когда есть конечное число возможностей, и у нас нет оснований предпочесть одну возможность другой.

В.  $M$  — вещественная прямая  $\mathbb{R}$  со стандартным семейством измеримых множеств и мерой Лебега  $\mu$ . Тогда  $\mu([a, b]) = b - a$  для любого отрезка  $[a, b]$ . Мера  $\mu$  показывает «априорную вероятность» отрезка, пропорциональную его длине. Конечно,  $\mu$  не является вероятностной мерой, потому что  $\mu(\mathbb{R}) = \infty$ . Заметим, что если  $\rho$  — произвольная интегрируемая функция, для которой

$$\int_{-\infty}^{\infty} \rho(x) dx = 1,$$

то  $\rho$  определяет вероятностную меру  $\rho dx$  на множестве  $\mathbb{R}$ , где

$$P([a, b]) = \int_{-\infty}^{\infty} \rho(x) dx.$$

Таким образом, мера  $\mu(= dx)$  определяет выделенный класс вероятностных мер на множестве  $\mathbb{R}$ . (На языке теории меры это класс тех вероятностных мер, которые абсолютно непрерывны по  $dx$ . Не все вероятностные меры на  $\mathbb{R}$  обладают этим свойством. Например, рассмотрим вероятностную меру  $P$ , которая приписывает значения  $P(\{0\}) = 1/2$ ,  $P(\{1\}) = 1/2$  и  $P(A) = 0$ , если  $\{0, 1\} \cap A = \emptyset$ . Тогда  $P$  обозначает вероятность в случае, когда достигаются точки либо 0, либо 1, причем их шансы одинаковы. При этом очевидно, что  $P$  не представляется<sup>12</sup> мерой вида  $\rho dx$ .)

Для любой меры  $\mu$  на измеримом пространстве и для любой интегрируемой неотрицательной функции  $\rho$  мы получаем новую меру, которую обозначаем  $\rho\mu$ , положив

$$\rho\mu(A) = \int_A \rho\mu.$$

Здесь следует ввести общее

**Определение:** пусть  $(M, \mathcal{A}, \mu)$  — пространство с мерой. *Состояние* (статистическое)  $(M, \mathcal{A}, \mu)$  — это вероятностная мера в пространстве  $M$  вида  $\rho\mu$ , где  $\rho \geq 0$  — интегрируемая функция. Другими словами, состояние соответствует такой интегрируемой функции  $\rho$ , для которой  $\int \rho\mu = 1$ .

<sup>12</sup>если не рассматривать обобщенных функций. — Прим. ред.

С. Пусть  $M = \mathbb{R}$  и  $\mu = H(\cdot) dx$ , где  $H(x) = 0$  при  $x \leq 0$  и  $H(x) = 1$  при  $x > 0$ .<sup>13</sup> Это соответствует ситуации, когда мы уверены, что данное вещественное число положительное, причем все положительные числа равновероятны. Конечно, можно с тем же успехом взять  $M = \mathbb{R}^+ = \{x \mid x > 0\}$ , поскольку здесь сконцентрирована мера  $\mu$ .

Д. Число заполнения для классических частиц. В этом случае  $M = \mathbb{N} = \{0, 1, 2, 3, \dots\}$  и  $\mu(\{k\}) = 1/k!$ . Мера  $\mu$  приписывает вес  $1/k!$  вероятности  $k$  частицам находиться в ящике. Объясним, почему это подходящая мера. Предположим, что у нас есть некоторое количество ящиков  $n$  и несколько частиц  $N$ . Число способов размещения частиц по ящикам, так чтобы  $k_1$  частиц попали в первый ящик,  $k_2$  — во второй и т. д., равно

$$\frac{N!}{k_1! k_2! \dots k_n!}.$$

(Конечно, здесь частицы предполагаются «различимыми»: обмен парой частиц между разными ящиками изменяет способ размещения частиц. Существует  $N!$  различных перестановок частиц, но мы должны разделить это число на  $k_1! \dots k_n!$ , потому что перестановка частиц внутри каждого ящика не дает новый способ их размещения.) Если нам не известны числа  $N$  и  $n$ , то лучшее, что мы можем сказать, — это что число способов размещения частиц так, чтобы в первом ящике было  $k_1$  частиц, пропорционально  $1/k_1!$ .<sup>14</sup> Это и есть наше *априорное* знание относительных вероятностей.

Е. Число заполнения для частиц Ферми–Дирака. В этом случае в каждый ящик можно поместить только одну частицу. Тогда  $M = \{0, 1\}$  и  $\mu(\{0\}) = \mu(\{1\}) = 1$ . Это, конечно, частный случай примера А. Мы также могли бы взять  $M = \mathbb{N} =$  неотрицательным целым числам, причем  $\mu(\{0\}) = \mu(\{1\}) = 1$  и

<sup>13</sup>Указанная функция  $H$  называется функцией Хевисайда. Она является первообразной для обобщенной дельта-функции Дирака (сосредоточенной в нуле). — Прим. ред.

<sup>14</sup>Данное утверждение справедливо лишь при  $N \lesssim n$ , в противном случае самым вероятным числом частиц в ящике будет ближайшее целое к  $N/n$ . — Прим. ред.

$\mu(\{k\}) = 0$  для всех  $k > 1$ . В этой схеме заполнения отражен «принцип запрета». В одном ящике не может быть больше одной частицы. Наличие частицы в ящике или ее отсутствие *априори* рассматриваются как равновероятные события<sup>15</sup>. Конечно, в реальной интерпретации ящик может быть достаточно абстрактным. Например, это может быть квантовое состояние одной частицы для частиц, подчиняющихся принципу запрета Паули.

Г. *Число заполнения для частиц Бозе-Эйнштейна*. Здесь  $M = N = \{0, 1, 2, 3, \dots\}$  и  $\mu(\{k\}) = 1$ . Любое число частиц в ящике может быть обнаружено с одинаковой вероятностью. Если, например, частица — это возмущение среды (пятно на экране) и две частицы представляют возмущение двойной интенсивности, то при распределении  $N$  «единиц возмущения» по  $n$  «ящикам» надо учесть, что «частицы» «не различимы» в том смысле, что все их распределения равновероятны. Таким образом *априори* мера, приписываемая наличию  $k_1$  частиц в первом ящике, равна 1.

В квантовой механике все частицы делятся на два класса: частицы Ферми-Дирака и частицы Бозе-Эйнштейна. Частицы Ферми-Дирака (электроны, протоны, нейтроны и т. д.) называются фермионами, а частицы Бозе-Эйнштейна (фотоны, пионы и т. д.) называются бозонами.

Г. *Биномиальная мера*. Предположим, что у нас есть большое число  $N$  частиц, каждая из которых может быть в двух положениях — «вверх» или «вниз». Причем состояние их совокупности описывается только полным числом частиц в положении «вверх» минус полное число частиц в положении «вниз». Тогда  $M = \{-N, -N + 2, \dots, N - 2, N\}$ . Для простоты предположим, что  $N$  — четное число. Тогда число частиц в положении «вверх» минус число частиц в положении «вниз» равняется  $2m$ , когда  $(N/2) + m$  частиц находятся в положении «вверх», а  $(N/2) - m$  частиц — в положении вниз. Число способов сделать это равно

$$\binom{N}{N/2 + m} = \frac{N!}{(N/2 + m)!(N/2 - m)!},$$

<sup>15</sup>при условии, что мы не знаем отношение числа частиц к числу ящиков. То же относится и к следующему примеру. — *Прим. ред.*

Если нам интересно знать относительную частоту появления  $2m$ , надо разделить на самый большой биномиальный коэффициент  $\binom{N}{N/2}$  и тогда

$$\mu(\{2m\}) = \binom{N}{N/2+m} \binom{N}{N/2} = \frac{[(N/2)!]^2}{(N/2+m)!(N/2-m)!}.$$

Использование формулы Стирлинга  $N! = \sqrt{2\pi N} \cdot N^N e^{-N}$  показывает, что для больших чисел  $N$  последнее выражение очень хорошо аппроксимируется функцией  $e^{-2m^2/N}$ . На этом основан следующий пример.

Н. *Дискретная гауссова мера*. Пусть  $M = \mathbb{Z}$  — множество всех целых чисел и

$$\mu(\{m\}) = e^{-2m^2/N}.$$

Заметим, что когда  $m = \sqrt{N/2}$ , значение  $\mu(\{m\})$  равняется к  $e^{-1}\mu(\{0\})$ . Предположим, что мы хотим знать величину  $m/N$ , т. е. что каждое состояние «вверх» или «вниз» дает вклад  $\pm 1/N$  в измеряемую величину. Тогда функция  $m/N$  имеет очень острый пик в окрестности 0 для больших чисел  $N$ . Например, если  $N \sim 10^{24}$ , то  $\mu(\{m\})$  уменьшается в  $e$  раз, когда  $m/N$  сдвигается от 0 на величину  $10^{-12}$ .

Чтобы «увидеть» распределение, мы должны изменить масштаб множителем  $N^{1/2}$ , и не  $N$ . Если взять  $x_k = 2k/N^{1/2}$ , то

$$\mu(\{a \leq x_k \leq b\}) = \sum_a^b e^{-x_k^2/2}.$$

Сумма справа растёт приблизительно как  $N^{1/2}$ . Действительно, если разделить это выражение на  $(2\pi N)^{-1/2}$ , то правая часть стремится к

$$\frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

Тем самым мы подошли к следующему примеру.

I.  $M = \mathbb{R}$  и  $\mu$  — (нормированная) непрерывная гауссова мера

$$\mu([a, b]) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

Ж. В качестве последнего примера рассмотрим *классическую механику*. В этом случае пространство  $M$  — фазовое пространство механической системы. Фазовое пространство имеет выделенные координаты (канонические координаты)  $q, p$  и выделенную меру (мера Лиувилля)

$$\mu = dp dq.$$

Например, точечная классическая частица движется в пространстве  $\mathbb{R}^3$ . Координаты  $q_1, q_2, q_3$  описывают ее положение, а координаты  $p_1, p_2, p_3$  задают ее импульс. Тогда  $M = \mathbb{R}^6$  и

$$dp dq = dp_1 dp_2 dp_3 dq_1 dq_2 dq_3,$$

т. е. пространство  $\mathbb{R}^6$  с обычной мерой.

## 22.7. Произведения и образы

В этом параграфе мы обсудим различные способы построения систем. Сначала обсудим понятие произведения. Предположим, что у нас есть две системы  $(M_1, \mathcal{A}_1, \mu_1)$  и  $(M_2, \mathcal{A}_2, \mu_2)$ . Можно образовывать их произведение  $(M_1 \times M_2, \mathcal{A}_1 \times \mathcal{A}_2, \mu_1 \times \mu_2)$ . Например, пусть  $(M_1, \mathcal{A}_1, \mu_1)$  и  $(M_2, \mathcal{A}_2, \mu_2)$  — системы, описанные в примере Е предыдущего параграфа, т. е. они представляют возможное число частиц в ящике (см. рис. 22.9). Тогда  $(M_1 \times M_2, \mathcal{A}_1 \times \mathcal{A}_2, \mu_1 \times \mu_2)$  представляет систему, состоящую из двух ящиков, соединенных вместе (см. рис. 22.10).

Мера  $\mu_1 \times \mu_2$  характеризуется выражением

$$(\mu_1 \times \mu_2)(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$$

для произведений множеств  $A = A_1 \times A_2$ , где  $A_i \in \mathcal{A}_i$ ,  $i = 1, 2$ .<sup>16</sup> Конечно, состояние  $\rho$  системы-произведения не обязано иметь

<sup>16</sup> Семейство измеримых подмножеств  $\mathcal{A}_1 \times \mathcal{A}_2$  произведения измеримых пространств  $(M_1, \mathcal{A}_1)$  и  $(M_2, \mathcal{A}_2)$  состоит из счетных объединений «прямоугольных» множеств вида  $A_1 \times A_2$ , где  $A_i \in \mathcal{A}_i$ , т. е. является пополнением семейства  $\{A_1 \times A_2\}$  до  $\sigma$ -алгебры. Произведение мер  $\mu_1 \times \mu_2$  определяется предыдущим равенством и свойством аддитивности. — *Прим. ред.*

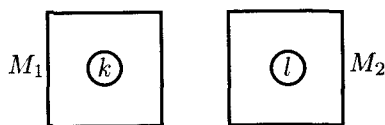


Рис. 22.9.  $\mu(\{k\}) = \frac{1}{k!}$ ,  $\mu(\{l\}) = \frac{1}{l!}$

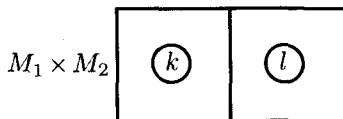


Рис. 22.10.  $(\mu_1 \times \mu_2)(k, l) = \frac{1}{k! l!}$

вид  $\rho_1 \times \rho_2$ . Если же это состояние типа  $\rho = \rho_1 \times \rho_2$ , то соответствующая вероятностная мера  $(\rho_1 \times \rho_2)(\mu_1 \times \mu_2) = \rho_1 \mu_1 \times \rho_2 \mu_2$  на множестве  $M_1 \times M_2$  тоже является произведением мер. Это означает, что событиям в пространствах  $M_1$  и  $M_2$  приписываются *независимые* вероятности.

Можно построить произведение любого<sup>17</sup> количества систем  $(M_i, \mathcal{A}_i, \mu_i)$ , где  $i$  пробегает произвольное множество индексов  $I$ .

Другой способ построения системы использует отображение. Пусть  $(M, \mathcal{A}, \mu)$  — одна система, и  $(N, \mathcal{B})$  — измеримое пространство, т. е.  $N$  — множество с выделенным классом подмножеств. Пусть  $f : M \rightarrow N$  — отображение. Предположим, что для каждого  $B \in \mathcal{B}$  подмножество  $f^{-1}(B)$  множества  $M$  принадлежит семейству  $\mathcal{A}$ .<sup>18</sup> Тогда можно определить правило  $f_*\mu$ , согласно которому каждому  $B \in \mathcal{B}$  приписывается значение

$$f_*\mu(B) = \mu(f^{-1}B).$$

<sup>17</sup>В принципе число сомножителей может быть бесконечным, но к определению измеримых множеств произведения в этом случае необходим более осторожный подход. — *Прим. ред.*

<sup>18</sup>Тогда отображение  $f$  называется измеримым. В частности, измеримы относительно обычной меры все кусочно-непрерывные отображения евклидовых пространств. — *Прим. ред.*

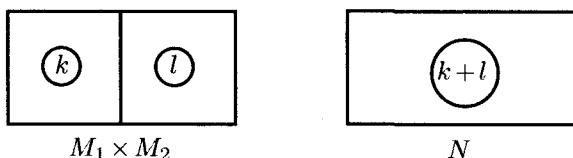


Рис. 22.11

Если функция  $f_*\mu$  на множествах из  $\mathcal{B}$  является мерой, то она называется *переносом*<sup>19</sup> меры  $\mu$  при отображении  $f$ . Тогда мы имеем новую систему  $(N, \mathcal{B}, f_*\mu)$ , называемую образом системы  $(M, \mathcal{A}, \mu)$ . Например, начнем с произведения  $M = M_1 \times M_2$  двух систем из того же примера Е. Пусть  $N$  — это большой ящик, из которого удалили перегородку (см. рис. 22.11). Тогда отображение  $f$  задается формулой

$$f((k, l)) = k + l.$$

Таким образом, при отображении  $f$  игнорируется то, что в ящике слева находится  $k$  частиц, а в ящике справа —  $l$  частиц. Оно оставляет только информацию, что в большом ящике находятся  $k + l$  частиц. Тогда для любого положительного целого числа  $s$  можно написать

$$(f_*\mu)(\{s\}) = \sum_{k+l=s} \frac{1}{k!l!} = \frac{1}{s!} \sum_{k+l=s} \frac{s!}{k!l!}.$$

Откуда

$$(f_*\mu)(\{s\}) = \frac{2^s}{s!}.$$

Оба способа построения — с помощью произведения или с помощью отображения — в данном случае становятся интуитивно более понятными, если немного изменить пример Е. Давайте определим систему, соответствующую ящику с объемом  $V$ , положив

$$M_V = N, \quad \mu_V(\{k\}) = \frac{V^k}{k!}.$$

<sup>19</sup>В оригинале *pushforward*, т. е. движение вперед. Действительно, в отличие от переноса дифференциальной формы (*pullback*, см. примечание на с. 291 в томе I), при этой операции область определения меры перемещается в том же направлении, что и точки при отображении. — *Прим. ред.*

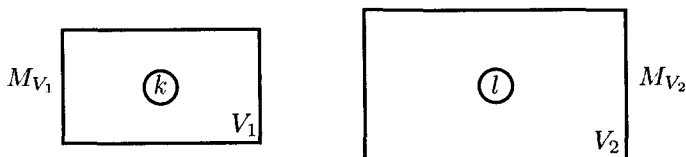


Рис. 22.12.  $\mu_{V_1}(\{k\}) = \frac{V_1^k}{k!}$ ,  $\mu_{V_2}(\{l\}) = \frac{V_2^l}{l!}$

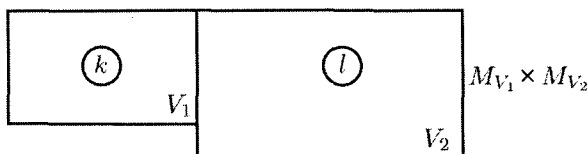


Рис. 22.13.  $\mu_{V_1} \times \mu_{V_2}(\{k, l\}) = \frac{V_1^k V_2^l}{k! l!}$

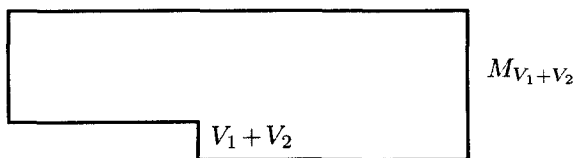


Рис. 22.14

Тогда если рассмотреть две системы, соответствующие ящикам с объемами  $V_1$  и  $V_2$  (см. рис. 22.12), то  $M_{V_1} \times M_{V_2}$  соответствует рассмотрению двух ящиков вместе (см. рис. 22.13). Уберем теперь перегородку и рассмотрим ящик с объемом  $V_1 + V_2$  (см. рис. 22.14). Опять положим  $N = M_{V_1 + V_2}$  и определим отображение  $f : M_{V_1} \times M_{V_2} \rightarrow N$  формулой  $f((k, l)) = k + l$ . Тогда

$$\begin{aligned}
 (f_*\mu)(\{s\}) &= \sum_{k+l=s} \frac{V_1^k V_2^l}{k! l!} \\
 &= \frac{1}{s!} \sum_{k+l=s} \frac{s!}{k! l!} V_1^k V_2^l \\
 &= \frac{(V_1 + V_2)^s}{s!}.
 \end{aligned}$$

Вернемся к рассмотрению общего случая. Пусть  $(M, \mathcal{A}, \mu)$  — система, и отображение  $F : M \rightarrow N$  дает нам образ системы  $(N, \mathcal{B}, f_*\mu)$ . Если  $\sigma$  — произвольное состояние из  $(N, \mathcal{B}, f_*\mu)$ , то функция  $f^*\sigma = \sigma \circ f$  определяет состояние из  $(M, \mathcal{A}, \mu)$ . Действительно,  $f^*\sigma \geq 0$ , тогда из определения интеграла следует, что

$$\int f^*\sigma \mu = \int \sigma(f_*\mu) = 1.$$

Таким образом, состояния, определенные на множестве  $N$ , задают состояния, определенные на множестве  $M$ . Состояние  $\rho$  на множестве  $M$  также определяет состояние  $f_*\rho$  на множестве  $N$ , но это определение лежит значительно глубже. Давайте сначала рассмотрим пример. Рассмотрим пространство  $M_1 = \mathbb{R}^2$  со стандартными семейством измеримых множеств  $\mathcal{A}$  и мерой  $\mu (= dx dy)$ . Пусть отображение  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^+$  задано формулой  $f(x, y) = (x^2 + y^2)^{1/2}$ . Тогда  $f_*\mu = 2\pi r dr$ . Пусть  $\rho(x, y)$  — произвольная непрерывная функция на плоскости, удовлетворяющая условию  $\int \rho dx dy = 1$ . Тогда  $\rho$  является состоянием. Определим состояние  $f_*\rho$  условием

$$f_*\rho(r) = \frac{1}{2\pi} \int_0^{2\pi} \rho(r \cos \theta, r \sin \theta) d\theta.$$

Таким образом,  $f_*\rho(r)$  — состояние, «усредненное» по окружности с радиусом  $r$ , т. е. среднее от  $\rho$  по  $f^{-1}(r)$ . Тогда

$$\begin{aligned} \int_{\mathbb{R}^+} (f_*\rho) \mu &= \int_0^\infty (f_*\rho)(r) 2\pi r dr \\ &= \int_0^\infty \int_0^{2\pi} \rho(r \cos \theta, r \sin \theta) dr d\theta \\ &= \int_{-\infty}^\infty \int_{-\infty}^\infty \rho(x, y) dx dy = 1. \end{aligned}$$

Отсюда следует, что  $f_*\rho$  — состояние, определенное на множестве  $N$ . Заметим, что  $f_*f^*\sigma = \sigma$ , но, вообще говоря,  $f^*f_*\rho$  не обязано быть равным  $\rho$ . В общем случае, когда  $\rho$  — состояние, определенное на множестве  $M$ ,  $\rho\mu$  является вероятностной мерой на  $M$ .

Поэтому  $f_*(\rho\mu)$  — вероятностная мера на множестве  $N$ . Кроме того, если  $B \in \mathcal{B}$  — любое множество, удовлетворяющее условию

$$(f_*\mu)(B) = 0, \quad \text{то} \quad \mu(f^{-1}(B)) = 0,$$

и, следовательно,

$$(f_*(\rho\mu))(B) = (\rho\mu)(f^{-1}(B)) = \int_B \rho\mu = 0.$$

Итак, равенство  $(f_*\mu)(B) = 0$  означает, что  $f_*(\rho\mu)(B) = 0$ . Теорема из теории меры, называемая теоремой Радона–Никодима, утверждает, что если  $\nu_1$  и  $\nu_2$  — две меры такие, что для всех  $B$  из  $\nu_2(B) = 0$  следует, что  $\nu_1(B) = 0$ , то существует такая функция  $\sigma$ , для которой  $\nu_1(B) = \int_B \sigma \nu_2$ . Если взять  $\nu_1 = f_*(\rho\mu)$  и  $\nu_2 = f_*\mu$ , то мы приходим к заключению, что существует неотрицательная интегрируемая функция  $\sigma$ , определенная на множестве  $N$ , для которой справедливо равенство

$$(f_*(\rho\mu))(B) = \int_B \rho(f_*\mu)$$

для любых  $B$ . Из того, что  $f_*(\rho\mu)$  — вероятностная мера, следует

$$\int_B \sigma(f_*\mu) = 1.$$

Тогда  $\sigma$  — состояние, определенное на множестве  $(N, \mathcal{B}, f_*\mu)$ , и мы получаем

$$f_*\rho = \sigma.$$

## 22.8. Наблюдаемые, ожидаемые значения и внутренняя энергия

Частным случаем отображений является обычная функция с численными значениями  $f : M \rightarrow \mathbb{R}$ . Такое отображение называется *наблюдаемым* для данной системы. Грубо говоря, наблюдаемая величина какой-то системы — это просто численное свойство

системы, которое можно измерить. Обсуждая чисто теоретические вопросы, будем обозначать наблюдаемое буквой  $J$ . Конкретные наблюдаемые, когда они будут появляться, обозначаются буквами, отражающими их физический смысл. Конечно, как часть нашего определения наблюдаемого, мы предполагаем, что  $J^{-1}([a, b])$  принадлежит семейству  $\mathcal{A}$  для каждого отрезка  $[a, b]$  на вещественной прямой, т. е. предполагается измеримость отображения  $J$ .

Пусть  $J$  — наблюдаемое и  $\rho$  — состояние. Для каждого отрезка  $[a, b]$  можно рассмотреть подмножество  $J^{-1}([a, b])$  множества  $M$ , и тогда интеграл

$$\int_{J^{-1}([a, b])} \rho \mu$$

есть как раз размер множества  $J^{-1}([a, b])$  по отношению к вероятностной мере  $\rho \mu$ . Другими словами, это есть вероятность того, что точка множества  $M$  лежит в  $J^{-1}([a, b])$ , когда мы приписываем вероятности в соответствии с вероятностной мерой  $\rho \mu$ . То же самое можно сказать по-другому: это есть вероятность того, что наблюдаемое  $J$  принимает значения на отрезке  $[a, b]$ , когда мы приписываем вероятности на множестве  $M$  согласно вероятностной мере  $\rho \mu$ . Запишем это утверждение в следующем виде.

$$\text{Prob}(J \in [a, b], \text{ когда система находится в состоянии } \rho) = \int_{J^{-1}([a, b])} \rho \mu,$$

или короче

$$\text{Prob}(J \in [a, b]; \rho) = \int_{J^{-1}([a, b])} \rho \mu. \quad (22.27)$$

Таким образом, наблюдаемое и состояние вместе приписывают значение вероятности на вещественной прямой. Вероятность, приписанная любому отрезку, есть вероятность измерения наблюдаемой  $J$  на этом отрезке, когда система находится в состоянии  $\rho$ ; она задается предыдущей формулой.

В теории вероятностей существует язык, который мы будем здесь использовать. Интеграл

$$\int_M J \rho \mu$$

называется *математическим ожиданием* или *ожидаемым значением* функции  $J$  по отношению к вероятностной мере  $\rho\mu$ . Это есть среднее значение  $J$ , когда вес подмножеств множества  $M$  берется в соответствии с вероятностной мерой  $\rho\mu$ . Таким образом, *ожидаемое значение*  $J$  в состоянии  $\rho$  определяется интегралом

$$E(J; \rho) = \int_M J \rho \mu. \quad (22.28)$$

Приведем пример. Пусть  $M$  — фазовое пространство классической частицы, так что  $M = \mathbb{R}^6$  с координатами  $q_1, q_2, q_3, p_1, p_2, p_3$  и стандартной мерой. Пусть  $J$  — это функция энергии, равная сумме кинетической и потенциальной энергий, так что  $J = H = K + \mathcal{V}$ , т. е.

$$H(q_1, q_2, q_3, p_1, p_2, p_3) = \frac{1}{2}(p_1^2 + p_2^2 + p_3^2) + \mathcal{V}(q_1, q_2, q_3),$$

где  $K$  — кинетическая энергия

$$K = \frac{1}{2}(p_1^2 + p_2^2 + p_3^2)$$

и  $\mathcal{V}$  — потенциальная энергия. В этом случае состояние — это неотрицательная функция  $\rho$ , определенная на пространстве  $\mathbb{R}^6$ , интеграл от которой по всему пространству равен 1. По определению

$$E(H; \rho) = \int_{\mathbb{R}^6} H \rho dq_1 dq_2 dq_3 dp_1 dp_2 dp_3. \quad (22.29)$$

Предположим, что мы хотим рассмотреть энергию «свободной частицы, помещенной в ящик  $B$  с объемом  $V$ ». Это можно сделать, предположив соответствующую форму потенциала  $\mathcal{V}$ . Пусть  $\mathcal{V}$  равняется 0, когда точка  $(q_1, q_2, q_3)$  лежит внутри ящика  $B$ , и  $\mathcal{V}$  становится чрезвычайно большим за пределами ящика. Таким образом, предполагается, что частица должна преодолеть очень высокий барьер, чтобы вылететь из ящика. Выбранный потенциал внесем в выражение для  $H$  в формуле (22.29). Предположим теперь, что состояние  $\rho$  таково, что оно приписывает пренебрежимо малую вероятность очень большим значениям

энергии, т. е. предположим, что  $\text{Prob}(H > E; \rho)$  практически равна нулю для очень больших  $E$ . Чтобы быть совсем точными, давайте предположим, что приписываемое значение вероятности настолько мало, что при вычислении интеграла (22.29) можно пренебречь вкладом от точек  $(q_1, q_2, q_3)$ , лежащих за пределами ящика. Внутри ящика  $V = 0$ , и интеграл принимает вид

$$\int_{B \times \mathbb{R}^3} \frac{1}{2} (p_1^2 + p_2^2 + p_3^2) \rho(q_1, q_2, q_3, p_1, p_2, p_3) dq_1 dq_2 dq_3 dp_1 dp_2 dp_3. \quad (22.30)$$

Чтобы двигаться дальше, необходимо иметь больше информации о состоянии  $\rho$ . Например, предположим, что когда частица находится в ящике, то она с одинаковой вероятностью находится в любой его точке, т. е.  $\rho$  не зависит от координат положения  $(q_1, q_2, q_3)$ , а зависит только от  $(p_1, p_2, p_3)$ . Тогда интеграл (22.30) принимает вид

$$V \int_{\mathbb{R}^3} \frac{1}{2} (p_1^2 + p_2^2 + p_3^2) \rho(p_1, p_2, p_3) dp_1 dp_2 dp_3. \quad (22.31)$$

Далее предположим, что в состоянии  $\rho$  распределение скоростей частицы задается так называемым распределением Максвелла с параметром  $\beta$ , т. е. что

$$\rho(p_1, p_2, p_3) = F_\beta^{-1} \exp[-\beta(p_1^2 + p_2^2 + p_3^2)/2],$$

где  $F_\beta$  — нормировочная константа, зависящая от  $\beta$ . Мы должны теперь выбрать  $F_\beta$  так, чтобы полный интеграл  $\int_M \rho \mu = 1$ , т. е. мы должны взять  $F_\beta = \int_M \rho \mu$ , что составляет

$$F_\beta = V(2\pi\beta)^{-3/2}.$$

Тогда формула (22.31) принимает вид

$$(\beta/2\pi)^{3/2} \int \frac{1}{2} (p_1^2 + p_2^2 + p_3^2) \exp[-\beta(p_1^2 + p_2^2 + p_3^2)/2] dp_1 dp_2 dp_3.$$

Это гауссов интеграл, значением которого является  $(3/2)\beta^{-1}$ . Тогда, с учетом всех сделанных предположений, мы имеем

$$E(H; \rho) = (3/2)\beta^{-1}.$$

Как это следует понимать? С точки зрения математики, формулы (22.29)–(22.31) представляют чисто интегральные выражения. Это стандартные процедуры *математической* теории вероятностей. С точки зрения *кинетической теории газов*, рассматривается огромное количество молекул газа, заключенных в ящик и не взаимодействующих друг с другом<sup>20</sup>. Тогда  $\rho(q_1, q_2, q_3, p_1, p_2, p_3)$  определяет относительную частоту появления молекул газа в окрестности точки  $(q_1, q_2, q_3, p_1, p_2, p_3)$  фазового пространства. В такой интерпретации процедура усреднения, входящая в определение  $E(H; \rho)$ , есть усреднение по молекулам газа, и тогда можно сказать, что  $E(H; \rho)$  — средняя энергия, приходящаяся на одну молекулу газа.

Понадобился гений Гиббса, чтобы осознать, что *не следует связывать себя такой интерпретацией*. Мы вполне можем исследовать систему, состоящую из одной молекулы в ящике. Теория вероятностей применима и в этом случае, вся математика остается без изменения. Выбрав некоторое состояние  $\rho$ , мы рассматриваем вероятность энергии принимать определенные значения в любом интервале или вероятность нахождения частицы в определенной области фазового пространства. Конечно, если мы захотим исследовать более глубокие проблемы, то нам придется изучить общий смысл термина «вероятность» на более высоком уровне. К сожалению, несмотря на то, что математическое обоснование теории вероятностей было сделано более пятидесяти лет назад, нельзя сказать то же самое о ее философской интерпретации. Когда пишется эта книга (1987 год), все еще существуют серьезные разногласия относительно «истинного» смысла вероятности. Сторонники одной крайности — «частотники» — ограничивают область применения вероятности только ситуациями, когда «одно и то же» явление повторяется много раз, так что вероятность имеет смысл частоты. С другой стороны, есть «последователи Байеса», которые рассматривают вероятность как способ мышления — теория вероятностей выражает, как должен думать каждый здравомыслящий «рациональный» гражданин. Экстремисты из каждой школы рассматривают членов другой школы

<sup>20</sup> Это одно из постулируемых свойств идеального газа, которое довольно хорошо выполняется в обычных условиях. — *Прим. ред.*

как людей, у которых в голове «каша» и т. д. Эти философские споры приводят к ощутимым последствиям в реальной статистической практике. Различные взгляды приводят к разным статистическим процедурам, особенно если рассматривается небольшое число объектов. К счастью, эти споры не имеют отношения к статистической механике, поскольку число рассматриваемых в ней объектов чрезвычайно велико. Важна только математическая сторона этой задачи, т. е. вычисления с данными определениями. Однако, Гиббс работал около ста лет назад, когда математическая теория еще не была разработана. Поэтому ему приходилось заниматься и философией тоже. Кажется, что он скорее разделял точку зрения «частотников». Поэтому, описывая вероятностную меру типа  $\rho$ , он говорил об «ансамбле», подтверждая интерпретацию «частотников». К сожалению, эта терминология и слово «ансамбль» как синоним вероятностной меры прочно засели в стандартных учебниках.

Но, независимо от интерпретации, мы уже готовы перекинуть первый мостик между понятиями статистической механики и термодинамики. Если в качестве наблюдаемого рассматривать энергию  $H$ , то

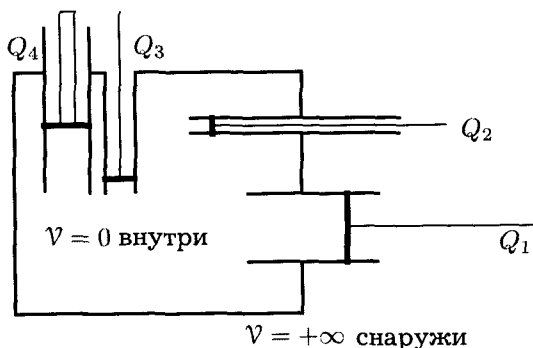
*внутренняя энергия состояния  $\rho$  определяется как  $E(H; \rho)$ .*

Другими словами, внутренняя энергия есть ожидаемое значение энергии в данном состоянии. Конечно, функция энергии  $H$  — фиксированная функция на множестве  $M$ . Существуют различные состояния  $\rho$ , дающие значения энергии  $E(H; \rho)$ . Запишем это определение в форме

$$U(\rho) = E(H; \rho), \quad (22.32)$$

так что внутренняя энергия — это функция на пространстве состояний.

Мы уже сказали, что  $H$  — фиксированная функция на множестве  $M$ . Но в общем случае мы позволим функции  $H$  зависеть от некоторых дополнительных параметров. Например, еще раз рассмотрим газ, находящийся в ящике. Но теперь пусть в ящике имеются различные поршни (рис. 22.15), так что его форма изменяется соответственно значениям  $Q_1, Q_2, \dots, Q_d$ , определяющим

Рис. 22.15. Зависимость  $\mathcal{V}$  от параметров

положение поршней. Тогда потенциальная энергия  $\mathcal{V}$  (предполагается, что она равна 0 внутри ящика и « $+\infty$ » за его пределами) зависит от формы ящика и, следовательно, от параметров  $Q_1, Q_2, \dots, Q_d$ .

Таким образом, энергия  $U$  также зависит от этих параметров:

$$U = U(\rho, Q_1, Q_2, \dots, Q_d).$$

Итак, у нас имеются понятия состояния и внутренней энергии. Что такое равновесные состояния? Мы можем определить их с помощью принципа максимальной энтропии. В общем случае мы сделаем это в параграфе 22.11. А сейчас вернемся к примеру газа, заключенного в ящик.

Для каждого положительного значения  $\beta$  рассмотрим функцию  $e^{-\beta H}$  в пространстве  $\mathbb{R}^6$ . (Здесь  $\beta$  — параметр, имеющий размерность обратной энергии, так что  $\beta H$  — безразмерная числовая функция. Очень скоро  $\beta$  станет играть роль, о которой говорилось в параграфе 22.4.) Поскольку  $\mathcal{V}$  и, следовательно,  $H$  обращается в  $+\infty$ , когда точка  $(q_1, q_2, q_3)$  находится за пределами ящика, функция  $e^{-\beta H}$  там равна нулю, и она везде неотрицательна. Поэтому в качестве состояния можно использовать  $F^{-1}e^{-\beta H}$ , где постоянная  $F$  выбирается так, чтобы интеграл от этой функции по всему пространству был равен 1, т. е.

$$F = \int_{\mathbb{R}^6} e^{-\beta H} dq_1 dq_2 dq_3 dp_1 dp_2 dp_3. \quad (22.33)$$

Мы уже вычисляли этот интеграл. Интегрирование по пространственным переменным  $(q_1, q_2, q_3)$  проводится только внутри ящика, где функция  $H$  от них не зависит. Поэтому это интегрирование дает множитель  $V$  — объем ящика. А интеграл по импульсам гауссов, что приводит к результату

$$F = V(2\pi\beta)^{-3/2}. \quad (22.34)$$

Здесь  $F$  зависит от  $\beta$ , а также от параметров  $Q_1, Q_2, \dots, Q_d$  только через объем  $V$ . В этом частном случае определим равновесные состояния как

$$\rho_{\beta, Q} = F^{-1} e^{-\beta H}, \quad (22.35)$$

где  $F = F(\beta, Q_1, Q_2, \dots, Q_d)$  задается формулой (22.33).

Давайте определим

$$Z = \log F. \quad (22.36)$$

Можно вычислить  $\partial Z / \partial \beta$ , дифференцируя (22.33) под знаком интеграла, и получить

$$\begin{aligned} \partial Z / \partial \beta &= F^{-1} \partial F / \partial \beta = F^{-1} \int_{\mathbb{R}^6} -H e^{-\beta H} dq_1 dq_2 dq_3 dp_1 dp_2 dp_3 \\ &= - \int H \rho_{\beta, Q} dq_1 dq_2 dq_3 dp_1 dp_2 dp_3 = -E(H; \rho_{\beta, Q}) \end{aligned} \quad (22.37)$$

или

$$\partial Z / \partial \beta = -U. \quad (22.38)$$

Вернемся к понятиям работы и обратимой кривой. Предположим, что мы медленно двигаем один из поршней, скажем,  $Q_1$ . Соответственно изменяется энергия  $H$ , и возникает функция  $\partial H / \partial Q_1$  на пространстве  $M$ , которая имеет смысл силы. Для любого заданного состояния  $\rho$  мы должны вычислить среднее значение и получить обобщенную силу

$$E(\partial H / \partial Q_1; \rho).$$

Обратимость процесса означает, что мы ограничиваем состояния  $\rho$  требованием, что они принадлежат множеству равновесных состояний. Следовательно, форма для работы имеет вид

$$\begin{aligned} \omega &= E(\partial H / \partial Q_1; \rho) dQ_1 + E(\partial H / \partial Q_2; \rho) dQ_2 + \dots \\ &\quad + E(\partial H / \partial Q_k; \rho) dQ_k. \end{aligned}$$

Дифференцируя  $Z$  по  $Q_i$ , получим

$$-\beta\omega = (\partial Z/\partial Q_1) dQ_1 + (\partial Z/\partial Q_2) dQ_2 + \cdots + (\partial Z/\partial Q_k) dQ_k \quad (22.39)$$

или в более символической форме

$$d_Q Z = -\beta\omega, \quad (22.39a)$$

где  $d_Q$  обозначает правую часть формулы (22.39), т. е. мы вычисляем оператор  $d$  в направлениях  $Q$  (при фиксированном значении  $\beta$ ). Заметим, что при выводе уравнений (22.37)–(22.39) мы нигде не использовали явный вид  $H$ . Если же теперь предположить, что рассматривается «свободный газ в ящике», где функция  $F$  и, следовательно,  $Z$  зависят только от  $Q_i$  через объем  $V$ , то можно написать  $\partial Z/\partial Q_i = (\partial Z/\partial V)(\partial V/\partial Q_i)$ , и формула (22.39) превращается в

$$-\beta\omega = (\partial Z/\partial V) dV.$$

Конечно, увеличение  $V$  вызывает уменьшение  $H$  (поскольку точки, которые раньше были за пределами ящика, где  $V = +\infty$ , теперь оказываются внутри ящика, где  $V = 0$ ). Как обычно, давление считается положительным, из чего следует, что  $\omega = -p dV$ . Тогда последнее уравнение принимает вид

$$\partial Z/\partial V = \beta p = \nu,$$

а это показывает, что  $Z$  обладает всеми свойствами функции Массье.

Итак, у нас теперь есть общий рецепт вычисления функции Массье. Для заданной системы  $(M, \mathcal{A}, \mu)$  и функции  $H$ , определенной на пространстве  $M$  и называемой «энергией», определим функцию разбиения  $F$  выражением

$$F = \int e^{-\beta H} \mu. \quad (22.40)$$

Функция  $H$  и, следовательно,  $F$  могут зависеть от дополнительных переменных  $Q_1, Q_2, \dots, Q_k$ . Кроме того, интеграл в формуле (22.40) будет сходиться только в определенной части пространства  $(\beta, Q_1, Q_2, \dots, Q_k)$ . Определим внутреннюю энергию формулой (22.32), равновесные состояния — формулой (22.35) (где  $F$

определяется формулой (22.40)) и функцию Массье — формулой (22.36). Тогда справедлива формула (22.38), и форма для работы  $\omega$  задается формулой (22.39). Это и есть рецепт для перехода от «микроскопической модели», закодированной функцией  $H$ , к «макроскопическим равновесным явлениям», описываемым функцией Массье. Получен четкий рецепт, но его еще надо обосновать. Это можно будет сделать, когда мы поймем вероятностное определение энтропии, которым займемся в следующем параграфе.

Есть еще одно небольшое обобщение наших рассуждения, которое окажется полезным. Предположим, что мы хотим рассмотреть не одно наблюдаемое, а несколько наблюдаемых  $(J_1, \dots, J_n)$  сразу. В этом случае можно ввести  $\mathbf{J} = (J_1, \dots, J_n)$  — векторное наблюдаемое, значения которого лежат в пространстве  $\mathbb{R}^n$ , т. е. наблюдаемое, принимающее значения в векторном пространстве  $W$ . Наблюдаемое со значениями в  $W$  — это просто функция из  $M$  в  $W$  (удовлетворяющая соответствующему условию измеримости, как и в случае  $\mathbb{R}$ ). Понятие интеграла от вектор-функции определено, поэтому определение (22.28) для математического ожидания вектор-наблюдаемого в данном состоянии сохраняется без изменения. Точно так же можно говорить про вероятность (в данном состоянии) наблюдаемого  $\mathbf{J}$  лежать в некотором подмножестве  $W$ . Но нам не придется этим заниматься.

## 22.9. Энтропия

Введем понятие энтропии состояния. Пусть  $(M, \mathcal{A}, \mu)$  — система. Мы хотим каждому состоянию системы  $\rho$  приписать число  $\text{Ent}(\rho)$ , которое будет измерять меру «беспорядка» состояния  $\rho$  по отношению к  $\mu$ . Таким образом,  $\text{Ent}$  будет функцией, определенной на пространстве всех состояний. Чтобы оправдать такое определение, сначала рассмотрим пространство конечной выборки, т. е. случай  $A$  из параграфа 22.6. Если  $M = \{e_1, \dots, e_k\}$ , то состояние  $\rho$  определяется заданием  $k$  вещественных чисел  $p_i = \rho(e_i)$ ,  $i = 1, \dots, k$ . Тогда  $p_i$  есть вероятность события  $\{e_i\}$ . В теории информации энтропия состояния  $\rho = (p_1, \dots, p_k)$  определяется

равенством

$$\text{Ent}_k(\rho) = - \sum_{i=1}^k p_i \log p_i. \quad (22.41)$$

Эта функция обладает различными свойствами, делающими ее мерой «беспорядка» или недостатка информации. Выпишем некоторые из них.

- (i)  $\text{Ent}_k(p_1, \dots, p_k)$  симметрична по  $(p_1, \dots, p_k)$ . Это соответствует требованию, что мера беспорядка не зависит от способа индексации результата.
- (ii)  $\text{Ent}_k(1, 0, \dots, 0) = 0$ . (Здесь мы определяем  $x \log x = 0$  для  $x = 0$  в силу непрерывности.) Это соответствует утверждению, что если в данном состоянии результат предопределен, то оно имеет нулевой беспорядок.
- (iii)  $\text{Ent}_{k+1}(p_1, \dots, p_k, 0) = \text{Ent}_k(p_1, \dots, p_k)$ . Это означает, что если нашу систему  $M = \{e_1, \dots, e_k\}$  заменить на систему  $M' = \{e_1, \dots, e_k, e_{k+1}\}$ , но при этом рассматривать только состояния, в которых  $e_{k+1}$  невозможны, то энтропия не изменяется. Другими словами, добавление фиктивной альтернативы не изменяет недостаток информации.
- (iv)  $\text{Ent}_k(p_1, \dots, p_k) \leq \text{Ent}\left(\frac{1}{k}, \dots, \frac{1}{k}\right)$ , где будет строгое неравенство, если  $p_i \neq 1/k$  для некоторого значения  $i$ .

Таким образом, недостаток информации достигает максимума, когда все альтернативы равновероятны. Для доказательства этого утверждения заметим, что для всех  $x \geq 0$  справедливо условие

$$x \log x \geq x - 1.$$

Равенство будет только для  $x = 1$ , что легко проверяется дифференцированием обеих частей. При  $x \neq 0$  это условие можно переписать в виде

$$-\log x \leq x^{-1} - 1.$$

Тогда

$$-p_i \log p_i - \left( -p_i \log \frac{1}{k} \right) = -p_i \log \frac{p_i}{k^{-1}} \leq -p_i \left( \frac{k^{-1}}{p_i} - 1 \right) = -\frac{1}{k} + p_i.$$

При вычислении суммы по  $i$  справа получается 0, потому что  $\sum (1/k) = 1 = \sum p_i$ . Итак, мы имеем

$$-\sum p_i \log p_i \leq -\sum p_i \log \frac{1}{k} = -\log \frac{1}{k} = -\sum \frac{1}{k} \log \frac{1}{k}.$$

Заметим, что  $\text{Ent}(p_1, \dots, p_k)$  отражает недостаток знания до выполнения эксперимента, который скажет нам, какие из  $k$  возможностей реализуются на самом деле. В этом смысле  $\text{Ent}(p_1, \dots, p_k)$  представляет *количество информации*, полученной при выполнении эксперимента.

- (v) Предположим, что  $M = M_1 \times M_2$ , где  $M_1$  содержит  $k$  элементов  $\{e_1, \dots, e_k\}$ , а  $M_2$  содержит  $l$  элементов, так что  $M$  состоит из  $kl$  элементов  $\{e_i, f_j\}$ . Предположим, что  $\rho = \rho_1 \times \rho_2$ , так что  $\rho(\{(e_i, f_j)\}) = p_i q_j$ . Тогда

$$\begin{aligned} \text{Ent}_{kl}(\rho) &= -\sum p_i q_j \log(p_i q_j) \\ &= -\sum p_i q_j (\log p_i + \log q_j) \\ &= -\sum p_i \log p_i - \sum q_j \log q_j. \end{aligned}$$

Но поскольку  $\sum p_i = \sum q_j = 1$ , мы имеем

$$\text{Ent}_{kl}(\rho_1 \times \rho_2) = \text{Ent}_k(\rho_1) + \text{Ent}_l(\rho_2).$$

Другими словами, если мы проведем два независимых эксперимента, то полное количество полученной информации будет равно сумме информаций, полученных от каждого эксперимента.

Оказывается, что свойства (i)–(iv) вместе с более сильным вариантом утверждения (v) определяют функцию  $\text{Ent}$  с точностью

до постоянного множителя. Таким образом, в пространстве конечной выборки есть только один способ измерения «беспорядка» состояния, если эта мера беспорядка удовлетворяет нескольким разумным аксиомам. Мера беспорядка задается выражением (22.41). Это выражение можно переписать таким образом, что оно будет иметь смысл для любой системы.

По определению, для произвольной системы  $(M, \mathcal{A}, \mu)$  и произвольного состояния  $\rho$  имеем

$$\text{Ent}(\rho) = - \int \rho \log \rho \, \mu, \quad (22.42)$$

при условии, что этот интеграл сходится.

## 22.10. Равновесие в статистических системах

Итак, у нас есть определение меры «беспорядка» в системе. Предположим, что нам задана система и последовательность вещественных наблюдаемых  $J_1, \dots, J_n$ . Мы можем наблюдать ожидаемые значения этих наблюдаемых в любом состоянии. Соберем наблюдаемые в векторную конструкцию  $\mathbf{J} = (J_1, \dots, J_n)$ . Используя принцип максимальной энтропии, рассмотрим состояние «статистического равновесия», в котором «беспорядок» достигает максимума в соответствии с нашим знанием ожидаемого значения  $\mathbf{J}$ . Сначала поставим точную математическую задачу.

Пусть  $\mathbf{J} : M \rightarrow V$  — векторное наблюдаемое системы, и пусть  $\bar{\mathbf{J}}$  — произвольное возможное ожидаемое значение  $\mathbf{J}$ . Рассмотрим следующую проблему.

Среди всех состояний  $\rho$ , обладающих свойством

$$E(\mathbf{J}; \rho) = \int \rho \mathbf{J} \, \mu = \bar{\mathbf{J}},$$

найти такое, в котором энтропия достигает своего максимума.

Иначе говоря, нужно решить задачу на условный максимум  $\text{Ent}(\rho)$  при ограничении  $\int \rho \mathbf{J} \, \mu = \bar{\mathbf{J}}$ .

Оказывается, что если для системы  $(M, \mathcal{A}, \mu)$  и наблюдаемого  $\mathbf{J}$  выполняются ряд предположений, то эта задача имеет единственное решение. Итак, начнем.

Пусть  $V^*$  обозначает пространство, дуальное к  $V$ . Тогда элемент  $\gamma \in V^*$  — линейная функция на пространстве  $V$ . Пусть  $\gamma \cdot \mathbf{v}$  обозначает значение  $\gamma$  на векторе  $\mathbf{v}$ . Тогда для любых  $\gamma \in V^*$  и  $m \in M$  можно рассмотреть  $\gamma \cdot \mathbf{J}(m)$ , что является числом, зависящим от  $\gamma$  и  $m$ . Но тогда  $\gamma \cdot \mathbf{J}$  — числовая функция на множестве  $M$ . Запишем интеграл

$$F(\gamma) = \int_M e^{-\gamma \cdot \mathbf{J}} \mu.$$

Отсюда видно, что  $0 < F(\gamma) \leq +\infty$  (интеграл может расходиться к  $+\infty$ ). Предположим, что  $\gamma$  таково, что интеграл  $F(\gamma)$  конечен. Тогда

$$\rho_\gamma = \frac{1}{F(\gamma)} e^{-\gamma \cdot \mathbf{J}} \quad (22.43)$$

определяет состояние системы. Действительно,  $\rho_\gamma$  — положительная функция, и по определению  $F(\gamma)$  мы имеем  $\int \rho_\gamma \mu = 1$ . Для каждого значения  $\gamma$ , когда  $F(\gamma) < +\infty$ , мы получаем состояние  $\rho_\gamma$ . Итак, мы получили набор состояний, параметризованный подмножеством  $V^*$ . Заметим, что если  $F(\gamma) < +\infty$ , то  $\text{Ent}(\rho_\gamma)$  конечна. Действительно,

$$-\int \rho_\gamma \log \rho_\gamma \mu = -\int \rho_\gamma (-\log F(\gamma) - \gamma \cdot \mathbf{J}) \mu.$$

Здесь  $F(\gamma)$  постоянна, если рассматривать ее как функцию на множестве  $M$ . Поэтому

$$\int \rho_\gamma \log F(\gamma) \mu = \log F(\gamma).$$

Кроме того,

$$\int \rho_\gamma \mathbf{J} \mu = E(\mathbf{J}; \rho_\gamma)$$

есть просто ожидаемое значение векторного наблюдаемого  $\mathbf{J}$  в состоянии  $\rho$ . Поэтому можно написать

$$\int \rho_\gamma \gamma \cdot \mathbf{J} \mu = \gamma \cdot E(\mathbf{J}; \rho_\gamma).$$

Но тогда

$$\text{Ent}(\rho_\gamma) = \log F(\gamma) + \gamma \cdot E(\mathbf{J}; \rho_\gamma). \quad (22.44)$$

Давайте теперь учтем, что функция  $\mathbf{J}$  может также зависеть от некоторых «конфигурационных» параметров  $Q$ . Если написать  $S(\gamma, Q) = \text{Ent}(\rho_\gamma, Q)$  и  $Z(\gamma, Q) = \log F(\gamma, Q)$ , то равенство (22.44) принимает вид

$$S = Z + \gamma \cdot \mathbf{J}(\gamma, Q).$$

В частном случае, когда  $\mathbf{J} = H$ , это равенство совпадает с (22.24). Но формула (22.24) представляет собой комбинацию первого и второго законов термодинамики.

Предположим, что для заданного значения  $\bar{\mathbf{J}} \in V$  мы можем найти такое  $\gamma \in V^*$ , что

$$\int \mathbf{J} \rho_\gamma \mu = E(\mathbf{J}; \rho_\gamma) = \bar{\mathbf{J}}.$$

(При этом предполагается, что  $\int \mathbf{J} \rho_\gamma \mu$  сходится абсолютно.)

Тогда утверждается, что  $\rho$  — единственное решение экстремальной задачи. Другими словами, справедлива

**Теорема.** Пусть  $\gamma \in V^*$  таково, что  $F(\gamma) < +\infty$  и справедливо равенство (22.44), где  $\rho_\gamma$  определено формулой (22.43). И пусть  $\rho$  — другое состояние, для которого

$$\int \rho \mathbf{J} \mu = \bar{\mathbf{J}}. \quad (22.45)$$

Тогда

$$\text{Ent}(\rho) \leq \text{Ent}(\rho_\gamma),$$

причем мы имеем строгое неравенство, если  $\rho$  и  $\rho_\gamma$  отличаются на множестве с положительной мерой  $\mu$ .

*Доказательство.* Здесь мы повторяем аргументы, которые приводили для доказательства, что  $(1/k, \dots, 1/k)$  дает максимум энтропии в конечной системе. Итак, для любого состояния  $\rho$  и любого  $m \in M$  мы имеем

$$-\rho(m) \log \rho(m) - (-\rho(m) \log \rho_\gamma(m)) \leq \rho_\gamma(m) - \rho(m). \quad (22.46)$$

В самом деле, если  $\rho(m) = 0$ , то левая часть равенства равна нулю, а правая положительна. Если  $\rho(m) > 0$ , то

$$-\rho(m) \log \rho(m) - (-\rho(m) \log \rho_\gamma(m)) = \rho(m) \log[\rho(m)/\rho_\gamma(m)] \\ \leq \rho(m)([\rho_\gamma(m)/\rho(m)] - 1) = \rho_\gamma(m) - \rho(m),$$

при этом используется неравенство  $-\log x \leq x^{-1} - 1$ .

Интегрирование по пространству  $M$  дает

$$-\int \rho \log \rho \mu - \int \rho \log \rho_\gamma \mu \leq 0,$$

потому что

$$\int \rho \mu = \int \rho_\gamma \mu = 1.$$

Но

$$\log \rho_\gamma = -\log F(\gamma) - \gamma \cdot \mathbf{J},$$

а следовательно,

$$-\int \rho \log \rho_\gamma \mu = \log F(\gamma) + \int \rho(\gamma \cdot \mathbf{J}) \mu = \log F(\gamma) + \gamma \cdot \left( \int \rho \mathbf{J} \mu \right) \\ = \log F(\gamma) + \gamma \cdot \mathbf{J} = \text{Ent}(\rho_\gamma),$$

поскольку  $E(\mathbf{J}; \rho_\gamma) = \bar{\mathbf{J}}$ .

Таким образом,

$$\text{Ent}(\rho) \leq \text{Ent}(\rho_\gamma).$$

(Заметим, также, что из неравенства (22.46) следует, что

$$-\rho \log \rho \leq -\rho[\log F(\gamma) + \gamma \cdot \mathbf{J}] + \rho_\gamma - \rho,$$

причем правая часть этого неравенства — абсолютно интегрируемая функция. Итак, если  $-\int \rho \log \rho \mu$  расходится, то это происходит на  $-\infty$ .)

Неравенство (22.46) будет строгим, если  $\rho(m)/\rho_\gamma(m) \neq 1$ . Тогда, если  $\rho(m) \neq \rho_\gamma(m)$  для множества с положительной мерой, то мы имеем  $\text{Ent}(\rho) < \text{Ent}(\rho_\gamma)$ , что и требовалось доказать.

В качестве следствия этой теоремы мы дадим статистическое определение равновесного состояния.

**Определение.** Пусть  $\mathbf{J}$  — наблюдаемое в системе  $(M, \mathcal{A}, \mu)$ . Если  $\gamma$  таково, что  $F(\gamma) = \int e^{-\gamma \cdot \mathbf{J}} \mu < \infty$ , то

$$\rho_\gamma = \frac{1}{F(\gamma)} e^{-\gamma \cdot \mathbf{J}}$$

называется *равновесным состоянием системы* (относительно наблюдаемого  $\mathbf{J}$ ).

Мы должны показать, что (при соответствующих предположениях) для любого значения  $\bar{\mathbf{J}}$  существует состояние  $\rho_\gamma$ , для которого  $E(\mathbf{J}; \rho_\gamma) = \bar{\mathbf{J}}$ . Из доказательства теоремы следует, что если такое состояние  $\rho_\gamma$  существует, то оно единственно. Мы вернемся к этому вопросу позднее. А сначала дадим несколько примеров равновесных состояний, используя системы, введенные в примерах параграфа 22.6.

A1. Пусть  $M = \{e_1, \dots, e_k\}$  и  $\mu$  точно такие же, как и в примере A параграфа 22.6. Пусть  $V = \mathbb{R}$  — одномерное пространство и пусть  $\mathbf{J} : M \rightarrow \mathbb{R}$  задается условием  $\mathbf{J}(e_i) = \varepsilon_i$ , где  $\varepsilon_i$  — вещественные числа (мы будем считать их «энергетическими уровнями»). Без потери общности для элементов множества  $M$  введем индексы

$$\varepsilon_1 \leq \varepsilon_2 \leq \dots \leq \varepsilon_k.$$

Тогда  $V^* = \mathbb{R}$  и для любого  $\beta \in V^*$  мы имеем

$$F(\beta) = \sum_{p=1}^k e^{-\beta \varepsilon_p},$$

так что равновесное состояние  $\rho_\beta$  задается формулой

$$\rho_\beta(e_i) = \frac{e^{-\beta \varepsilon_i}}{\sum_{p=1}^k e^{-\beta \varepsilon_p}}.$$

Оно называется «распределением Максвелла–Больцмана».

Вспомним вычисление ожидаемого значения «энергии»

$$E(\mathbf{J}; \rho_\beta) = \frac{\sum_i \varepsilon_i e^{-\beta \varepsilon_i}}{\sum_i e^{-\beta \varepsilon_i}},$$

или короче

$$E(\mathbf{J}; \rho_\beta) = -\frac{\partial \log F(\beta)}{\partial \beta}.$$

(Как мы уже видели, это последнее равенство дает общее правило.) Тогда

$$\begin{aligned} -\frac{\partial E(\mathbf{J}; \rho_\beta)}{\partial \beta} &= \frac{\sum \varepsilon_i^2 e^{-\beta \varepsilon_i}}{\sum e^{-\beta \varepsilon_i}} - \left( \frac{\sum \varepsilon_i e^{-\beta \varepsilon_i}}{\sum e^{-\beta \varepsilon_i}} \right)^2 \\ &= \int \|\mathbf{J}\|^2 \rho_\beta \mu - \left( \int \mathbf{J} \rho_\beta \mu \right)^2 \\ &= \int (\mathbf{J} - \bar{\mathbf{J}})^2 \rho_\beta \mu \leq 0, \end{aligned}$$

причем будет строгое неравенство, если только не все  $\varepsilon_i$  равны между собой. (Если же все  $\varepsilon_i$  равны, то существует только одно возможное значение для  $\mathbf{J}$ , равное этому общему значению всех  $\varepsilon_i$ . Тогда все  $\rho_\beta$  одинаковы и равны  $\rho_\beta(e_i) = 1/k$ . Таким образом, в этом случае теорема сводится к утверждению (iii).) Если  $\varepsilon_1 \neq \varepsilon_k$ , то  $E(\mathbf{J}; \rho_\beta)$  — строго убывающая функция от  $\beta$ . Очевидно, что

$$\lim_{\beta \rightarrow -\infty} E(\mathbf{J}; \rho_\beta) = \varepsilon_k \quad \text{и} \quad \lim_{\beta \rightarrow +\infty} E(\mathbf{J}; \rho_\beta) = \varepsilon_1.$$

Итак, любое значение  $\bar{\mathbf{J}}$  такое, что  $\varepsilon_1 < \bar{\mathbf{J}} < \varepsilon_k$ , может быть получено только единственным выбором  $\beta$ . Для любого значения  $\beta$  мы имеем

$$\begin{aligned} \text{Ent}(\rho_\beta) &= -\rho_\beta(e_i) \log \rho_\beta(e_i) \\ &= -\sum_i \frac{e^{-\beta \varepsilon_i}}{\sum_p e^{-\beta \varepsilon_p}} \left[ \log(e^{-\beta \varepsilon_i}) - \log \left( \sum_j e^{-\beta \varepsilon_j} \right) \right] \\ &= \frac{\sum \beta \varepsilon_i e^{-\beta \varepsilon_i}}{\sum e^{-\beta \varepsilon_i}} + \log F(\beta) = \beta \cdot \bar{\mathbf{J}} + \log F(\beta). \end{aligned}$$

Это, конечно, частный случай формулы (22.24).

A2. Рассмотрим случай конечной системы  $M$ , элементы которой для удобства обозначим  $e_0, \dots, e_k$ , так что в ней насчитывается  $k+1$  элемент. Пусть  $\mathbf{J} = (J_1, \dots, J_k)$  отображает  $M$  в

пространство  $\mathbb{R}^k$ , где

$$J_i(e_j) = \delta_i^j.$$

Таким образом,  $\mathbf{J}$  отображает  $M$  в вершины симплекса в пространстве  $\mathbb{R}^k$ , заданного базисными векторами.

В этом случае  $\gamma = (\gamma_1, \dots, \gamma_k)$  и

$$F(\gamma) = 1 + e^{-\gamma_1} + \dots + e^{-\gamma_k}$$

сходится для всех значений  $\gamma$  с

$$\rho_\gamma(e_0) = \frac{1}{F(\gamma)}, \quad \rho_\gamma(e_1) = \frac{e^{-\gamma_1}}{F(\gamma)}, \quad \dots, \quad \rho_\gamma(e_k) = \frac{e^{-\gamma_k}}{F(\gamma)}.$$

Если мы обозначим  $\rho_\gamma(e_i)$  через  $p_i$ , то понятно, что в любую точку  $\mathbf{p} = (p_1, \dots, p_k)$  внутри симплекса можно попасть с помощью соответствующего выбора  $\gamma_i$ . В этом случае все состояния, соответствующие внутренней точке  $\mathbf{p}$ , являются равновесными состояниями.

Отображение  $\gamma \rightarrow E(\mathbf{J}; \rho_\gamma)$  переводит все элементы пространства  $\mathbb{R}^n = V^*$  во внутреннюю область единичного симплекса. Читатель может проверить, что это отображение взаимнооднозначно и имеет дифференцируемое обратное.

В. Пусть  $(\mathbb{R}, \mathcal{A}, \mu)$  — вещественная прямая с обычной мерой. Если  $\mathbf{J}$  — тождественное отображение  $\mathbf{J}(x) = x$ , то соответствующая функция  $Z(\gamma) = \int_{-\infty}^{\infty} e^{-\gamma x} dx$  расходится для всех значений  $\gamma$ . Однако, если рассмотреть отображение  $\mathbf{J} : \mathbb{R} \rightarrow \mathbb{R}^2$ , заданное равенством  $\mathbf{J}(x) = (x, x^2)$ , то для всех  $\gamma = (\gamma_1, \gamma_2)$  с  $\gamma_2 > 0$  функция

$$F(\gamma) = \int_{-\infty}^{\infty} e^{-\gamma_1 x - \gamma_2 x^2} \mu$$

сходится. Соответствующее равновесное состояние

$$\rho_\gamma = F(\gamma)^{-1} e^{-\gamma_1 x - \gamma_2 x^2} = F(\gamma)^{-1} e^{\gamma_1^2/4\gamma_2} e^{-(x+\gamma_1/2\gamma_2)^2/2(1/2\gamma_2)}$$

имеет нормальную плотность, математическое ожидание  $m = -\gamma_1/2\gamma_2$  и дисперсию<sup>21</sup>  $\sigma = (2\gamma_2)^{-1/2}$ . Очевидно, что для любого

<sup>21</sup>Напомним, что дисперсия случайной величины (функции на пространстве с вероятностной мерой)  $f$  с математическим ожиданием  $m$  определяется как математическое ожидание величины  $(f - m)^2$ . — Прим. ред.

$\sigma > 0$  и любого  $m$  можно найти такой вектор  $\gamma = (\gamma_1, \gamma_2)$ , что

$$\rho_\gamma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}.$$

Итак, для всех случайных величин с производными, имеющих заданные математическое ожидание и дисперсию, энтропия максимальна при нормальных распределениях. Чтобы вычислить энтропию, заметим, что в силу трансляционной инвариантности можно считать, что  $m = 0$ . Тогда

$$\begin{aligned} \text{Ent}(\rho_\gamma) &= \frac{1}{\sigma\sqrt{2\pi}} \int e^{-x^2/2\sigma^2} \left[ \frac{1}{2} \log 2\pi + \log \sigma + \frac{x^2}{2\sigma^2} \right] dx \\ &= \log \sigma + \frac{1}{2} \log 2\pi + 1. \end{aligned}$$

Энтропия стремится к  $-\infty$  при  $\sigma \rightarrow 0$ . Это соответствует тому, что при  $\sigma \rightarrow 0$  состояние  $\rho$  становится более сосредоточенным вокруг точки и поэтому требуется «бесконечное количество информации, чтобы выделить эту точку из континуума».

С. Давайте рассмотрим наблюдаемые  $J_1 = x$  и  $J_2 = \log x$  в пространстве  $\mathbb{R}^+$ . Тогда  $\gamma = (\gamma_1, \gamma_2)$ ,  $\mathbf{J} = (J_1, J_2)$  и

$$F(\gamma) = \int_0^\infty e^{-\gamma_1 x - \gamma_2 \log x} dx = \int_0^\infty e^{-\gamma_1 x} x^{-\gamma_2} dx$$

сходится при  $\gamma_1 > 0$  и  $\gamma_2 < -1$ .

Если положить  $k = -\gamma_2 + 1$  и  $y = \gamma_1 x$ , то интеграл принимает вид

$$F(\gamma) = \int_0^\infty e^{-\gamma_1 x} x^{k-1} dx = \gamma_1^{-k} \int_0^\infty e^{-y} y^{k-1} dy = \gamma_1^{-k} \Gamma(k).$$

Соответствующие распределения называются *гамма-распределениями*. В этом случае плотность  $\rho_\gamma$  равна

$$\rho_\gamma = \frac{1}{\sqrt{Z(\gamma)}} e^{-\gamma_1 x - \gamma_2 \log x}.$$

Мы хотим вычислить векторное ожидаемое значение

$$E(\mathbf{J}) = (E(J_1), E(J_2)).$$

При этом мы имеем

$$E(J_1) = \int_0^\infty x e^{-\gamma_1 x - \gamma_2 \log x} dx$$

и

$$E(J_2) = \int_0^\infty \log x \cdot e^{-\gamma_1 x - \gamma_2 \log x} dx.$$

Тогда

$$E(J_1) = \int_0^\infty x^{1-\gamma_2} e^{-\gamma_1 x} dx.$$

Сново обозначим  $y = \gamma_1 x$  и  $k = 1 - \gamma_2$ , что дает

$$E(J_1) = \gamma_1^{k-1} \int_0^\infty e^{-y} y^k dy = \gamma_1^{k-1} \Gamma(k+1).$$

Вычисление  $E(J_2)$  оставляем читателю в качестве упражнения.

D. Пусть  $\mathbf{J}$  — наблюдаемое, заданное формулой  $\mathbf{J}(k) = k$ , а  $\beta \in \mathbb{R}^1$ . Тогда мы видим, что ряд

$$F(\beta) = \sum \frac{e^{-\beta \cdot k}}{k!} = e^{-\beta}$$

сходится для всех значений  $\beta$ . Введем величину  $e^{-\beta} = \lambda$ ; тогда  $F(\beta) = e^\lambda$ . (Параметр  $\lambda$  часто называют «активностью».) Тогда

$$\rho_\beta = \frac{1}{F(\beta)} e^{-\beta \cdot k}.$$

Поэтому

$$\text{Prob}\{\mathbf{J}\rho_\beta = k\} = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Это распределение вероятностей называется *распределением Пуассона*. Заметим, что

$$E(\mathbf{J}; \rho_\beta) = -\partial \log F(\beta) / \partial \beta = -\partial e^{-\beta} / \partial \beta = e^{-\beta}$$

или

$$E(\mathbf{J}; \rho_\beta) = \lambda.$$

Другими словами, параметр  $\lambda$  показывает ожидаемое число частиц в ящике.

Давайте теперь еще раз перечислим полученные теоретические результаты.

- Система  $(M, \mathcal{A}, \mu)$  является пространством с мерой.
- Состояние — неотрицательная функция  $\rho$  на  $M$  с условием  $\int \rho \mu = 1$ .
- Энтропия состояния есть  $\text{Ent}(\rho) = - \int \rho \log \rho \mu$ .
- Наблюдаемое есть (векторная) функция на пространстве  $M$ .
- Ожидаемое значение наблюдаемого  $K$  в состоянии  $\rho$  есть  $E(K; \rho) = \int K \rho \mu$ .

Если  $\mathbf{J}$  — наблюдаемое со значениями в пространстве  $V$ , то функция разбиения, связанная с  $\mathbf{J}$ , есть функция, определенная для  $\gamma \in V^*$  условием

$$F(\gamma) = \int e^{-\gamma \cdot \mathbf{J}} \mu.$$

Она определена на подмножестве  $C$  пространства  $V^*$ , где этот интеграл сходится.

Функция Массье  $Z$  определена на  $C$  условием  $Z = \log F$ . Равновесные состояния относительно  $\mathbf{J}$  — это состояния вида

$$\rho_\gamma = F(\gamma)^{-1} e^{-\gamma \cdot \mathbf{J}}$$

для  $\gamma \in C$ . С помощью функции Массье можно вычислить ожидаемое значение  $\mathbf{J}$  в состоянии  $\rho_\gamma$  при условии, что  $\gamma$  находится внутри подмножества  $C$ , по формуле

$$J_i = -\partial Z / \partial \gamma_i,$$

что для краткости можно записать в виде

$$\mathbf{J} = -\partial Z / \partial \gamma.$$

Тогда

$$S(\gamma) = \text{Ent}(\rho_\gamma) = Z + \gamma \cdot \mathbf{J}.$$

В частности,  $\gamma$  можно получить из  $\mathbf{J}$  по формуле

$$\gamma = \partial S / \partial \mathbf{J}.$$

Здесь мы рассматриваем  $S$  как функцию от  $\mathbf{J}$ , поскольку  $\gamma$  можно выразить как функцию от  $\mathbf{J}$ , обращая заданную функцию  $\mathbf{J} = \mathbf{J}(\gamma)$ . Наблюдаемое  $\mathbf{J}$  может зависеть от дополнительных координат  $Q_1, \dots, Q_d$ , т. е.  $\mathbf{J}$  может быть функцией на пространстве  $M \times B$ , где  $B$  — пространство с координатами  $Q_1, \dots, Q_d$ . Тогда функция  $F$  и равновесные состояния  $\rho_\gamma$  также зависят от этих дополнительных координат.

В случае, когда одной из компонент  $\mathbf{J}$  является  $H$  — «энергия», *внутренняя энергия* состояния  $\rho$  дается  $E(H; \rho)$ . Имея все эти определения и после обсуждения работы в параграфе 22.8 (см. (22.39a)), мы подготовились в переходе от «микроскопической» модели к «макроскопически» наблюдаемым явлениям.

## 22.11. Квантовые и классические газы

Начнем со сравнения функций разбиения для трех систем D, E и F, рассмотренных в параграфе 22.8. Как и раньше, для каждой из этих систем будем рассматривать двумерное наблюдаемое  $\mathbf{J} = (N, H)$ , где  $N$  — «число заполнения»,  $H$  — «уровень энергии», причем  $N$  и  $H$  связаны соотношением

$$H = N\varepsilon,$$

где  $\varepsilon$  — «энергия одного заполненного состояния». В предыдущем параграфе мы уже проделали вычисления для случая D, а теперь сделаем все сразу в виде таблицы 22.2.

Во всех трех случаях мы вычисляем последний столбец по формуле  $E(N) = -\partial \log F / \partial \beta_1$ . Сумма, возникающая при вычислении случая F, является геометрической прогрессией, а в случае E — это просто сумма двух слагаемых. Во всех трех случаях  $E(N)$  мала тогда и только тогда, когда велика  $e^{-(\mu-\varepsilon)/T}$ . Тогда мы пренебрегаем слагаемым  $\pm 1$ , появляющимся в знаменателе выражения для  $Z$ . Но тогда распределения вероятностей наблюдаемого  $N$  очень близки. В таблице 22.3 мы введем обозначение

$$\lambda = e^{(\mu-\varepsilon)/T}.$$

Если рассматривать ожидаемое число заполнения как меру «концентрации», то для «разреженных» систем распределения вероятностей Больцмана, Ферми–Дирака и Бозе–Эйнштейна

Таблица 22.2

Система	$\mu(k)$	Функция разбиения	$E(N)$
D	$1/k!$	$\exp[e^{-(\beta_1+\beta_2\varepsilon)}]$ $= \exp[e^{(\mu-\varepsilon)/T}]$	$e^{(\mu-\varepsilon)/T}$
E	$\mu(0) = \mu(1) = 1$ $\mu(k) = 0, k > 1$	$1 + e^{-(\beta_1+\beta_2\varepsilon)}$ $= 1 + e^{(\mu-\varepsilon)/T}$	$\frac{1}{1+e^{-(\mu-\varepsilon)/T}}$
F	1	$\frac{1}{1 - e^{-(\beta_1+\beta_2\varepsilon)}}$ $= \frac{1}{1 - e^{(\mu-\varepsilon)/T}}$	$\frac{1}{e^{-(\mu-\varepsilon)/T} - 1}$

Таблица 22.3

Система	Название	$\text{Prob}(N = k)$
D	Больцмана–Пуассона	$e^{-\lambda} \lambda^k / k!$
E	Ферми–Дирака	$\text{Prob}(N = 0) = 1/(1 + \lambda)$ $\text{Prob}(N = 1) = \lambda/(1 + \lambda)$ $\text{Prob}(N > 1) = 0$
F	Бозе–Эйнштейна	$(1 - \lambda) \lambda^k$

очень близки, т. е. все члены в правом столбце таблицы 22.3 будут близки друг к другу, когда  $\lambda$  очень мало. Но для больших «концентраций» они сильно отличаются. На рис. 22.16 изображены графики трех функций  $e^{-t}$ ,  $1/(e^t + 1)$  и  $1/(e^t - 1)$ . Заметим, что при  $t > 2$  кривые почти сливаются, но при малых значениях  $t$  они сильно расходятся.

Давайте рассмотрим ситуацию, когда имеется  $d$  экземпляров одной из этих систем, причем каждая имеет различное значение энергии одного заполненного состояния, так что мы имеем  $d$  значений  $\varepsilon_1, \dots, \varepsilon_d$ . Система представляет собой прямое произведение

$$M = M_1 \times \dots \times M_d.$$

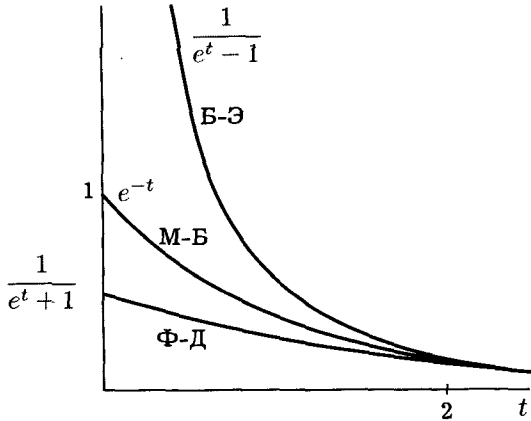


Рис. 22.16. Графики ожидаемых значений числа частиц для распределений Больцмана, Ферми-Дирака и Бозе-Эйнштейна в зависимости от  $(\mu - \varepsilon)/T = t$

Мы предполагаем, что «энергия взаимодействия» пренебрежимо мала, т. е.

$$H = H_1 + \dots + H_d.$$

Кроме того, определим

$$N = N_1 + \dots + N_d.$$

Тогда функция разбиения всей системы есть просто произведение функций разбиения всех компонент:

$$F = F_1 \times \dots \times F_d,$$

где

$$F_i = F_{\varepsilon_i}(\beta_1, \beta_2).$$

В таблице 22.2 мы должны в выражениях для функций разбиения вместо  $\varepsilon$  подставить  $\varepsilon_i$ .

Давайте еще раз обсудим, что это означает. У нас есть  $d$  систем. Мы собрали их в одну систему. Следовательно, они могут «взаимодействовать». Равновесные состояния  $i$ -й подсистемы определялись бы ее собственными параметрами  $\beta_{1i}$  и  $\beta_{2i}$ . Тот факт, что системы находятся в равновесии, означает, что у нас

имеются общие для всех значения  $\beta_1$  и  $\beta_2$ . Таким образом, мы еще раз открыли критерий Гиббса (22.26).

До сих пор шел весьма абстрактный разговор, и мы не давали физического описания компонент нашей системы. Сейчас рассмотрим следующую модель «газа в ящике» (отличную от прежней модели). В фазовом пространстве рассмотрим подобласть  $B \times \square$ , где  $B \subset \mathbb{R}^3$  — наш ящик и  $\square \in \mathbb{R}^3$  — небольшой кубик с центром в точке  $\mathbf{p}$  в пространстве импульсов. Итак, когда мы говорим, что молекула находится в нашей области, то имеем в виду, что она находится в ящике и ее импульсы близки к  $\mathbf{p}$ . Подсистемы «взаимодействуют» посредством столкновений, т. е. за счет обмена импульсами. Предположим, что энергия одной частицы в выбранной подобласти есть ее кинетическая энергия, т. е.

$$\varepsilon(\mathbf{p}) = \frac{1}{2} \|\mathbf{p}\|^2 = \frac{1}{2} (p_x^2 + p_y^2 + p_z^2).$$

Это и есть «энергия заполненного состояния» для подсистемы с центром в точке  $\mathbf{p}$ . Согласно утверждению Больцмана–Пуассона «ожидаемое число частиц в нашей подсистеме» пропорционально

$$\exp \left[ \left( \mu - \frac{1}{2} \|\mathbf{p}\|^2 \right) / T \right].$$

Это и есть распределение Максвелла–Больцмана, где дополнительный параметр  $\mu$  зависит от плотности газа. Конечно, мы перепрыгнули через две технических ступени — мы рассмотрели бесконечное число подсистем, заполняющих все фазовое пространство, и устремили размер ячейки в фазовом пространстве к нулю. Но это не представляет большой сложности.

Эти же соображения можно было бы применить к статистикам Ферми–Дирака и Бозе–Эйнштейна. Тогда вместо распределения Максвелла–Больцмана мы получили бы распределение Ферми–Дирака

$$f(\mathbf{p}) = \frac{1}{\exp[(\frac{1}{2} \|\mathbf{p}\|^2 - \mu)/T] + 1}.$$

Аналогичная замена происходит и в случае статистики Бозе–Эйнштейна.

## 22.12. Определители и следы

Давайте перепишем формулу

$$F(\beta_1, \beta_2) = F_1(\beta_1, \beta_2) \times F_2(\beta_1, \beta_2) \times \cdots \times F_d(\beta_1, \beta_2) \quad (22.47)$$

(в каждом из трех случаев) на первый взгляд странным образом, но так, чтобы из нее были видны следствия. Пусть  $V = \mathbb{R}^d$  и  $X$  — диагональная матрица,  $i$ -й элемент которой равен  $\varepsilon_i$ .

Сначала рассмотрим случай Больцмана–Пуассона. Поскольку произведение экспонент равняется экспоненте от суммы, мы имеем

$$\begin{aligned} F(\beta_1, \beta_2) &= \exp[e^{-(\beta_1 + \beta_2 \varepsilon_1)} + \cdots + e^{-(\beta_1 + \beta_2 \varepsilon_d)}] \\ &= \exp(\operatorname{tr} e^{-(\beta_1 \mathbb{I} + \beta_2 X)}) \\ &= \operatorname{Det}(\exp e^{-(\beta_1 \mathbb{I} + \beta_2 X)}). \end{aligned}$$

Здесь использовано, что для любой матрицы имеет место равенство  $\operatorname{Det}(\exp A) = \exp(\operatorname{tr} A)$ .

Давайте рассмотрим случай Ферми–Дирака. Поскольку определитель диагональной матрицы равен произведению ее диагональных элементов, можно написать

$$F(\beta_1, \beta_2) = \operatorname{Det} \begin{pmatrix} F_1(\beta_1, \beta_2) & 0 & \cdots & 0 \\ 0 & F_2(\beta_1, \beta_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & F_d(\beta_1, \beta_2) \end{pmatrix} \quad (22.48)$$

Матрица, стоящая справа, — это матрица

$$\mathbb{I} + e^{-(\beta_1 \mathbb{I} + \beta_2 X)},$$

и поэтому

$$F = \operatorname{Det}(\mathbb{I} + e^{-(\beta_1 \mathbb{I} + \beta_2 X)}).$$

Аналогично, в случае Бозе–Эйнштейна соответствующее равенство имеет вид

$$F = \operatorname{Det}[\mathbb{I} - e^{-(\beta_1 \mathbb{I} + \beta_2 X)}]^{-1}.$$

Другими словами, мы получили таблицу 22.4.

Таблица 22.4

Система	Функция разбиения
Больцмана	$\text{Det}(\exp e^{-(\beta_1 \mathbb{I} + \beta_2 X)})$
Ферми-Дирака	$\text{Det}(\mathbb{I} + e^{-(\beta_1 \mathbb{I} + \beta_2 X)})$
Бозе-Эйнштейна	$\text{Det}[\mathbb{I} - e^{-(\beta_1 \mathbb{I} + \beta_2 X)}]^{-1}$

Если сравнить таблицы 22.2 и 22.4, то увидим, что изменение состоит в том, что скаляр  $\beta_1 + \beta_2 \varepsilon$  заменяется на матрицу  $\beta_1 \mathbb{I} + \beta_2 X$  и вычисляется определитель.

Опишем несколько по-другому процедуру взятия определителя. Но сначала сделаем ряд замечаний относительно линейной алгебры. Пусть  $V$  — векторное пространство, и  $A$  — его линейное преобразование. Тогда  $A$  индуцирует линейное преобразование  $A \otimes A$  пространства  $V \otimes V$  согласно формуле

$$(A \otimes A)(u \otimes v) = Au \otimes Av.$$

Если  $A = e^{tY}$ , то, дифференцируя уравнение

$$e^{tY} \otimes e^{tY}(u \otimes v) = e^{tY} u \otimes e^{tY} v$$

по  $t$  в точке  $t = 0$ , получаем

$$\begin{aligned} (d/dt)[e^{tY} \otimes e^{tY}(u \otimes v)]|_{t=0} &= Yu \oplus v + u \otimes Yv \\ &= (Y \otimes \mathbb{I} + \mathbb{I} \otimes Y)[u \otimes v]. \end{aligned}$$

Это показывает, что  $e^{tY} \otimes e^{tY} = \exp(tZ)$ , где  $Z$  — оператор, определенный равенством

$$Y \otimes \mathbb{I} + \mathbb{I} \otimes Y.$$

Другими словами,

$$e^{tY} \otimes e^{tY} = e^{t[Y \otimes \mathbb{I} + \mathbb{I} \otimes Y]}.$$

Аналогично, если рассмотреть индуцированное действие на  $V \otimes V \otimes V$ , то получим

$$e^{tY} \otimes e^{tY} \otimes e^{tY} = e^{t[Y \otimes \mathbb{I} \otimes \mathbb{I} + \mathbb{I} \otimes Y \otimes \mathbb{I} + \mathbb{I} \otimes \mathbb{I} \otimes Y]}.$$

Далее идут аналогичные выражения для  $V \otimes V \otimes V \otimes V$  и т. д. Введем обозначение  $D_k(Y)$  для матрицы, стоящей в экспоненте справа. Тогда

$$\begin{aligned} D_2(Y) &= Y \otimes \mathbb{I} + \mathbb{I} \otimes Y && \text{на } V \otimes V, \\ D_3(Y) &= Y \otimes \mathbb{I} \otimes \mathbb{I} + \mathbb{I} \otimes Y \otimes \mathbb{I} + \mathbb{I} \otimes \mathbb{I} \otimes Y && \text{на } V \otimes V \otimes V \end{aligned}$$

и т. д. Например, если взять  $t = 1$  на  $V \otimes V \otimes V$ , то получим

$$e^Y \otimes e^Y \otimes e^Y = e^{D_3(Y)}.$$

Вычисление следа матрицы, стоящей слева, дает  $(\operatorname{tr} e^Y)^3$ , поскольку  $\operatorname{tr}(A \otimes B \otimes C) = (\operatorname{tr} A)(\operatorname{tr} B)(\operatorname{tr} C)$ , т. е. под знаком тензорного произведения следы умножаются. Итак,

$$(\operatorname{tr} e^Y)^3 = \operatorname{tr} e^{D_3(Y)}.$$

Давайте положим

$$D_0(Y) = 0, \quad D_1(Y) = Y,$$

так что для всех  $k$

$$(\operatorname{tr} Y)^k = \operatorname{tr} e^{D_k(Y)}.$$

Теперь функцию разбиения для больцмановской составной системы можно написать в виде

$$\begin{aligned} \operatorname{Det}(\exp e^{-(\beta_1 \mathbb{I} + \beta_2 X)}) &= \exp(\operatorname{tr} e^{-(\beta_1 \mathbb{I} + \beta_2 X)}) \\ &= \sum (\operatorname{tr}[e^{-(\beta_1 \mathbb{I} + \beta_2 X)}])^k / k! \\ &= \sum (1/k!) \operatorname{tr}\{\exp(D_k[-(\beta_1 \mathbb{I} + \beta_2 X)])\}, \end{aligned}$$

где мы взяли  $Y = e^{-(\beta_1 \mathbb{I} + \beta_2 X)}$ . Запишем эту формулу более компактно. Пусть  $T_k(V)$  —  $k$ -кратное тензорное произведение  $V$  на себя, так что

$$\begin{aligned} T_0(V) &= \mathbb{R}, \\ T_1(V) &= V, \\ T_2(V) &= V \otimes V, \\ T_3(V) &= V \otimes V \otimes V, \quad \text{и т. д.} \end{aligned}$$

Рассмотрим полную тензорную алгебру

$$T(V) = T_0(V) \oplus T_1(V) \oplus T_2(V) \oplus \dots$$

(Это бесконечномерное пространство, но не надо беспокоиться.) Любая матрица на этом пространстве разлагается на блоки

$$A = \begin{pmatrix} A_{00} & A_{01} & A_{02} & A_{03} & \dots \\ A_{10} & A_{11} & A_{12} & A_{13} & \dots \\ A_{20} & A_{21} & A_{22} & A_{23} & \dots \\ A_{30} & A_{31} & A_{32} & A_{33} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

где  $A_{00}$  — преобразование  $T_0(V)$  (т.е. скаляр),  $A_{11}$  — линейное преобразование  $T_1(V)$ ,  $A_{22}$  — линейное преобразование  $T_2(V)$  и т.д. Определим

$$\text{tr}^\otimes A = \sum (1/k!) \text{tr} A_{kk}.$$

Конечно, это выражение не обязательно сходится. Мы будем считать, что оно определено только для бесконечных рядов, сходящихся абсолютно.

Для любого оператора  $Y$  в пространстве  $V$  можно записать

$$D(Y) = \begin{pmatrix} D_0(Y) & 0 & 0 & 0 & \dots \\ 0 & D_1(Y) & 0 & 0 & \dots \\ 0 & 0 & D_2(Y) & 0 & \dots \\ 0 & 0 & 0 & D_3(Y) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Полагаем теперь  $D_0(Y) = 0$  и  $D_1(Y) = Y$ . Тогда получаем

$$\text{tr}^\otimes e^{D(Y)} = \sum (1/k!) \text{tr} \exp\{D_k(Y)\},$$

следовательно,

$$F(\beta_1, \beta_2) = \text{tr}^\otimes e^{-D(\beta_1 \mathbb{I} + \beta_2 X)}.$$

Построим аналогичную конструкцию для случая Ферми–Дирака. Вместо  $T(V)$  (тензорной алгебры) рассмотрим  $\Lambda(V)$  (внешнюю алгебру). Например, давайте посмотрим на  $\Lambda^2(V) = V \wedge V$  — пространство внешних 2-векторов над  $V$ . В этом случае определим  $D_2(Y)$  как линейное преобразование  $\Lambda^2(V)$  согласно формуле

$$D_2(Y)(u \wedge v) = Yu \wedge v + u \wedge Yv.$$

Это та же формула, что и раньше, только знак  $\otimes$  заменен на  $\wedge$ . Как и раньше, мы имеем

$$e^Y \wedge e^Y = e^{D_2(Y)} \quad \text{на } \Lambda^2(V).$$

Точно так же определяем

$$D_3(Y) = Y \wedge \mathbb{I} \wedge \mathbb{I} + \mathbb{I} \wedge Y \wedge \mathbb{I} + \mathbb{I} \wedge \mathbb{I} \wedge Y \quad \text{на } \Lambda^3(V)$$

и т. д. Так же, как и раньше, мы определяем  $D(Y)$  как прямую сумму всех  $D_i(Y)$  — матрицу на  $\Lambda(V)$ , диагональные компоненты которой суть  $D_i(Y)$ . Если  $V$  — конечномерное векторное пространство, то теперь это конечная матрица.

Предположим, что  $Z$  — диагонализуемая матрица, собственные значения которой равны  $z_1, \dots, z_d$ , а соответствующие собственные векторы —  $e_1, \dots, e_d$ . Тогда мы знаем, что векторы  $e_i \wedge e_j$ , где  $i < j$ , образуют базис пространства  $\Lambda^2(V)$  и

$$(Z \wedge Z)(e_i \wedge e_j) = Ze_i \wedge Ze_j = z_i z_j e_i \wedge e_j,$$

так что  $z_i z_j$  для  $i < j$  — собственные значения  $Z \wedge Z$  на  $\Lambda^2(V)$ . Аналогично,  $z_i z_j z_k$  для  $i < j < k$  — собственные значения  $Z \wedge Z \wedge Z$  на  $\Lambda^3(V)$  и т. д. Поэтому, если  $Z^\wedge$  обозначает прямую сумму рассмотренных операторов, т. е. оператор на пространстве  $\Lambda(V) = \mathbb{R} \oplus V \oplus (V \wedge V) \oplus \dots$ , где

$$Z^\wedge = \begin{pmatrix} \mathbb{I} & 0 & 0 & 0 & \dots \\ 0 & Z & 0 & 0 & \dots \\ 0 & 0 & Z \wedge Z & 0 & \dots \\ 0 & 0 & 0 & Z \wedge Z \wedge Z & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

то

$$\begin{aligned}\mathrm{tr} Z^\wedge &= 1 + \sum z_i + \sum z_i z_j + \sum z_i z_j z_k + \dots \\ &= \prod (1 + z_i) = \mathrm{Det}(\mathbb{I} + Z).\end{aligned}$$

Пусть  $Z = e^Y$ . Воспользуемся формулой

$$(e^Y)^\wedge = e^{D(Y)},$$

которая является прямой суммой всех формул

$$\begin{aligned}1 &= e^0 = e^{D_0(Y)} && \text{на } \mathbb{R}, \\ e^Y &= e^{D_1(Y)} && \text{на } V, \\ e^Y \wedge e^Y &= e^{D_2(Y)} && \text{на } \Lambda^2(V) \text{ и т. д.}\end{aligned}$$

Тогда получаем

$$\mathrm{tr}^\wedge e^{D(Y)} = \mathrm{Det}(\mathbb{I} + e^Y),$$

где в левой части уравнения мы использовали оператор  $\mathrm{tr}^\wedge$ , чтобы отразить то, что след вычисляется в пространстве  $\Lambda(V)$ . Если взять  $Y = e^{-(\beta_1 \mathbb{I} + \beta_2 X)}$ , то получим

$$\mathrm{Det}(\mathbb{I} + e^{-(\beta_1 \mathbb{I} + \beta_2 X)}) = \mathrm{tr}^\wedge e^{-D(\beta_1 \mathbb{I} + \beta_2 X)}.$$

Давайте еще раз обсудим странные математические манипуляции, которыми мы здесь занимались. Начнем с простой системы, которая может находиться в  $d$  состояниях с «энергиями»  $\varepsilon_1, \dots, \varepsilon_d$ . Соответствующая функция разбиения имеет вид

$$Z(\beta) = e^{-\beta \varepsilon_1} + \dots + e^{-\beta \varepsilon_d}.$$

Если  $X$  — диагональная матрица с собственными значениями  $\varepsilon_1, \dots, \varepsilon_d$ , то эту формулу можно переписать в виде

$$Z(\beta) = \mathrm{tr} e^{-\beta X}. \quad (22.49)$$

Пусть  $V$  обозначает векторное пространство  $\mathbb{R}^d$ , и след есть просто след матрицы. Мы хотим рассмотреть более сложную систему, в которой на каждом «энергетическом уровне» может находиться различное число частиц — «оккупантов». У нас есть две

равновесных наблюдаемых — полное число оккупантов и полная энергия. Теперь функция разбиения зависит от двух переменных, и переход от простой системы к этой сложной состоит в следующем:

Заменяем  $V$  на  $T(V)$ , и тогда

$$Z(\beta_1, \beta_2) = \text{tr}^{\otimes} e^{-D(\beta_1 \mathbb{I} + \beta_2 X)} \quad (\text{статистика Больцмана–Пуассона});$$

заменяем  $V$  на  $\Lambda(V)$ , и тогда

$$Z(\beta_1, \beta_2) = \text{tr}^{\wedge} e^{-D(\beta_1 \mathbb{I} + \beta_2 X)} \quad (\text{статистика Ферми–Дирака}).$$

Таким образом, различие между случаями Больцмана–Пуассона и Ферми–Дирака состоит в том, что используется внешняя алгебра  $\Lambda(V)$  вместо полной тензорной алгебры  $T(V)$ . В случае Бозе–Эйнштейна мы заменяем внешнюю алгебру симметрической алгеброй  $S(V)$  (алгеброй всех полиномов на пространстве  $V^*$ ). Используются те же соображения, что и в случае Ферми–Дирака, но теперь у нас будет

$$\text{tr}^{\text{sym}} Z = \text{Det}(\mathbb{I} - Z)^{-1}.$$

Все полученные результаты собраны в таблице 22.5.

Таблица 22.5

Система	Векторное пространство	Функция разбиения
Простая	$V$	$\exp(-\beta X)$
Больцмана–Пуассона	$T(V)$	$\text{tr}^{\otimes} e^{-D(\beta_1 \mathbb{I} + \beta_2 X)}$ $= \text{Det}(e^{-(\beta_1 \mathbb{I} + \beta_2 X)})$
Ферми–Дирака	$\Lambda(V)$	$\text{tr}^{\wedge} e^{-D(\beta_1 \mathbb{I} + \beta_2 X)}$ $= \text{Det}(\mathbb{I} + e^{-(\beta_1 \mathbb{I} + \beta_2 X)})$
Бозе–Эйнштейна	$S(V)$	$\text{tr}^{\text{sym}} e^{-D(\beta_1 \mathbb{I} + \beta_2 X)}$ $= \text{Det}(\mathbb{I} - e^{-(\beta_1 \mathbb{I} + \beta_2 X)})^{-1}$

### 22.13. Квантовые состояния и квантовая логика

Давайте еще раз рассмотрим построения предыдущего параграфа в простейшем случае. Пусть система будет конечным множеством, в котором  $\mu$  приписывает одинаковые веса всем точкам, т. е. систему  $A$  из параграфа 22.6. В такой системе «интегрирование» есть просто суммирование. Наблюдаемое — это функция, определенная на этом конечном множестве, а состояние — неотрицательная функция, сумма значений которой равна 1. Но в этот раз мы будем записывать функции как диагональные матрицы, так что суммирование становится вычислением следов этих матриц. Итак, состояние  $\rho$  записывается в виде

$$\rho = \begin{pmatrix} \rho(1) & 0 & 0 & \dots \\ 0 & \rho(2) & 0 & \dots \\ 0 & 0 & \rho(3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

и наблюдаемое  $J$  аналогично

$$J = \begin{pmatrix} J(1) & 0 & 0 & \dots \\ 0 & J(2) & 0 & \dots \\ 0 & 0 & J(3) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

так что

$$E(J; \rho) = \text{tr } J\rho. \quad (22.50)$$

Функция разбиения задается в форме

$$F(\beta) = \text{tr } e^{-\beta X}, \quad (22.51)$$

где

$$X = \begin{pmatrix} \varepsilon_1 & 0 & 0 & \dots \\ 0 & \varepsilon_2 & 0 & \dots \\ 0 & 0 & \varepsilon_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Тогда равновесные состояния имеют вид

$$\rho_\beta = F(\beta)^{-1} \begin{pmatrix} e^{-\beta \varepsilon_1} & 0 & 0 & \dots \\ 0 & e^{-\beta \varepsilon_2} & 0 & \dots \\ 0 & 0 & e^{-\beta \varepsilon_3} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

или

$$\rho_\beta = F(\beta)^{-1} e^{-\beta X}. \quad (22.52)$$

Переход от простой системы к случаю с числами заполнения состоял в замене конечномерного векторного пространства  $V$  на одно из более широких пространств  $T(V)$ ,  $\Lambda(V)$  или  $S(V)$ , в зависимости от того, какая «статистика» исследуется. При этом все рассматриваемые матрицы были диагональными, поэтому использование матриц казалось весьма искусственным. Здесь же давайте рассмотрим эту роль матриц более серьезно.

Пусть  $V$  — векторное пространство, где определено скалярное произведение. (В силу важных физических причин мы будем рассматривать комплексное векторное пространство  $V$ , где  $(\cdot, \cdot)$  — положительно определенное скалярное произведение, так что  $V$  является унитарным пространством. В реальных задачах часто приходится рассматривать бесконечномерное пространство  $V$ . Но для понимания ключевых идей нет необходимости включать сюда еще и технические проблемы теории (унитарных) гильбертовых пространств, поэтому все понятия мы продемонстрируем на конечномерном пространстве.) Вспомним, что оператор  $A$  в пространстве  $V$  называется *самосопряженным*, если

$$(Au, v) = (u, Av) \quad \text{для всех } u, v \in V.$$

Рассуждения, приведенные в параграфе 4.3 (выполненные для комплексного векторного пространства), показывают, что любой самосопряженный оператор можно привести к диагональному виду, т.е. что существует такой ортонормированный базис  $e_1, \dots, e_d$ , что

$$Ae_i = \lambda_i e_i, \quad i = 1, \dots, d,$$

где  $\lambda_i$  — вещественные числа. (В случае бесконечномерного пространства соответствующая теорема, модифицированная определенным образом, называется спектральной теоремой.) Оператор

$A$  называется неотрицательным, или  $A \geq 0$ , если все  $\lambda_i \geq 0$ . Другими словами,  $A$  неотрицателен, если и только если

$$(Au, u) \geq 0 \quad \text{для всех } u \in V,$$

что можно легко проверить, записав  $u$  в виде линейной комбинации всех  $e_i$ . Теперь можно ввести следующие определения.

*Квантовая система* — унитарное (гильбертово) пространство  $V$ .

*Квантовое (статистическое) состояние* — неотрицательный самосопряженный оператор  $\rho$ , для которого

$$\text{tr } \rho = 1.$$

*Энтропия* состояния  $\rho$  задается формулой

$$\text{Ent}(\rho) = -\text{tr } \rho \log \rho.$$

*Квантовое наблюдаемое* — самосопряженный оператор.

*Ожидаемое значение* наблюдаемого  $A$  в состоянии  $\rho$  задается формулой

$$E(A; \rho) = \text{tr } A\rho.$$

Пусть  $J_1, \dots, J_k$  —  $k$  коммутирующих наблюдаемых, т.е. предположим, что

$$J_i J_j = J_j J_i \quad \text{для всех } i \text{ и } j.$$

Пусть  $\beta = (\beta_1, \dots, \beta_k)$ , где  $\beta_i$  — вещественные числа, и пусть  $\beta \mathbf{J}$  — наблюдаемое

$$\beta \mathbf{J} = \beta_1 J_1 + \dots + \beta_k J_k.$$

Тогда соответствующая *функция разбиения* определяется выражением

$$F(\beta) = \text{tr } e^{-\beta \mathbf{J}}, \quad (22.53)$$

и *равновесные состояния*  $\rho_\beta$  определяются выражением

$$\rho_\beta = F(\beta)^{-1} e^{-\beta \mathbf{J}}.$$

Итак, ожидаемое значение наблюдаемого  $A$  в равновесном состоянии, соответствующем  $\beta$ , имеет вид

$$E(A; \rho_\beta) = F(\beta)^{-1} \operatorname{tr} A e^{-\beta J}. \quad (22.54)$$

Уравнения (22.53) и (22.54) являются теоретическим базисом квантовой статистической механики. Приведем цитату с первой страницы книги Фейнмана «Статистическая механика»: «Это фундаментальное соотношение является вершиной статистической механики; остальное ее содержание есть либо спуск с вершины, когда основные принципы применяются к частным вопросам, либо восхождение на нее, когда выводятся основные соотношения...»

В случае, когда  $V$  — бесконечномерное пространство, след оператора становится бесконечным рядом, который может и не сходиться. Поэтому не все операторы имеют след. Соответственно, функция  $F(\beta)$  будет сходиться только в ограниченной области  $\beta$ . Сейчас мы оставим эти проблемы в стороне и займемся моделью конечномерного пространства  $V$ .

Основное различие между квантовой и классической системами состоит в том, что в классической системе наблюдаемая является функцией (на конечном множестве), которую будет диагональной матрицей, а для квантовой системы мы допускаем *все* (самосопряженные) матрицы. На первый взгляд это может показаться только технической модификацией (которая подтверждается экспериментами на протяжении шестидесяти лет). На самом деле это наиболее глубокая революция в истории науки, потому что она изменяет элементарные правила логики. Давайте это объясним.

Опять начнем с классической системы. Наблюдаемое типа «да или нет» является функцией  $f$ , принимающей значения 0 или 1. Это соответствует любому эксперименту, в котором ответ задается неким индикатором (например, выключателем или светом), включен он или нет. Когда мы говорим, что функция  $f$  принимает значения 0 или 1, — это то же самое, что

$$f(m)^2 = f(m)$$

для всех  $m$ , или короче

$$f^2 = f. \quad (22.55)$$

Каждая такая функция соответствует подмножеству  $B \subset M$  множества, где  $f = 1$ . Поэтому для любого заданного подмножества  $B$  мы имеем функцию<sup>22</sup>  $f_B$ , для которой

$$f_B(m) = 1, \text{ если } m \in B, \text{ и } f_B = 0, \text{ если } m \notin B.$$

Если  $f$  и  $g$  — две такие функции «да или нет», то в классической системе

$$fg \text{ — тоже функция «да или нет».} \quad (22.56)$$

Действительно, если  $C$  и  $D$  — подмножества, то

$$f_C f_D = f_{C \cap D}, \quad (22.57)$$

поскольку  $f_C(m)f_D(m) = 1$ , если и только если для обеих функций мы имеем  $f_C(m) = 1$  и  $f_D(m) = 1$ . Такое умножение функций соответствует пересечению подмножеств, т. е. логической конъюнкции — точка принадлежит подмножеству  $C \cap D$ , если и только если она принадлежит и  $C$ , и  $D$ . Например, подмножества  $C$  и  $D$  не пересекаются, т. е.

$$C \cap D = \emptyset, \text{ если и только если } f_C f_D = 0.$$

Если  $C$  и  $D$  не пересекаются, то

$$f_C + f_D = f_{C \cup D}, \quad (22.58)$$

поскольку обе части равенства равны 1, когда аргумент принадлежит  $C$  или  $D$  (и не существует точек, принадлежащих обоим подмножествам). Дистрибутивный закон *умножения*

$$f_B(f_C + f_D) = f_B f_C + f_B f_D. \quad (22.59)$$

— это перевод дистрибутивного закона *теории множеств*

$$B \cap (C \cup D) = (B \cap C) \cup (B \cap D), \quad (22.60)$$

что, в свою очередь, является записью дистрибутивного закона логики. Пусть  $\mathbf{B}$  обозначает утверждение  $m \in B$  и т. д., тогда  $\cap$

<sup>22</sup>называемую характеристической функцией, или индикатором множества  $B$ . — Прим. ред.

обозначает конъюнкцию: формула  $\mathbf{B} \cap \mathbf{C}$  утверждает, что и  $\mathbf{B}$ , и  $\mathbf{C}$  верны. Аналогично,  $\cup$  обозначает (неисключающую) дизъюнкцию, т. е. формула  $\mathbf{C} \cup \mathbf{D}$  означает, что верно либо  $\mathbf{C}$ , либо  $\mathbf{D}$ , либо и то, и другое. (Далее мы будем употреблять «или» в неисключающем смысле, не уточняя «или оба».) Итак, дистрибутивный закон просто отражает элементарный логический принцип:

Если мы говорим, что верны  $\mathbf{B}$  и  $(\mathbf{C}$  или  $\mathbf{D})$ , то это то же самое, что сказать: верно  $(\mathbf{B}$  и  $\mathbf{C})$  или  $(\mathbf{B}$  и  $\mathbf{D})$ .

Рассмотрим соответствующую ситуацию в квантовой системе. Наши наблюдаемые уже являются не функциями, а самосопряженными операторами. Мы уже не можем говорить о «значениях» наблюдаемых, а только о их собственных значениях. Поэтому назовем самосопряженный оператор  $\pi$  наблюдаемым типа «да или нет», если его собственные значения равны 0 или 1. Это произойдет, если и только если

$$\pi^2 = \pi, \quad (22.61)$$

что аналогично (22.55). Условие (22.61) означает, что  $\pi$  является проекцией. Более того, если  $\mathbf{e}_1, \dots, \mathbf{e}_d$  — собственные векторы  $\pi$ , то  $\pi$  — *ортогональная проекция* на подпространство, натянутое на собственные векторы, соответствующие собственному значению 1. Таким образом, каждое наблюдаемое «да или нет» соответствует подпространству (вместо подмножества), и можно писать  $\pi_A$ , обозначая ортогональную проекцию на подпространство  $A$ . Заметим, что теперь нулевой оператор соответствует ортогональной проекции на нулевое подпространство  $\{0\}$  и

$$\pi_C \pi_D = 0, \quad \text{если и только если } C \cap D = \{0\}, \quad (22.62)$$

поэтому нулевое подпространство  $\{0\}$  играет роль, аналогичную пустому множеству. Заметим, что если  $\pi_C \pi_D = 0$ , то, перейдя к сопряженным операторам, имеем

$$0 = (\pi_C \pi_D)^* = \pi_D^* \pi_C^* = \pi_D \pi_C,$$

откуда получаем, что  $\pi_D \pi_C = 0$  и

$$(\pi_C + \pi_D)^2 = \pi_C^2 + \pi_D^2 = \pi_C + \pi_D.$$

Таким образом,  $\pi_C + \pi_D$  является наблюдаемым «да или нет». Очевидно, что это ортогональная проекция на прямую сумму  $C \oplus D$  ортогональных подпространств  $C$  и  $D$ . Итак, условие

$$\pi_C + \pi_D = \pi_{C \oplus D}, \quad \text{если} \quad \pi_C \pi_D = 0 \quad (22.63)$$

является полной аналогией (22.58), где  $\oplus$  заменяет  $\cup$ . Однако (22.56) полностью нарушается. Вообще говоря, для произвольных подпространств  $C$  и  $D$  равенство  $\pi_C \pi_D = \pi_D \pi_C$  не выполняется, и поэтому оператор  $\pi_C \pi_D$  не является самосопряженным, и, следовательно, он не является наблюдаемым. Например, если  $C$  и  $D$  — две неортогональные прямые на плоскости, то  $\pi_C \pi_D \neq 0$ . Но образом  $\pi_C \pi_D$  является  $C$ , а образом  $\pi_D \pi_C$  —  $D$ . Поэтому  $\pi_C \pi_D \neq \pi_D \pi_C$ . Проблема возникает из некоммутативности умножения матриц. Заметим, что аналогом дистрибутивного закона теории множеств являлось бы утверждение

$$B \cap (C \oplus D) = (B \cap C) \oplus (B \cap D),$$

которое верно далеко не всегда. Действительно, если в качестве  $C$  взять ось  $x$ , а в качестве  $D$  — ось  $y$  на плоскости, то  $C \oplus D$  — вся плоскость, и поэтому  $B \cap [(C \oplus D)] = B$  для любой прямой  $B$  на этой плоскости. Но если  $B$  — любая прямая, отличная от осей  $x$  и  $y$ , то  $B \cap C = B \cap D = \{0\}$ , поэтому  $(B \cap C) \oplus (B \cap D) = \{0\} \oplus \{0\} = \{0\}$ , и мы уже не имеем равенства двух частей.

Итак, в квантовой логике не выполняется дистрибутивный закон. Как уже упоминалось, на протяжении многих лет экспериментально подтверждается справедливость квантовой механики. Поэтому эксперимент показывает, что, работая с квантовыми наблюдаемыми, необходимо отказаться от одного из наиболее драгоценных принципов логики.

## Резюме

### А. Формулировка Каратеодори второго закона термодинамики

Вы должны уметь объяснить, при каких обстоятельствах существуют точки в окрестности точки  $P$ , которые нельзя соединить с ней нулевой кривой 1-формы  $\alpha$ .

Вы должны уметь формулировать и объяснять формулу Каратеодори для второго закона термодинамики.

Вы должны уметь объяснять понятия абсолютной температуры и энтропии в терминах дифференциальных форм и контурных интегралов, уметь описывать, как эти величины можно вычислять по экспериментальным данным.

## Задачи

- 22.1. Пусть  $\alpha = (y^3 + y) dx + (xy^2 + x) dy$ . Дайте характеристику множества точек в пространстве  $\mathbb{R}^2$ , которые можно соединить с точкой  $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$  с помощью нулевой кривой.
- 22.2. Пусть  $\alpha = y dx + x dz$ . Найдите нулевую кривую формы  $\alpha$ , которая соединяет начало координат с точкой  $\begin{pmatrix} 1 \\ b \\ c \end{pmatrix}$ , где  $b > 0$  и  $c > 0$ .  
Кривая не должна проходить через какие-либо точки, у которых  $x$ ,  $y$  или  $z$  равны нулю, за исключением начала координат.
- 22.3. Пусть  $\alpha = x dy$ . Покажите, что любые две точки в пространстве  $\mathbb{R}^3$  можно соединить нулевой кривой формы  $\alpha$ , несмотря на то, что  $\alpha \wedge d\alpha = 0$ . Покажите, что для формы  $\beta = (1 + x^2) dy$  две точки с различными координатами  $y$  не могут быть соединены нулевой кривой.
- 22.4. Форму  $\alpha = 2ye^x dx$  запишите в виде  $\alpha = f dg$  и в виде  $\alpha = F d\Omega$ , где  $F$  не пропорционально  $f$ .
- 22.5. Рассмотрите один моль одноатомного газа, для которого  $pV = RT$  и  $U = \frac{3}{2}pV$ . Предположим, что начальные параметры газа:  $p = 32$ ,  $V = 1$ , но потом он расширяется до параметров:  $p = 1$ ,  $V = 8$ . Вычислите работу, совершенную газом, и количество тепла, поглощенного им, в следующих процессах.
- (а) Сначала газ охлаждается при постоянном объеме, и его давление  $p$  понижается до 1; затем объем  $V$  увеличивается до 8 при постоянном давлении.
  - (б) Сначала газ изотермически расширяется до объема  $V = 8$ ; потом при фиксированном объеме давление снижается до 1.
  - (в) Происходит процесс, в ходе которого  $pV^{5/3} = 32$ .

- 22.6. Для одноатомного идеального газа из упражнения 22.5 напишите форму для тепла в виде  $\alpha = T dS$ . Определите энтропию  $S$  (с точностью до аддитивной константы) как функцию  $p$  и  $V$ . Проверьте, что  $dT \wedge dS = dp \wedge dV$ .
- 22.7. Сосуд объема  $V$  заполнен электромагнитным излучением при температуре  $T$ , давлении  $p = \frac{1}{3}CT^4$  (где  $C$  — постоянная). Его энергия  $U = 3pV$ .
- Записав формулу тепла  $dU + p dV$  в виде  $\alpha = T dS$ , найдите выражение для  $S$  через  $p$  и  $V$ .
  - Предположим, что начальный объем сосуда равен  $V_0$ , а температура  $T_0$ . Объем увеличивается адиабатически до величины  $64V_0$ . Определите конечную температуру и работу, выполненную при расширении.
- 22.8. (a) В машине Карно тепло поглощается при температуре  $4T_0$ , а отдается при температуре  $T_0$ . Какова максимальная доля поглощенного тепла, которая может быть преобразована в полезную работу?
- (b) Предположим, что эта же машина работает как холодильник. Если тепло  $Q$  поглощается при более низкой температуре  $T_0$ , то какую работу необходимо затратить для этого холодильника?
- 22.9. Система из  $N$  протонов, каждый из которых имеет магнитный дипольный момент  $\mu$ , помещена во внешнее магнитное поле  $H$ . Разумно полагать, что  $H$  — единственная конфигурационная переменная. Магнитный момент системы протонов  $M$  и соответствующая энергия  $U$  при больших температурах  $T$  хорошо аппроксимируются формулами

$$M = \frac{N\mu^2 H}{kT}, \quad U = -\frac{N(\mu H)^2}{kT}.$$

Работа, совершенная системой, определяется выражением  $W = \int M dH$ . Поглощенное тепло равняется  $\Delta U + W$ . Получите выражение для энтропии  $S$  через  $T$  и  $H$ .

- 22.10. Рассмотрим систему, у которой объем  $V$  — единственная конфигурационная переменная. Докажите следующие соотношения.

$$(a) \left( \frac{\partial T}{\partial V} \right)_{\text{адиабат.}} = \left( \frac{\partial p}{\partial S} \right)_{V=\text{const}};$$

$$(b) \left( \frac{\partial S}{\partial V} \right)_{\text{изотерм.}} = \left( \frac{\partial p}{\partial T} \right)_{V=\text{const}}.$$

(Указание: используйте равенство  $dT \wedge dS = dp \wedge dV$ .)

22.11. «Свободная энергия Гельмгольца»  $F$  для системы связана с внутренней энергией  $U$  соотношением  $F = U - TS$ .

- (a) Покажите, что если система взаимодействует с окружением изотермически, то свободная энергия  $F$  равна работе, совершенной над системой.
- (b) Предположим, что  $F$  записана как функция  $V$  и  $T$ . Вычислите ее частные производные по этим переменным.
- (c) Найдите функцию  $F(V, T)$  для идеального одноатомного газа из упражнения 22.5.

22.12. Пусть  $K = \frac{dV \wedge dT}{V dT \wedge dp}$  обозначает изотермическую сжимаемость, и  $\alpha = \frac{dV \wedge dp}{V dt \wedge dp}$  обозначает коэффициент теплового расширения при постоянном давлении. Докажите соотношение

$$C_p = C_v + VT \frac{\alpha^2}{K}.$$

22.13. Для одноатомного газа из упражнения 22.5 получите явное выражение для функции  $Z(\beta, V)$ . Вычислите ее частные производные и покажите, что они равны  $-U$  и  $\beta p$  соответственно.

22.14. Предположим, что энтропия системы  $S$  зависит только от ее внутренней энергии  $U$  согласно формуле  $S = N(U/U_0)^{1/2}$ . Получите выражения для теплоемкостей этой системы в зависимости от температуры.

22.15. Рассмотрим систему трех частиц. Две из них — неразличимые бозоны, а третья — какая-то другая частица. Каждая частица может находиться в одном из трех состояний, причем все они имеют одинаковую энергию.

- (a) Какова вероятность того, что все три частицы находятся в одном состоянии?
- (b) Какова вероятность того, все три частицы находятся в разных состояниях?

- 22.16. Для классического магнитного диполя с моментом  $\mu$  в магнитном поле  $\mathbf{H}$  энергия равна  $U = -\mu H \cos \theta$ , где  $\theta$  — угол между этими векторами. Вычислите функцию разбиения этой системы, интегрируя по всевозможным направлениям.
- 22.17. Рассмотрим систему, у которой есть три состояния с энергиями  $-\varepsilon$ ,  $0$  и  $\varepsilon$ . Предположим, что ожидаемое значение энергии равно  $-\frac{3}{7}\varepsilon$ . Определите вероятности этих трех состояний.
- 22.18. Предположим, что в системе есть одно состояние с энергией  $-\varepsilon$ , два состояния с энергией  $0$  и одно состояние с энергией  $\varepsilon$ . Если ожидаемое значение энергии равно  $-\frac{1}{5}\varepsilon$ , то каковы вероятности этих состояний?
- 22.19. Пусть в системе есть только два состояния: одно с энергией  $-\varepsilon$ , другое с энергией  $\varepsilon$ . Найдите функцию разбиения этой системы и используйте ее для определения энергии  $U$  и энтропии  $S$  в зависимости от температуры  $T$ .

## Приложение: доказательство теоремы о нормальной форме

### 1. Ранг 1-формы

Сначала определим ранг 2-формы. Пусть  $\Omega$  — 2-форма. Для каждого положительного целого числа  $k$  можно рассмотреть  $2k$ -форму

$$\Omega^k = \Omega \wedge \Omega \wedge \cdots \wedge \Omega \quad (k \text{ раз}).$$

Для каждой точки  $x$  пусть  $k$  будет таким неотрицательным целым числом, что  $(\Omega^k)_x \neq 0$ , но  $(\Omega^{k+1})_x = 0$ . Вообще говоря, целое число  $k$  может меняться от точки к точке. Ранг формы  $\Omega$  в точке  $x$  равен  $2k$ . Нам, главным образом, интересны формы постоянно-го ранга, когда  $2k$  не зависит от  $x$ . Если  $2k = 0$ , то  $\Omega = 0$ . Тогда форма  $\Omega = dx \wedge dy$  имеет ранг 2, поскольку  $\Omega \neq 0$ , в то время как  $\Omega^2 = 0$ . Форма

$$\Omega = dx \wedge dy + du \wedge dv$$

имеет ранг 4, поскольку

$$\Omega^2 = 2dx \wedge dy \wedge du \wedge dv \neq 0,$$

но  $\Omega^3 = 0$ . (Здесь мы предполагаем, что  $x, y, u, v$  и т.д. — независимые функции.)

А теперь определим ранг 1-формы. Это немного сложнее, поэтому воспользуемся индукцией. Если  $\omega = 0$  в точке  $x$ , то мы говорим, что ранг  $\omega$  в  $x$  равен 0. Если  $\omega_x \neq 0$ , но  $d\omega_x = 0$ , то мы говорим, что ранг  $\omega$  в точке  $x$  равен 1. Если  $\omega_x \neq 0$  и  $d\omega_x \neq 0$ , но  $\omega_x \wedge d\omega_x = 0$ , то ранг  $\omega$  в точке  $x$  равен 2. В общем случае мы ищем такое положительное целое число  $k$ , что

$$\omega_x \wedge (d\omega)_x^{k-1} \neq 0,$$

но

$$\omega_x \wedge (d\omega)_x^k = 0.$$

В этой ситуации существует две возможности.

Если  $(d\omega)_x^k \neq 0$ , то ранг  $\omega$  в точке  $x$  равен  $2k$ .

Если  $(d\omega)_x^k = 0$ , то ранг  $\omega$  в точке  $x$  равен  $2k - 1$ .

Конечно, таким образом ранг  $\omega$  определен в каждой точке  $x$ . Нам интересны только те случаи, когда ранг постоянный. Поэтому на самом деле мы имеем в виду, что: ранг  $\omega$  равен  $2k + 1$ , если  $\omega \wedge (d\omega)^k$  нигде не обращается в нуль, а  $(d\omega)^{k+1}$  тождественно равна нулю; ранг  $\omega$  равен  $2k$ , если  $(d\omega)^k$  нигде не обращается в нуль, но везде выполняется тождество  $\omega \wedge (d\omega)^k \equiv 0$ . Если к этому тождеству применить оператор  $d$ , то  $(d\omega)^{k+1}$  тождественно равна нулю. Итак, если 1-форма  $\omega$  имеет ранг  $2k$  или  $2k + 1$ , то 2-форма  $d\omega$  имеет ранг  $2k$ .

**Примеры.** Нулевая форма имеет ранг 0;  $dx_0$  имеет ранг 1, потому что  $d(dx_0) = 0$ , откуда следует  $k = 0$ ;  $x_1 dx_2$  имеет ранг 2 в области  $x_1 \neq 1$ , потому что форма  $d(x_1 dx_2) = dx_1 \wedge dx_2$  имеет  $k = 1$  и  $x_1 dx_2 \wedge dx_1 \wedge dx_2 = 0$ ;  $dx_0 + x_1 dx_2$  имеет ранг 3, потому что  $d(dx_0 + x_1 dx_2) = dx_1 \wedge dx_2$  имеет  $k = 1$ , в то время как

$$(dx_0 + x_1 dx_2) \wedge dx_1 \wedge dx_2 = dx_0 \wedge dx_1 \wedge dx_2 \neq 0.$$

В общем случае очевидно, что

$$dx_0 + x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k}$$

имеет ранг  $2k + 1$ , а

$$x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k}$$

имеет ранг  $2k$  в области, где не все  $x_{2k-1}$  равны нулю.

В этом приложении мы хотим доказать, что это фактически единственные примеры. А именно, если  $\omega$  — линейная дифференциальная форма постоянного ранга  $2k$ , то всегда можно найти такие локальные координаты  $x_1, x_2, \dots, x_{2k}$ , что

$$\omega = x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k},$$

а если ранг  $\omega$  равен  $2k + 1$ , то всегда можно найти такие координаты  $x_0, x_1, x_2, \dots, x_{2k}$ , что

$$\omega = dx_0 + x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k}.$$

Это и есть тот факт, по модулю которого мы доказали теорему Каратеодори.

## 2. Сведение к теореме Дарбу

Как можно заключить из предыдущей теоремы, всегда можно ввести такие координаты, что

$$d\omega = dx_1 \wedge dx_2 + dx_3 \wedge dx_4 + \cdots + dx_{2k-1} \wedge dx_{2k}.$$

Теорема Дарбу (ее мы докажем позднее в этом приложении) утверждает, что если  $\Omega$  — любая замкнутая 2-форма постоянного ранга  $2k$ , то всегда можно ввести такие координаты, что

$$\Omega = dx_1 \wedge dx_2 + dx_3 \wedge dx_4 + \cdots + dx_{2k-1} \wedge dx_{2k}.$$

Давайте допустим, что эта теорема доказана, и с ее помощью получим теорему о нормальной форме. Если ранг  $\omega$  нечетный, то это совсем просто. Действительно,  $d\omega$  — замкнутая 2-форма ранга  $2k$ , поэтому можно записать ее в виде

$$\begin{aligned} d\omega &= dx_1 \wedge dx_2 + dx_3 \wedge dx_4 + \cdots + dx_{2k-1} \wedge dx_{2k} \\ &= d(x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k}), \end{aligned}$$

так что

$$d(\omega - (x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k})) = 0.$$

Откуда, в силу леммы Пуанкаре, имеем

$$\omega - (x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k}) = dx_0,$$

где  $x_0$  — некоторая функция. Тогда

$$\omega = dx_0 + x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k}.$$

Но мы предположили, что

$$\omega \wedge (d\omega)^k = dx_0 \wedge dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{2k-1} \wedge dx_{2k}$$

не равняется нулю. Следовательно,  $dx_0$  не зависит от остальных  $dx_i$ , и можно использовать  $x_0, x_1, x_2, \dots, x_{2k}$  в качестве координат. Итак, с учетом теоремы Дарбу о замкнутых 2-формах, это доказывает теорему о нормальной форме для линейных дифференциальных форм нечетного ранга.

Для форм четного ранга доказательство несколько сложнее. Мы докажем, что если форма  $\omega$  имеет постоянный ранг  $2k$ , то можно найти такую положительную функцию  $f$ , что

$$\omega = f\sigma,$$

где форма  $\sigma$  имеет ранг  $2k - 1$ . Если теперь применить теорему о нормальных формах нечетного ранга, то можно написать

$$\sigma = dx_0 + w_1 dx_2 + \cdots + w_{2k-3} dx_{2k-2},$$

где  $x_0, w_1, x_2, w_3$  и т. д. — координаты. Тогда

$$\begin{aligned} \omega &= f dx_0 + f w_1 dx_2 + \cdots \\ &= x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k}, \end{aligned}$$

если положить

$$\begin{aligned} x_1 &= f w_1, & x_3 &= f w_3, & \dots, & x_{2k-3} &= f w_{2k-3}, \\ x_{2k-1} &= f, & x_{2k} &= x_0. \end{aligned}$$

В силу того, что  $(d\omega)^k \neq 0$ , внешнее произведение всех  $dx_i$  не равно нулю, так что  $dx_i$  линейно независимы во всех точках, и  $x_i$

образуют часть системы координат. Следовательно, мы должны доказать, что  $\omega = f\sigma$ , где  $\sigma$  имеет ранг  $2k-1$  и  $f$  положительна. С этой целью воспользуемся теоремой Дарбу (пока не доказанной) и напишем

$$d\omega = dy_1 \wedge dy_2 + dy_3 \wedge dy_4 + \cdots + dy_{2k-1} \wedge dy_{2k},$$

где  $y_1, \dots, y_n$  — подходящие локальные координаты. Можно использовать эти координаты и написать

$$\omega = a_1 dy_1 + a_2 dy_2 + \cdots + a_n dy_n.$$

Тогда поскольку  $\omega \wedge (d\omega)^k \equiv 0$ , все  $a_i = 0$  для  $i > 2k$ . Из записи формы  $d\omega$  следует, что все  $a_i$  для  $i \leq 2k$  не зависят от переменных  $y_{2k+1}, \dots, y_n$ , поскольку иначе в выражении для  $d\omega$  мы получили бы ненулевые коэффициенты при  $dy_i \wedge dy_j$ , где  $i \leq 2k$  и  $j > 2k$ . Итак, доказывая теорему о нормальных формах четного ранга  $2k$ , можно предположить, что мы работаем в пространстве  $\mathbb{R}^{2k}$ . (Все остальные переменные роли не играют.) Итак,  $\omega \wedge (d\omega)^{k-1}$  — ненулевая форма степени  $2k-1$  в пространстве  $\mathbb{R}^{2k}$ . Следовательно, в каждой точке своей области определения она задает прямую, т. е. одномерное пространство всех решений уравнения

$$i(\xi)[\omega \wedge (d\omega)^{k-1}] = 0.$$

Таким образом,  $\omega \wedge (d\omega)^{k-1}$  определяет «элемент поля направлений», другими словами, систему обыкновенных дифференциальных уравнений. Давайте выберем  $(2k-1)$ -мерную поверхность и будем рассматривать точки этой поверхности как начальные условия. Тогда каждая точка вблизи гиперповерхности лежит на единственной кривой решения, и поэтому соответствует определенному моменту времени  $t$  (параметр кривой) и определенной точке (начальное условие) на гиперповерхности.

Если на гиперповерхности ввести координаты  $z_1, \dots, z_{2k-1}$ , то  $t, z_1, \dots, z_{2k-1}$  будут локальными координатами в пространстве  $\mathbb{R}^{2k}$ . В этих координатах  $\xi = \partial/\partial t$ . Другими словами, формы

$$\omega \wedge (d\omega)^{k-1} \quad \text{и} \quad dz_1 \wedge dz_2 \wedge \cdots \wedge dz_{2k-1}$$

определяют один и тот же элемент поля направлений. Тогда

$$\omega \wedge (d\omega)^{k-1} = g dz_1 \wedge dz_2 \wedge \cdots \wedge dz_{2k-1},$$

где  $g$  — функция, нигде не обращающаяся в нуль. Можно при необходимости заменить  $z_1$  на  $-z_1$  и считать, что  $g$  — положительная функция. Пусть  $f^k = g$ . Если определить форму  $\sigma$  равенством  $\omega = f\sigma$ , то

$$d\omega = df \wedge \sigma + f d\sigma,$$

откуда

$$\omega \wedge d\omega = f^2 \sigma \wedge d\sigma.$$

Далее, находим

$$\omega \wedge (d\omega)^{k-1} = f^k \sigma \wedge (d\sigma)^{k-1} = g dz_1 \wedge dz_2 \wedge \cdots \wedge dz_{2k-1}.$$

Тогда

$$\sigma \wedge (d\sigma)^{k-1} = dz_1 \wedge dz_2 \wedge \cdots \wedge dz_{2k-1}.$$

Действуя оператором  $d$  на обе части этого уравнения, получаем

$$(d\sigma)^k = 0.$$

Другими словами,  $\sigma$  — форма ранга  $2k - 1$ . Это доказывает теорему о нормальной форме в предположении теоремы Дарбу.

### 3. Доказательство теоремы Дарбу

Для доказательства этой теоремы сделаем ряд редукций. Начнем с замкнутой 2-формы  $\Omega$  ранга  $2k$  в  $n$ -мерном пространстве. Предположим, что  $n > 2k$ . Можно выбрать такое векторное поле  $\xi$ , что  $i(\xi)\Omega \equiv 0$ . Тогда, как и раньше, можно (решая систему обыкновенных дифференциальных уравнений) ввести координаты  $y_1, \dots, y_{n-1}, t$  такие, что  $\xi = \partial/\partial t$ . Поскольку  $i(\xi)\Omega$  тождественно обращается в нуль, форма  $\Omega$  не содержит  $dt$ , если записать ее в этих локальных координатах, т. е.

$$\Omega = \sum a_{ij} dy_i \wedge dy_j.$$

Но тогда уравнение  $d\Omega = 0$  означает, что все  $\partial a_{ij}/\partial t = 0$ , так как иначе в  $d\Omega$  будут несокращаемые члены  $(\partial a_{ij}/\partial t) dt \wedge dy_i \wedge dy_j$ . Таким образом,  $a_{ij}$  зависят только от  $y_1, \dots, y_{n-1}$ . Другими словами, на самом деле  $\Omega$  — 2-форма в  $(n - 1)$ -мерном пространстве

координат  $y_i$ . Если  $n - 1 > 2k$ , то можно повторить процедуру и убрать дополнительные координаты, т. е. при доказательстве теоремы Дарбу можно предположить, что  $2k = n$ .

Утверждается, что это условие означает, что форма  $\Omega$  несингулярна в каждой точке в том смысле, что если

$$i(\xi_p)\Omega_p = 0$$

для касательного вектора  $\xi_p$  в точке  $p$ , то  $\xi_p = 0$ . Действительно, из этого уравнения очевидным образом следует, что

$$i(\xi_p)[\underbrace{\Omega \wedge \Omega \wedge \cdots \wedge \Omega}_{k \text{ раз}}]_p = 0.$$

Но  $[\Omega \wedge \Omega \wedge \cdots \wedge \Omega]_p$  —  $n$ -форма в  $n$ -мерном пространстве, а пространство  $n$ -форм одномерно. Поскольку мы предположили, что  $n$ -форма  $[\Omega \wedge \Omega \wedge \cdots \wedge \Omega]_p \neq 0$ , то  $\xi_p = 0$ . (Это следует на самом деле из того, что в  $n$ -мерном пространстве условие несингулярности  $\Omega$  эквивалентно условию, что эта форма имеет ранг  $n$ .) Сказать, что  $\Omega$  несингулярна, можно и по-другому:  $n \times n$  матрица  $(a_{ij}(y))$ , возникающая в выражении  $\Omega = \sum a_{ij} dy_i \wedge dy_j$ , несингулярна для всех  $y$ .

Доказательство теоремы Дарбу разобьем на две части.

(а) Дифференциально-геометрическая часть, в которой утверждается, что всегда можно провести локальную замену координат так, чтобы в новых координатах матрица  $(a_{ij}(y))$  была постоянной, т. е. все матричные элементы  $a_{ij}$  не зависели от  $y$ .

(б) Алгебраическая часть, в которой утверждается, что можно сделать линейную замену координат так, чтобы постоянная матрица  $(a_{ij})$  имела вид

$$\begin{pmatrix} 0 & -1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & -1 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & -1 \\ & & & & 1 & 0 \end{pmatrix}$$

Это — теорема о кососимметричных билинейных формах в векторном пространстве, которую мы докажем по индукции.

Предположим, что  $V$  — двумерное векторное пространство и  $\Omega$  — невырожденная билинейная форма в этом пространстве. Пусть  $\mathbf{u}$  — произвольный ненулевой вектор в  $V$ . Тогда линейная форма

$$i(\mathbf{u})\Omega = \Omega(\mathbf{u}, \cdot)$$

является ненулевой линейной формой в силу невырожденности  $\Omega$ . Следовательно, можно найти такой вектор  $\mathbf{v}$ , что

$$\Omega(\mathbf{u}, \mathbf{v}) = 1.$$

Если в качестве базиса пространства  $V$  взять  $\mathbf{u}$  и  $\mathbf{v}$ , а  $dy_1$  и  $dy_2$  образуют дуальный базис, то

$$\Omega = dy_1 \wedge dy_2,$$

что мы и хотели получить. Пусть теперь  $V$  —  $n$ -мерное пространство, и  $n$  — четное число  $> 2$ . Возьмем те же  $\mathbf{u}$  и  $\mathbf{v}$ , что и раньше, и пространство  $V_1$ , натянутое на эти вектора. Тогда

$$V = V_1 \oplus V_1^\perp,$$

где  $V_1^\perp$  обозначает ортогональное дополнение по отношению к  $\Omega$ . Ограничение  $\Omega$  до  $V_1^\perp$  невырождено. Поэтому желаемый базис пространства  $V_1^\perp$  можно найти по предположению индукции. Тогда разложение в прямую сумму вместе с результатом для двумерного случая доказывает теорему для пространства  $V$ .

Получим теперь геометрический результат (а). Если матричные элементы  $a_{ij}(y)$  зафиксировать в какой-то точке  $y_0$ , то получим невырожденную 2-форму  $\Omega_0$ , совпадающую с  $W$  в точке  $y_0$ . Таким образом, достаточно доказать следующее.

Пусть  $\Omega_1$  и  $\Omega_0$  — две замкнутые невырожденные 2-формы, которые совпадают, например, в 0. Тогда существует диффеоморфизм  $f$ , определенный в окрестности 0, для которого выполняется равенство

$$f^*\Omega_1 = \Omega_0.$$

Докажем это утверждение, снова решая дифференциальные уравнения. Но здесь потребуется некоторая подготовка. Пусть

$$\Omega_t = (1-t)\Omega_0 + t\Omega_1 = \Omega_0 + t\sigma, \quad \text{где } \sigma = \Omega_1 - \Omega_0.$$

Тогда  $d\sigma = 0$  и  $\sigma$  равно нулю в 0. Поскольку  $\Omega_0$  несингулярна, то и  $\Omega_t(0) = \Omega_0$  несингулярна для всех  $t$ . Следовательно, можно найти окрестность начала координат, где  $\Omega_t$  несингулярна для всех  $0 \leq t \leq 1$ . Переходя к меньшей окрестности (если это необходимо) можно предположить, что наша окрестность является звездной областью. Следовательно, существует определенная в этой окрестности 1-форма  $\beta$  такая, что

$$\sigma = d\beta, \quad \beta(0) = 0.$$

Следовательно, форма  $\Omega_t$  невырождена, и существует единственное (зависящее от  $t$ ) векторное поле  $\xi_t$ , для которого

$$i(\xi_t)\Omega_t = -\beta.$$

Поле  $\xi_t$  мы рассматриваем как зависящую от времени систему обыкновенных дифференциальных уравнений. Решение этих уравнений дает нам однопараметрическое семейство диффеоморфизмов  $f_t$ , для которого  $f_0 = \text{id}$  и касательный вектор к кривой  $f_t(x)$  в момент времени  $t$  есть  $\xi_t(f_t(x))$ . Кроме того,  $f_t(0) = 0$  для всех  $t$ . А теперь из фундаментальной формулы дифференциального исчисления форм (глава 18) мы имеем

$$\frac{d}{dt}f_t^*\Omega_t = f_t^*\left(\frac{d}{dt}\Omega_t + i(\xi_t)d\Omega_t + di(\xi_t)\Omega_t\right) = 0,$$

поскольку

$$d\Omega_t = 0, \quad \frac{d}{dt}\Omega_t = \sigma$$

и

$$di(\xi_t)\Omega_t = -d\beta = -\sigma.$$

Следовательно,

$$f_1^*\Omega_1 - \Omega_0 = f_1^*\Omega_1 - f_0^*\Omega_0 = \int_0^1 \frac{d}{dt}(f_t^*\Omega_t) dt = 0.$$

Итак, мы получили искомый диффеоморфизм  $f_1$ , что завершает доказательство теоремы Дарбу и вместе с ней доказательство теоремы о нормальной форме, на которую мы ссылались при доказательстве теоремы Каратеодори.

---

## Рекомендуемая литература

---

Повторим, что приведенный ниже перечень книг следует считать не библиографией, а дополнительной литературой. Книга LOOMIS AND STERNBERG еще заметнее, чем в случае тома I, претендует на роль параллельного учебника. Мы ссылаемся на нее, когда нужны точные (или более абстрактные) математические доказательства. Точно так же ФЕЙНМАНОВСКИЕ ЛЕКЦИИ можно считать универсальным справочником по физике.

Основное алгебраическое содержание тома II составляет переход от линейной алгебры к полилинейной. Книга GREUB является ясным и аккуратным введением в эту область. В частности, в ней дается подробное описание тензорной алгебры, симметрической алгебры и алгебр Клиффорда, которых мы только касаемся.

Главы 12 и 13 посвящены теории электрических цепей. Классический учебник по теории цепей — это GUILLEMIN. В книге LORRAIN AND CORSON обсуждаются теоремы о цепях из главы 8. Мы рекомендуем ознакомиться с вероятностным аспектом теории цепей опять-таки по книге DOYLE AND SNELL, а для более серьезного изучения почитать KELLY. По теории графов и ее приложениям мы рекомендуем книгу BONDY AND MURTY.

В главе 14 мы даем определения гомологий и когомологий с целью ввести соответствующий язык и не вдаемся в подробности. Для знакомства с этой областью на уровне этой книги см. книгу GIBLIN, а на более высоком математическом уровне находится книга MUNKRES. После ее прочтения можно перейти к книге BOTTL AND TU, которая представляет собой великолепную и элегантную трактовку многих глубоких вопросов алгебраической

топологии с точки зрения внешнего исчисления. Еще раз рекомендуем книгу GUILLEMIN AND POLLACK как геометрически интуитивное и математически строгое изложение дифференциальной топологии. Приложения внешнего исчисления, изучаемого в главе 15, можно найти в книге STERNBERG.

Главы 16–19 посвящены электричеству и магнетизму. Есть много замечательных книг по этому предмету. Наиболее доступен на уровне этой книги учебник PURCELL, элегантно и предельно ясно написанный, но в нем изложение ведется по принципу, отличному от нашего. В книге MISNER ET AL. электромагнетизм обсуждается с использованием дифференциальных форм.

Имеется две серии книг, где математическая физика излагается на существенно более высоком уровне, чем у нас. Это REED AND SIMON и THIRRING. Первая из них математическая, вторая физическая. Обе серии читать нелегко, но они того стоят. В THIRRING используется внешнее исчисление.

Классическим учебником по комплексному анализу является AHLFORS; его полезно прочесть после нашей главы 20. Еще одна хорошая книга — PÓLYA AND LATTA. Конечно, есть и многие другие. Про асимптотические методы и их связь с комплексным анализом можно прочесть в книге OLVER, имеющейся в большой и маленькой версии. Книга GUILLEMIN AND STERNBERG берет обсуждение асимптотики излучения из главы 21 в качестве отправной точки для изложения современной теории интегральных операторов Фурье и геометрической теории асимптотик. При этом от читателя требуется весьма серьезная математическая подготовка.

Имеются сотни книг по термодинамике и статистической механике. История этой науки в форме иконоборчества имеется в книге TRUESDELL. Кстати, он категорически не разделяет ту точку зрения на термодинамику, которая принята у нас здесь. Книга BORN является классическим полупопулярным изложением указанных тем. Наша трактовка статистической механики довольно близка к весьма объемной книге TRIBUS. Книга KITTEL не такая толстая и более элементарна, чем наше изложение. Стандартным учебником является REIF, а в книге THOMPSON хорошо

изложены различные вопросы с более выраженных математических позиций.

В этом курсе мы не занимались квантовой механикой, не считая краткого упоминания о квантовой логике в самом конце. Лучшее введение в этот предмет с математической точки зрения — книга MASKEY. Подробное обсуждение квантовой логики на высоком научном уровне содержится в VARADARAJAN. Книга ВОНМ — это сбалансированное и выверенное изложение, затрагивающее многие приложения. Другое доступное введение в данную область — SUDBERY.

Хинчин А. Я. Математические основания теории информации. — Физматгиз, 1954.

AHLFORS L. V. Complex Analysis. An Introduction to the Theory of Functions of One Complex Variable. — New York: McGraw-Hill, 1978 (3rd ed.). (International Series in Pure and Applied Mathematics.)

ВОНМ А. Quantum Mechanics: Foundations and Applications. — New York: Springer, 1993 (3rd ed.); reprinted 2001. (Texts and Monographs in Physics.)

BONDY J. A. AND MURTY U. S. R. Graph Theory with Applications. — New York: American Elsevier, 1976.

BORN M. The Natural Philosophy of Cause and Chance. — Oxford Univ. Press, 1964.

БОТТ Р. AND ТУ Л. W. Differential Forms in Algebraic Topology. — New York: Springer, 1982. (Graduate Texts in Mathematics, 82.)

*Русский перевод:* БОТТ Р., ТУ Л. В. Дифференциальные формы в алгебраической топологии. — М.: Наука, 1989.

DOYLE P. G. AND SNELL J. L. Random Walks and Electric Networks. — Washington, DC: Mathematical Association of America, 1985. (Carus Mathematical Monographs, 22.)

FEYNMAN R. P., LEIGHTON R. B. AND SANDS M. The Feynman Lectures on Physics. Vol. 1–3. — Reading, MA: Addison-Wesley, 1963–1965.

*Русский перевод:* ФЕЙНМАН Р. П., ЛЕЙТОН Р. Б. и СЭНДС М. Фейнмановские лекции по физике. Вып. 1–9. — М.: Мир, 1965–1967.

FEYNMAN R. P. Statistical Mechanics. A Set of Lectures. — Reading, MA: Perseus Books, 1998 (reprint of the 1972 original).

GIBLIN P. J. Graphs, Surfaces and Homology. An Introduction to Algebraic Topology. — London: Chapman & Hall, 1981 (2nd ed.).

- GREUB W. Multilinear Algebra. — New York: Springer, 1978 (2nd ed.). (Universitext.)
- GUILLEMIN E. A. Introductory Circuit Theory. — Wiley, 1953.
- GUILLEMIN V. W. AND POLLACK A. Differential Topology. — Englewood Cliffs, NJ: Prentice-Hall, 1974.
- GUILLEMIN V. W. AND STERNBERG S. Geometric Asymptotics. — Providence, RI: Amer. Math. Soc., 1977. (Mathematical Surveys, 14.)  
*Русский перевод*: ГИЙЭМЭН В. и СТЕРНБЕРГ Ш. Геометрические асимптотики. — М.: Мир, 1981.
- KELLY J. P. Reversibility and Stochastic Networks. — Chichester: Wiley, 1979. (Wiley Series in Probability and Mathematical Statistics.)
- KITTEL C. Thermal Physics. — W. H. Freeman, 1980.
- LAX P. D. AND PHILLIPS R. S. Scattering Theory. — Academic Press, 1967; 1989 (2nd ed.).  
*Русский перевод*: ЛАКС П. Д. и ФИЛЛИПС Р. С. Теория рассеяния. — М.: Мир, 1971.
- LOOMIS L. H. AND STERNBERG S. Advanced Calculus. — Boston, MA: Jones and Bartlett, 1990 (revised ed.).
- LORRAIN P. AND CORSON D. R. Electromagnetic Fields and Waves. — W. H. Freeman, 1988 (3rd ed.).
- MACKEY G. W. The Mathematical Foundations of Quantum Mechanics. — Mineola, NY: Dover, 2004 (reprint of the 1963 original).
- MARION J. B. AND HEALD M. A. Classical Electromagnetic Radiation. — New York: Academic Press, 1980 (2nd ed.).
- MISNER C. W., THORNE K. S. AND WHEELER J. A. Gravitation. — San Francisco, CA: W. H. Freeman, 1973.  
*Русский перевод*: МИЗНЕР Ч., ТОРН К. и УИЛЕР ДЖ. Гравитация. Т. 1–3. — М.: Мир, 1977.
- MUNKRES J. R. Elements of Algebraic Topology. — Menlo Park, CA: Addison-Wesley, 1984.
- OLVER F. W. J. Asymptotics and Special Functions. — Wellesley, MA: AK Peters, 1997 (reprint of the 1974 original).  
*Русский перевод*: ОЛВЕР Ф. Асимптотика и специальные функции. — М.: Наука, 1990.
- PÓLYA G. AND LATTA G. Complex Variables. — New York: Wiley, 1974.
- PURCELL E. Electricity and Magnetism. — McGraw-Hill, 1965.
- REED M. AND SIMON B. Methods of Modern Mathematical Physics. Vol. I–IV. — New York: Academic Press, 1972–79.

*Русский перевод:* РИД М. и САЙМОН Б. Методы современной математической физики. Т. 1–4. — М.: Мир, 1977–82.

REIF F. Fundamentals of Statistical and Thermal Physics. — McGraw-Hill, 1965.

ROSS S. M. Introduction to Probability and Statistics for Engineers and Scientists. — Burlington, MA: Academic Press, 2000 (2nd ed.).

SCOTT W. Physics of Electricity and Magnetism. — Wiley, 1977.

SOMMERFELD A. Electrodynamics. Lectures on Theoretical Physics, Vol. III. — New York: Academic Press, 1952.

STERNBERG S. Lectures on Differential Geometry. — New York: Chelsea, 1983 (2nd ed.).

*Русский перевод:* СТЕРНБЕРГ Ш. Лекции по дифференциальной геометрии. — М.: Мир, 1970.

SUDBERY A. Quantum Mechanics and the Particles of Nature. — Cambridge University Press, 1986.

THIRRING W. A Course in Mathematical Physics. Vol. 1–4. — New York–Vienna: Springer, 1978–83.

THOMPSON C. J. Mathematical Statistical Mechanics. — Princeton Univ. Press, 1972; reprinted 1979.

TRIBUS M. Thermostatistics and Thermodynamics. — Van Nostrand, 1961.

TRUEDELL C. A. The Tragical History of Thermodynamics, 1822–1854. — New York–Berlin: Springer, 1980. (Studies in the History of Mathematical and Physical Sciences, 4.)

VARADARAJAN V. S. Geometry of Quantum Theory. — New York: Springer, 1985 (2nd ed.).

---

# Предметный указатель

---

$\sigma$ -алгебра, 1137  
 $k$ -форма, см. дифференциальная форма степени  $k$   
0-коцепь, 604  
0-цепь, 591  
1-коцепь, 604  
1-форма, см. дифференциальная форма линейная  
— критическая, см. экстремаль  
1-цепь, 591  
2-форма, 386, 412  
4-импульс, 228  
—, закон сохранения, 229  
4-потенциал, 983

## А

абберация, 436  
— сферическая, 437  
абсолютная константа, см. скорость света  
абсолютная температура, 1097, 1111  
абстрактная температура, 1103  
адиабатическая кривая (адиабата), 1104  
адиабатически замкнутая система, 1104  
алгебра, 948  
—  $\mathbb{Z}_2$ -градуированная, 958  
— ассоциативная, 928, 948  
— внешняя ( $\Lambda(V)$ ), 947, 952  
— Дирака, 931  
— измеримых множеств, 1137

— кватернионов, 930  
— Клиффорда, 928, 951  
— матриц (и линейных преобразований), 34  
— с единицей, 948  
— тензорная ( $T(V)$ ), 949  
алгебраическое дополнение элемента матрицы, 560  
ампер, 576  
Ампера закон (второй), 886  
— в форме Максвелла, 971  
аналитическая функция, 1004  
аналитическое продолжение, 1025  
аннулирующее пространство (аннулятор), 492, 606  
ансамбль статистический, 1154  
антикоммутативность, 756  
антикоммутационные соотношения, 926  
аппроксимация дискретного распределения непрерывным, 819  
атлас карт многообразия, 790  
аффинная плоскость, 2  
аффинное преобразование, 23  
аффинное пространство (ассоциированное с векторным пространством), 12

## Б

базис, 14, 55, 488  
— дуальный, 491  
— ортонормированный, 177

## базисы

- подобные, 398
- противоположные, 398
- Бесселя функция, 1087, 1088
- биения, 194
- билинейная функция, 379, 750
- билинейное отображение, 937
- близнецов парадокс, 217
- Бозе–Эйнштейна распределение, 1172
- бозоны, 1142
- Больцмана–Пуассона распределение, 1172
- борновское разложение, 270
- Ботта периодичность, 968
- бутылка Клейна, 723

## В

## вакуум

- , электрические свойства, 840
- , конститутивные соотношения, 972
- вариационный принцип, 984
- вариация постоянных, 149
- ведущий элемент строки матрицы, 507
- Вейля метод ортогональных проекций, 645, 848
- вектор
  - касательный, 286
  - Пойнтинга, 987
  - энергии-импульса, см. 4-импульс
- векторное поле, 866
- векторное произведение, 934, 937
- векторное пространство, 6, 10, 476
  - , базис, 488
  - двумерное, 55
  - дуальное, 479
  - касательное, 290
  - линейных преобразований, 944
  - , натянутое на множество векторов, 484
  - одномерное, 13
  - , подпространство, 480
  - , размерность, 486
  - свободное над множеством, 941

- симплектическое, 461
- универсальное, 942
- вероятность (мера), 1139
- вечный двигатель второго рода, 1113
- взаимодействие, 1101
  - адиабатическое, 1104
- вихрь (rot), 934
- внешнее произведение, 383, 756
- внешняя алгебра ( $\Lambda(V)$ ), 947, 952
- внешняя производная, 379
- внешняя степень, 749, 754, 946
- внутреннее произведение, 872
- внутренняя энергия, 1096, 1105, 1171
- состояния, 1154
- волна
  - обратная (падающая), 990, 993
  - прямая (испускаемая), 990
  - распространяющаяся, 988
  - расходящаяся, 992
- волновой оператор, 974
- волновой пакет, 1080
- волновой фронт, 988
- вольт, 576
- вольт-амперная характеристика, 577
- время
  - истинное (по Ньютону), 15
  - собственное, 372
- вторичное излучение, 993, 1085
- второй закон термодинамики, 1097, 1107
- вынужденные колебания, 148
  - , резонанс, 154
- вырожденная матрица, см. линейное преобразование сингулярное
- вычет функции, 1012

## Г

- газовая постоянная, 1123
- Галилея преобразование, 221
- Гамильтона теорема, 452
- Гамильтона уравнения, 452
- Гамильтона–Кэли теорема, 86

гамма-распределение, 1168  
гамма-функция, 1066  
гармоническая функция, 844  
— сопряженная, 1053  
Гаусса закон (о потоке поля), 663,  
731, 825  
гауссово разложение, 448  
гауссово распределение, 1039  
Гельмгольца формула, 991, 1084  
генри, 577  
геометрия  
— конформная, 932  
— риманова, 933  
гессиан, 313  
Гиббса принцип минимальной  
энергии, 1132  
гиперболические синус и косинус,  
213  
главная ветвь, 1036, 1039  
главное значение Коши, 1042  
главные плоскости, 448  
голоморфная функция, 1004  
гомеоморфные пространства, 715  
гомологии, 709  
градиент (grad), 934  
Грама-Шмидта процесс, 176  
граница, 708  
граница области, 415  
граничное отображение, 591  
граничный оператор, 707, 820  
Грина оператор, 678  
Грина теорема взаимности, 657  
Грина формула, 415  
— вторая, 843  
— дискретная, 679  
— дискретная вторая, 682  
— первая, 842  
Грина функция, 851  
— дискретная, 678  
групповая скорость, 1080  
Гюйгенса принцип, 993  
— в формулировке Френеля, 1086

## Д

давление, 1106  
Дарбу теорема (о 2-формах), 1194

движение евклидово, 213  
двухступенчатый фильтр, 634  
дельта-функция Дирака, 823  
дерево, 597  
— максимальное, 599  
детерминант, см. определитель  
дефект линейного преобразования,  
503  
диаграмма Моллье, 1126  
диатермальный контакт, 1101  
дивергенция (div), 758, 934  
Дирака алгебра, 931  
Дирака дельта-функция, 823  
Дирихле задача, 830, 849  
— дискретная (для цепи), 668  
дискриминант квадратичной  
формы, 184  
дисперсия, 1167  
дифракция, 434  
дифференциал, 250, 254  
дифференциальная форма  
— замкнутая, 367, 385, 792  
— комплекснозначная, 1003  
— линейная (1-форма), 277  
— степени  $k$ , 749, 757  
— степени 2 (2-форма), 386, 412  
— точная, 363, 792  
дифференциальное уравнение  
—, интегральная кривая, 135  
— Лапласа, 317  
— линейное неоднородное, 148  
— —, установившееся решение, 152  
— линейное однородное, 125  
— —, начальные условия, 128  
— —, общее решение, 127  
—, фазовый портрет, 135  
дифференцируемое отображение,  
250  
—, регулярная точка, 524  
—, сохраняющее ориентацию, 401  
диэлектрическая постоянная, 840  
диэлектрическая проницаемость,  
579, 836  
длина  
— вектора, 166, 371  
— — лоренцева, 372

- кривой, 372
- оптическая, 452
- доктрина скрытого и явного тепла, 1117
- дуальное пространство, 479
- дуальный базис, 491

## Е

- евклидова плоскость, 165
- евклидово преобразование, 23, 166
- единичная матрица, 35
- емкость, 577

## З

- задача
  - Дирихле, 830, 849
  - — дискретная (для цепи), 668
  - Пуассона
  - — дискретная (для цепи), 669
- закон
  - Ампера (второй), 886
  - — в форме Максвелла, 971
  - Гаусса (о потоке поля), 663, 731, 825
  - Кеплера второй, 274
  - об отсутствии «магнитных зарядов», 886
  - сложения скоростей, 214
  - Снеллиуса, 436
  - — линеаризованный, 442
  - сохранения, 986
  - — 4-импульса, 229
  - — углового момента, 273
  - — энергии, 987
  - термодинамики
  - — второй, 1097, 1107
  - — нулевой (о равновесии), 1103
  - — первый, 1096, 1104
  - умножения определителей, 559
  - Фарадея, 731, 969
- законы
  - идеального газа, 1123
  - Кирхгофа, 580, 592, 612, 882
  - запаздывающие потенциалы, 982

- заплатки, см. атлас карт
- многообразия
- зародыш коцепи, 746, 817
- заряд, 576
  - «магнитный», 886
  - , распределение (плотность), 818
  - , скорость потока, 875
- затухание
  - аperiодическое, 146
  - критическое, 145
- звездное множество, 799
- звездный оператор (\*), 837, 903
- Зоммерфельда условия излучения, 992

## И

- идеал, 950
- идеальный газ, 1123
- идемпотентный оператор, 650
- излучение
  - вторичное, 993, 1085
  - , Зоммерфельда условия, 992
  - монохроматическое, 993
- измеримое пространство, 1137
  - , произведение, 1144
- изоморфизм векторных пространств, 14, 56
- изотерма, 1104
- изотропная среда, 837
- импеданс, 590
- инволюция, 956
- индекс строки матрицы, 507
- индикатор (функция), 1186
- индуктивность, 577
- индуктор, 576
- интеграл
  - , главное значение Коши, 1042
  - двойной
  - — абсолютный (первого рода), 391
  - — от 2-формы (второго рода), 389
  - криволинейный
  - — абсолютный (первого рода), 371

- от 1-формы (второго рода), 355
- Лебега, 1139
- несобственный, 428
- от  $k$ -формы, 765, 774
- по мере, 1138
- Френеля, 1039
- интерференция, 435
- источник питания, 612

## К

- калориметр, 1095
- «калорическая» функция, 1095
- каноническое преобразование, *см.* симплектическое отображение
- Каратеодори теорема, 1094
- Карно цикл, 1114
- касательное пространство, 290
- касательный вектор, 286
- квадратичная форма, 184
- квадратичная функция, 191
- квазистатистический процесс, 1106
- квантовая логика, 1187
- квантовая система, 1184
- кватернионы, 930
- кельвин (градус), 1123
- Кеплера закон второй, 274
- Кирхгофа законы, 580, 592, 612, 882
- Кирхгофа теорема, 603, 626
- Клейна бутылка, 723
- Клиффорда алгебра, 928, 951
- коаксиальный луч, 438
- когомологии, 731
- де Рама, 793
- кограница, 793
- кограничное отображение
  - ограниченное, 608
- кограничный оператор, 729
- количество информации, 1160
- коммутатор пары матриц, 271
- комплекс
  - двумерный, 705
  - одномерный, 590
  - размерности  $n$ , 706
  - связный, 595
- композиция отображений, 20

- конденсатор, 577
- конститутивные соотношения (для вакуума), 972
- контакт, 1131
- диатермальный, 1101
- конфигурационные переменные, 1104
- конформная геометрия, 932
- конформная матрица, 80
- кососимметричность, 750, 752
- коцель, 604, 728
- коцикл, 792
- Коши теорема интегральная, 1003
- Коши теорема о вычетах, 1012
- Коши формула интегральная, 1015
- Коши–Римана уравнения, 1005
- Коши–Шварца неравенство, 233
- коэффициент
  - полезного действия (КПД), 1114
  - преломления, 436
  - сжимаемости, 1130
  - теплового расширения, 1130
- кривая
  - адиабатическая, 1104
  - нулевая, 1090
  - обратимая, 1106
- кривизна, 441, 443
- круговой ток, 601
- кулон, 576

## Л

- Лагранжа множитель, 318
- лагранжево подмногообразие, 1128
- лагранжиан, 983
- Лапласа принцип, 1072
- Лапласа уравнение, 317
- лапласиан, 317
- $\square$ , 911
- дискретный (для цепи), 664
- ограниченный, 666
- Лебега мера и интеграл, 1139
- Лежандра полиномы, 180
- Лейбница правило, 881
- Лейбница формула (производная произведения функций), 263

лемма

— Морса, 339

— Пуанкаре, 797

линейная зависимость векторов, 9,  
47, 483

линейное преобразование, 28, 57

—, изменяющее ориентацию, 401

— конформное, 77, 1007

— Лоренца, 210

— — собственное, 209

—, матрица, 29

—, норма, 308

— ортогональное, 166

— регулярное (несингулярное), 28

— симплектическое, 231, 460

— сингулярное, 28

— скорости, 221

— сопряженное, 526

—, сохраняющее ориентацию, 400

— транспонированное, 183

линза

— тонкая, 443

—, фокальные плоскости, 445

—, фокусное расстояние, 444

лист Мёбиуса, 723, 737

Лиувилля теорема, 1021

логики дистрибутивный закон,  
1186, 1187

Лорана ряд, 1051

Лоренца преобразование, 210

— собственное, 209

Лоренца сила, 988

## М

магнитная восприимчивость, 886

магнитная индукция, 885

магнитное поле (возбуждение,  
напряженность), 886, 971

магнитные силовые линии, 885

Максвелла методы анализа цепей,  
580

Максвелла распределение, 1152

Максвелла уравнения, 972, 984

Максвелла–Больцмана

распределение, 1165, 1174

Маркова процесс, 93

—, матрица вероятностей перехода,  
95

масса покоя, 228

Массье функция, 1130, 1157, 1170

масштабное преобразование, 932

математическое ожидание, 1151

матрица

—, ведущий элемент строки, 507

— вероятностей перехода, 95, 684

— вращения, 564

— Гессе, 313

— единичная, 35

— емкостных коэффициентов, 688

— замены базиса, 58, 490

—, индекс строки, 507

— конформная, 80

— линейного преобразования, 29

— —, зависимость от базиса, 29

— обратная, 45

—, определитель, 549

— оптическая

— —, гауссово разложение, 448

— — общего вида, 448

— — преломления, 442

— — распространения, 441

— ортогональная, см. линейное  
преобразование ортогональное

— потенциальных коэффициентов,  
688

— присоединенная, 562

—, редукция по строкам, 506

— симплектическая, см. линейное  
преобразование  
симплектическое

— сингулярная (вырожденная), см.  
линейное преобразование  
сингулярное

— транспонированная, 183

—, функция от, 115

— Якоби, 303, 407

Мёбиуса лист, 723, 737

мера, 1137

— «беспорядка», 1159

— вероятностная, 1139

— гауссова (дискретная и непрерывная), 1143  
— Жордана, 767  
— Лиувилля, 1143  
—, перенос (pushforward), 1146  
мероморфная функция, 1012  
метод  
— вариации постоянных, 149  
— Гаусса, *см.* редукция по строкам  
— Кирхгофа расчета цепи, 650  
— круговых токов (Максвелла), 587, 613, 622  
— множителей Лагранжа, 317  
— Ньютона, 323  
— ортогональных проекций Вейля, 645, 848  
— Пикара, 325  
— узловых потенциалов (Максвелла), 588, 614, 624  
— эквивалентных сопротивлений, 584  
механический эквивалент тепла, 1105  
Минковского расстояние, 216  
многообразие, *см.* подмногообразие  
— равновесных состояний, 1101  
Моллье диаграмма, 1126  
Морса лемма, 339  
Муавра теорема, 81

## Н

наблюдаемое, 1149, 1170  
— «да или нет», 1185, 1187  
— квантовое, 1184  
накрытие, 1035, 1039  
напряжение, 576  
— холостого хода, 612  
напряженность  
— магнитного поля, 971  
— электрического поля, 824  
неравенство  
— Коши–Шварца, 233  
— треугольника, 216  
— — обратное, 217  
Нётер теорема, 985  
низкочастотный фильтр, 632

нильпотент, 54  
норма линейного преобразования, 308  
нормальная мода, 628  
нормальное распределение, 1039  
нормальные колебания, 192  
нулевая кривая, 1090  
нулевой вектор, 9  
нулевой конус, *см.* световой конус  
Ньютона метод, 323  
Ньютона уравнение (оптическое), 450  
Ньютона–Лейбница формула, 362, 415

## О

область  
— абсолютного будущего, 208  
— влияния, 976  
— зависимости, 976  
— звездная, 366, 799  
— односвязная, 1053  
обобщенные функции, 823  
образ  
— линейного преобразования, 51, 502  
— отображения, 18  
обратимая кривая, 1106  
обратная матрица, 45  
обратное отображение, 19  
обратное преобразование Фурье, 989, 1083  
однородная среда, 837  
ожидаемое значение, 1151, 1170, 1184  
окупанты, *см.* число заполнения  
ом, 577  
оператор  
— волновой, 974  
— — приведенный, 990  
— граничный, 707, 820  
— Грина, 678  
— звездный (\*), 837, 903  
— идемпотентный, 650  
— кограничный, 729  
— Лапласа, *см.* лапласиан

— линейный дифференциальный  
первого порядка, 921

— рождения фермионов, 926

— самосопряженный, 1183

— уничтожения фермионов, 926

— формально сопряженный, 920

определитель (детерминант), 42, 549

оптика

— волновая (теория Френеля), 434

— гауссова, 437

— геометрическая, 436

— линейная, 436

оптическая ось, 438

оптическая сила, 443

ориентация

— векторного пространства, 399, 902

— комплекса, 726, 778

— многообразия, 791

— области (на плоскости), 389, 416

— параллелограмма, 387

— поверхности (в пространстве), 391

— траектории, 351

ориентированная площадь  
(параллелограмма), 387

ориентированный граф, 590

ориентированный объем ( $n$ -мерного  
параллелепипеда), 547

ортогональная проекция, 179, 235

ортогональное преобразование, 166

ортогональность векторов, 166

ортонормированный базис, 177

ортонормировочная процедура, *см.*

Грама–Шмидта процесс

основная теорема алгебры, 1022

особая точка

— полюс, 1012

— существенная, 1030

— устранимая, 1018

осциллятор

— без затухания, 144

— с затуханием, 145

отображение, *также* функция,  
оператор, преобразование, 18

— аффинное, *см.* аффинное  
преобразование

— биективное, 19

— билинейное, 937

— вычисления, 939

— граничное, 591

— дифференцируемое, 250

— —, регулярная точка, 524

— —, сохраняющее ориентацию, 401

— евклидово, *см.* евклидово  
преобразование

— измеримое, 1145

— инъективное, 19

— Клиффорда, 951

— кограничное ограниченное, 608

— конформное, 1006

— кососимметричное, 945

— линейное, *см.* линейное  
преобразование

— обратное, 19

— ограниченное, 250

— открытое, 1031

— полилинейное ( $k$ -линейное), 945

— симплектическое, 467

— сюръективное, 19

— универсальное, 937

— экспоненциальное, 1086

## П

пакет волновой, 1080

параллельное соединение, 580

параметризация траектории, 352

первый закон термодинамики,  
1096, 1104

перенос

— меры (pushforward), 1146

— формы (pullback), 402, 761, 763

— функции (pullback), 291

периодичность Ботта, 968

перпендикулярность, *см.*  
ортогональность векторов

Пикара метод, 325

Планка постоянная, 1082

Планка функция, 1127

- плоскость
- аффинная, 2
  - проективная, 72
- плотность, 393
- заряда, 818
  - тока (скорость потока заряда), 875
- поверхность
- риманова, 1035
  - эквипотенциальная, 832, 1091
- подкомплекс, 596
- подмногообразие
- векторного пространства, 523
  - лагранжево, 1128
  - ориентируемое, 791
- подпространство векторного пространства, 480
- Пойнтинга вектор, 987
- показатель преломления, 438
- полюс функции, 1012
- поляризация света, 436
- полярные координаты, 291, 370
- порядок
- нуля, 1026
  - полюса, 1012
- последовательное соединение, 582
- постоянная Планка, 1082
- потенциал
- запаздывающий, 982
  - Льенара-Вихерта, 983
  - химический, 1134
  - электростатический, 686
- поток, 868
- энергии, 987
- правило
- Лейбница, 881
  - фаз, 1136
- преобразование
- масштабное, 932
  - Фурье, 1076
  - — обратное, 989, 1083
- приближение Борна (первое), 270
- принцип
- вариационный, 984
  - Гюйгенса, 993
  - — в формулировке Френеля, 1086
  - Дирихле (минимума энергии), 849
  - Лапласа, 1072
  - максимальной энтропии, 1132
  - максимума (для гармонических функций), 846
  - минимальной энергии Гиббса, 1132
  - причинности, 976
  - устранимых особенностей, 1018
  - Ферма (наименьшего времени), 457
  - электрической экранировки, 831
- присоединенная матрица, 562
- проводимость, 580, 864
- проективная плоскость, 72
- проекция, 52
- ортогональная, 179, 235
- произведение матриц, 31
- производная
- внешняя, 379, 881
  - Ли (по направлению), 288, 880
  - многомерная (отображения векторных пространств), см. также дифференциал, 247
  - произведения функций, см. Лейбница формула
  - смешанная (вторая), 282
  - частная, 275
- производящая функция, 1129
- простой (порядка 1) нуль или полюс, 1026
- пространство
- $k$ -форм, 754, 946
  - аннулирующее, 606
  - билинейных функций, 750
  - гомологий, 709
  - границ, 593, 708
  - когомологий, 731
  - циклов, 708
- пространство-время, 205
- противоположный вектор, 10
- процесс
- Грама-Шмидта, 176
  - квазистатический, 1106

- Маркова, 93
- —, матрица вероятностей перехода, 95

прямая сумма векторных пространств, 596

Пуанкаре лемма, 797

Пуанкаре преобразование, 213

Пуассона распределение, 1169

Пуассона уравнение, 837, 915

— дискретное (для цепи), 664

— ограниченное, 664

Пуассона формула интегральная, 855, 1055

Пуассона ядро, 680

## Р

работа, 1095, 1156

— силового поля, 349

равновесие тепловое, 1102

равномерная дифференцируемость, 326

равномерная сходимость последовательности голоморфных функций, 1047

радиус сходимости, 1050

разбиение единицы, 785

разложение

— борновское, 270

— гауссово, 448

размерность векторного пространства, 486

ранг

— дифференциальной (1- или 2-) формы, 1192

— линейного преобразования, 503

распределение

— Бозе-Эйнштейна, 1172

— Больцмана-Пуассона, 1172

— гамма-, 1168

— гауссово (нормальное), 1039

— Максвелла, 1152

— Максвелла-Больцмана, 1165, 1174

— Пуассона, 1169

— Ферми-Дирака, 1172

распространяющаяся волна, 988

расстояние Минковского, 216

регулярная точка

(дифференцируемого отображения), 524

редукция по строкам, 506

резистор, 577

резисторный порт, 659

резонанс, 154, 194

резонансная частота, 628, 633

риманова геометрия, 933

риманова поверхность, 1035

род поверхности, 722

росток голоморфной функции, 1025

ротатор (rot), 934

Руше теорема, 1028

ряд

— Лорана, 1051

— степенной, 116

— — абсолютно сходящийся, 119

— — сходящийся, 118

— Тейлора, 1020, 1049

— Фурье, 1052

## С

световой конус, 210

свободное векторное пространство, 941

связное множество, 846

сила

— Лоренца, 988

— обобщенная, 1156

силовое поле, 348

—, работа, 349

силовые линии магнитные, 885

сименс, 580

симплектическая группа, 232, 461

симплектическая форма, 461

симплектическое векторное пространство, 461

симплектическое отображение, 467

сингулярная (вырожденная)

матрица, см. линейное

преобразование сингулярное

система, 1100, 1170

— адиабатически замкнутая, 1104

— квантовая, 1184

скалярное произведение, 167  
— в пространстве  $k$ -форм, 900  
— дуальное (1-форм), 899  
— евклидово, 175  
— из векторного анализа (dot product), 934  
—, свойства, 169  
— симплектическое, 231  
скорости преобразование, 221  
скорость  
— групповая, 1080  
— звука, 1119  
— потока заряда, 875  
— света, 206, 973  
сложения скоростей закон, 214  
случайное блуждание (по графу), 684  
смещение, см. трансляция  
аффинная  
Снеллиуса закон, 436  
— линеаризованный, 442  
собственное время, 372  
собственное значение, 82  
собственный вектор, 82  
сопротивление, 577  
— комплексное, см. импеданс  
— нелинейное, 578  
— удельное, 864  
сопряженное отображение, 526  
сопряженные плоскости, 443  
состояние, 1100, 1140, 1170  
— квантовое, 1184  
— равновесное, 1101, 1165, 1170, 1184  
специальная теория  
относительности, 206  
статистика (Бозе–Эйнштейна, Ферми–Дирака), см. распределение  
Стирлинга формула, 1065  
Стокса теорема, 757, 777  
суперкоммутативность, 756  
существенная особенность, 1030

**Т**

Тейлора ряд, 1020, 1049

Тейлора формула (многомерная), 311  
телескоп астрономический, 447  
телесный угол, 822  
Теллегена теорема, 605  
температура  
— абсолютная, 1097, 1111  
— абстрактная, 1103  
— эмпирическая, 1103  
тензор  
— типа  $(k, 0)$ , 754  
— энергии-импульса, 988  
тензорная алгебра  $(T(V))$ , 949  
тензорное произведение, 940, 945  
— скрученное, 958  
теорема  
— алгебры основная, 1022  
— взаимности Грина, 657, 687  
— Гамильтона, 452  
— Гамильтона–Кэли, 86  
— Грина, см. Грина формула  
— Дарбу (о 2-формах), 1194  
— Каратеодори, 1094  
— Кирхгофа, 603, 626  
— Коши интегральная, 1003  
— Коши о вычетах, 1012  
— коэффициентов Морзе, 457  
— Лиувилля, 1021  
— Муавра, 81  
— Нётер, 985  
— о кососимметрических билинейных формах, 1198  
— о неявной функции, 334  
— о нормальной 1-форме, 1099, 1194  
— о постоянном ранге, 517  
— о ранге и дефекте, 503  
— о среднем значении, 305  
— — гармонической функции, 845  
— — многомерная (неравенство), 309  
— об обратной функции, 332  
— Радона–Никодима, 1149  
— Руше, 1028  
— спектральная (для операторов), 1183

- Стокса, 757, 777
- Теллегена, 605
- Эйлера для комплекса, 714
- тепло, 1095
- механический эквивалент, 1105
- скрытое (латентное), 1118
- тепловое равновесие, 1102
- теплоемкость, 1117
- термометр, 1103
- ток, 576
- короткого замыкания, 612
- круговой, 601
- через поверхность, 864
- тонкая линза, 443
- тор, 718
- точечная характеристика  
(оптической системы), 452
- точка ветвления, 1035
- траектория
- кусочно-дифференцируемая, 352
- ориентированная, 351
- , параметризация, 352
- простая (в цепи), 585
- трансляция аффинная, 6, 25
- транспонирование, 183
- формальное (оператора), 821
- трилинейная функция, 752

## У

- увеличение
- линзы, 446
- телескопа угловое, 447
- угловой момент, 272
- , закон сохранения, 272
- угол между векторами, 167
- удельное сопротивление, 864
- умножение матриц, 31
- универсальное отображение, 937
- универсальное пространство, 942
- уравнение
- волновое
- — неоднородное, 981
- — одномерное (колебаний  
струны), 974
- — приведенное (Гельмгольца),  
991

- — трехмерное, 977
- дифференциальное, см.  
дифференциальное уравнение
- Лапласа, 317
- — двумерное, 317, 665
- — дискретное (для цепи), 664
- Ньютона (оптическое), 450
- Пуассона, 837, 915
- — дискретное (для цепи), 664
- — ограниченное, 664
- характеристическое, 83
- уравнения
- Гамильтона, 452
- Коши-Римана, 1005
- Максвелла, 972, 984
- условия излучения Зоммерфельда,  
992
- условный экстремум, 317
- устраняемая особенность, 1018

## Ф

- фаз правило, 1136
- фаза волны, 988
- факторалгебра, 950
- факторпространство, 497
- фарад, 577
- Фарадея закон, 731, 969
- Ферма принцип (наименьшего  
времени), 457
- Ферми-Дирака распределение, 1172
- фермионы, 1142
- фильтр
- двухступенчатый, 634
- низкочастотный, 632
- фокальные плоскости, 445
- фокусное расстояние, 444
- формальное сопряжение, 920
- формальное транспонирование  
(оператора), 821
- формула
- Гельмгольца, 991, 1084
- Грина, 415
- — вторая, 843
- — дискретная, 679
- — дискретная вторая, 682
- — первая, 842

- де Бройля, 1081
- дифференциального исчисления
  - фундаментальная, 881
- замены переменных в  $k$ -мерном интеграле, 768
- Кирхгофа, 993
- Коши интегральная, 1015
- Лейбница (производная произведения функций), 263
- Ньютона–Лейбница, 362, 415
- обратного преобразования Фурье, 1083
- Пуассона интегральная, 855, 1055
- стационарной фазы, 1074
- Стирлинга, 1065
- Тейлора (многомерная), 311
- Эйнштейна, 1081
- Френеля интеграл, 1039
- фронт волновой, 988
- функция
  - аналитическая, 1004
  - Бесселя, 1087, 1088
  - билинейная, 379, 750
  - гамма-, 1066
  - гармоническая, 844
  - — сопряженная, 1053
  - голоморфная, 1004
  - Грина, 851
  - — дискретная, 678
  - Дирака (дельта-), 823
  - Дирихле (разрывная интегрируемая), 1139
  - «калорическая», 1095
  - квадратичная, 191
  - комплекснозначная дифференцируемая, 999
  - Массье, 1130, 1157, 1170
  - мероморфная, 1012
  - обобщенная, 823
  - от матрицы, 115
  - Планка, 1127
  - производящая, 1129
  - разбиения, 1157, 1170, 1184
  - с компактным носителем, 818
  - трилинейная, 752

- характеристическая, 1186
- Хевисайда, 1141
- Фурье преобразование, 1076
- обратное, 989, 1083
- Фурье ряд, 1052
- Фурье-образ, 1083

## Х

- характеристический полином, 83
- характеристическое уравнение, 83
- химический потенциал, 1134

## Ц

- цепное правило, 256
- цепь комплекса, 591
- цикл, 593, 708
- Карно, 1114
- циркуляция магнитного поля, 971

## Ч

- частная производная, 275
- число заполнения, 1141

## Э

- эйконал, см. точечная характеристика
- Эйлера теорема для комплекса, 714
- эйлерова характеристика, 714
- эквипотенциальная поверхность, 832, 1091
- экспоненциальное отображение, 1086
- экстремаль (лагранжиана), 983
- электрическое смещение, 825
- электродвижущая сила (ЭДС), 578
- электростатическая энергия, 842, 849
- электростатический потенциал, 686
- эмпирическая температура, 1103
- энергии-импульса тензор, 988
- энергия
  - взаимодействия, 1131
  - внутренняя, см. внутренняя энергия

- кинетическая, 1151
- свободная Гельмгольца, 1191
- электростатическая, 843, 849
- энтальпия, 1125
- энтропия, 343, 1097, 1108, 1159,  
1170, 1184
- эффективность (КПД), 1114

## Я

- ядро
  - линейного преобразования, 52,  
502
  - Пуассона, 680
- якобиан, *см.* матрица Якоби
- ячейка
  - комплекса, 706
  - электрической цепи, 585, 592

