Дж.Блейкмор Физика твердого тела

Дж.Блейкмор

Физика твердого тела

Издательство «Мир»

Solid State Physics

Second edition

J. S. BLAKEMORE
Department of Applied Physics
and Electrical Engineering
Oregon Graduate Center

Cambridge University Press
Cambridge
London New York New Rochelle
Melbourne Sydney

Дж. Блейкмор

Физика твердого тела

Перевод с английского
под редакцией
д-ра физ.-мат. наук, проф. Д. Г. АНДРИАНОВА
и д-ра физ.-мат. наук, проф. В. И. ФИСТУЛЯ

Москва «Мир» 1988

ББК 22.37
Б 68
УДК 539.2

Переводчики: д-р физ.-мат. наук, проф. Д. Г. Андрианов, Е. А. Гулько, канд. физ.-мат. наук Т. Ю. Лисовская, канд. физ.-мат. наук Н. С. Рытова

Блейкмор Дж.

Б68 Физика твердого тела: Пер. с англ.— М.: Мир, 1988.— 608 с., ил.

ISBN 5-03-001256-7

В книге американского автора излагаются общие вопросы кристаллического строения и динамики решетки, рассматриваются наиболее существенные особенности металлов и полупроводников, а также диэлектрические и магнитные свойства твердых тел и резонансные явления в них. Книга написана с большим педагогическим мастерством; задачи, помещенные в конце каждой главы, дополняют текст и могут быть использованы при самостоятельном изучении предмета.

Книга предназначена для студентов-физиков и металловедов, начинающих знакомство с физикой твердого тела. Она принесет пользу также научным сотрудникам и инженерам, желающим самостоятельно изучать физику твердого тела или расширить свои знания.

Б $\frac{1704060000—495}{041(01)—88}$ 77—88, ч. 1 ББК 22.37

Редакция литературы по физике и астрономии

ISBN 5-03-001256-7 (русск.)
ISBN 0-521-30932-8 (англ.)

© Cambridge University Press, 1985
This book was originally published in the English language by Cambridge University Press of Cambridge, England
© перевод на русский язык, «Мир», 1988

Предисловие редакторов перевода

Предлагаемая вниманию читателей книга представляет собой перевод второго американского издания, значительно расширенного и дополненного по сравнению с первым изданием, опубликованным на английском языке в 1970 г. и переведенным на русский язык в 1972 г. в издательстве «Металлургия» (*Дж. Блейкмор,* Физика твердого состояния).

За прошедшие 15 лет значительно расширились области практического применения науки о твердом теле. Появились новые типы твердотельных лазеров, возникла целая отрасль, производящая материалы на основе гранатов, существенно преобразовалась полупроводниковая электроника, в которой все шире используются сложные структуры, представляющие собой контакты твердых тел различной природы, например, металл — диэлектрик — полупроводник. Наметились перспективы применения других твердотельных материалов — на основе редкоземельных металлов, магнитных полупроводников, аморфных металлов и полупроводников. Наконец, в 1987 г. научный мир был буквально потрясен открытием высокотемпературной сверхпроводимости *.

За истекшие 15 лет не стояла на месте и теория твердого состояния вещества. Отметим многолетнюю дискуссию о природе проводимости аморфных тел и о характере перехода метал — диэлектрик. Большое внимание ученых привлекают различные стимулированные процессы в твердых телах: светоускоренная диффузия примесей, рекомбинационно-ускоренные атомные перестройки в диэлектрических и полупроводниковых кристаллах. По-прежнему вызывают интерес различного рода как близко-, так и дальнодействующие взаимодействия примесей и дефектов в твердых телах. Много исследований посвящено

---

* *Bednorz J. G., Müller K. A.*— *Z.* Phys. B — Condensed Matter, **64**, 189 (1986); *Uchida S., Takagi H., Kitizawa K., Tanaka S.*— Jpn. J. Appl. Phys, **26**, L1 (1987); *Chu C. W., Hor P. H., Meng R. L., Gao L., Huang Z. J., Wang Y. Q.*— Phys. Rev. Lett., **58**, 405 (1987).

совокупности физических явлений в растворах магнитных примесей в диамагнитных металлах и полупроводниках. В металлах эти явления обусловливают аномалии гальваномагнитных эффектов и приводят к возникновению состояния спинового стекла. Значительная величина обменного взаимодействия носителей заряда с локализованными на магнитных атомах магнитными моментами в полупроводниках дает основание выделить такие твердые растворы в отдельный класс полумагнитных полупроводников. В этих веществах помимо аномалий гальваномагнитных эффектов и возникновения состояния спинового стекла имеют место значительные магнетооптические эффекты, а также зависимость ширины запрещенной зоны от магнитного поля.

Наряду с теоретическими и практическими достижениями науки о твердом теле за истекшие 15 лет развивались и совершенствовались магнитные, магнито-резонансные, локальные и другие методы всестороннего исследования твердых тел. По-видимому, перечисленные причины и побудили автора выпустить второе издание своей книги. Во всяком случае, на наш взгляд, ее появление весьма своевременно.

Как отмечалось в предисловии к переводу первого издания, книги, посвященные общим вопросам физики твердого тела, можно разделить на две группы. К первой относятся те, которые рассчитаны на физиков, хорошо знакомых с квантовой механикой. Эти книги трудны для металлургов, технологов и материаловедов. Для такой категории читателей предназначены книги, составляющие вторую группу, но они, как правило, либо носят описательный характер, либо рассматривают лишь структурный аспект твердых тел.

Книга Дж. Блейкмора удачно дополняет книги второй группы. В ней автор, не злоупотребляя математическим аппаратом, излагает теоретические вопросы физики твердого тела и подробно рассматривает связь электронной и кристаллической структур, дефекты твердых тел, их электро- и теплопроводность, упругие свойства и другие важные особенности. Таким образом, книга Дж. Блейкмора перекидывает мост через «пропасть», которая сегодня еще, к сожалению, существует между теорией твердого тела и практическим материаловедением. Именно поэтому перевод первого издания книги был хорошо встречен советским читателем. Об этом говорят и отзывы о ней, и отсутствие ее на прилавках букинистических магазинов.

Как уже говорилось, второе издание существенно отличается от первого: автором значительно расширена глава, посвященная диэлектрическим и магнитным свойствам твердых тел, введены новые разделы, в которых рассматриваются резонансные явления, прыжковая проводимость. Практически переработаны

заново многие разделы в главе «Электроны в металлах». Наконец, автор переработал многие задачи и значительно увеличил их число. Следует отметить, что решение задач потребует значительных усилий от читателей; в ряде случаев понадобится даже проведение своего рода теоретических мини-исследований. Вместе с тем, задачи могут служить хорошим пособием для преподавателей и аспирантов вузов и научных работников. В целом книга вполне доступна для инженеров и научных сотрудников, желающих самостоятельно изучить физику твердого тела или повысить свои знания в этой области.

При переводе сделано небольшое число примечаний и несколько дополнены ссылки на литературу в тех случаях, когда имеются более доступные для читателя советские источники. Перевод книги выполнили Е. А. Гулько (гл. 1), канд. физ.-мат. наук Т. Ю. Лисовская (гл. 2), канд. физ.-мат. наук Н. С. Рытова (гл. 3, 4) и д-р физ.-мат. наук Д. Г. Андрианов (гл. 5). Редактирование гл. 1 и 2 выполнено проф. д-ром физ.-мат. наук Д. Г. Андриановым, гл. 3—5 — проф. д-ром физ.-мат. наук В. И. Фистулем.

Д. Г. Андрианов
В. И. Фистуль

Предисловие

Эта книга была написана как пособие к вводному курсу лекций по физике твердого тела, рассчитанному на одну четверть или один семестр. Обычно подобный курс входит в программу последних двух лет обучения студентов, специализирующихся в области физики. Однако эта книга задумывалась с тем расчетом, чтобы она оказалась полезной и для специалистов в других областях: химии, электротехнике, материаловедении и т. п., у которых до аспирантуры могло и не возникнуть потребности в изучении соответствующего курса. При рассмотрении некоторых вопросов (например, зонной теории) предполагается, что читатель уже знаком с языком и основными положениями квантовой механики в объеме односеместрового курса по «современной физике». Если читатель прослушал общий курс квантовой механики, то этого вполне достаточно для восприятия этой книги, однако во многих случаях у читателя, очевидно, такой подготовки нет. Поэтому, если уравнение Шредингера время от времени и появляется в тексте, то доказательства выводимых соотношений на основе квантовой механики полностью опускаются.

Таким образом, цель этой книги состоит в том, чтобы достаточно строго, но в то же время и понятно изложить основы физики твердого тела для впервые знакомящегося с этой темой читателя. Особое внимание уделено рассмотрению кристаллов, начиная от симметрии решетки и кончая понятиями обратного пространства и зон Бриллюэна. В гл. 2—4 эти понятия применяются соответственно при изложении теории колебаний решетки, теории металлов и кристаллических полупроводников. Примерно 75 % всего объема книги отведено следствиям периодичности в расположении атомов. Для того, чтобы полный объем книги не выходил за рамки, приемлемые для начального курса физики твердого тела, рассмотрение ряда вопросов физики конденсированного вещества, включенных в книгу, по необходимости является кратким и носит обзорный характер. Это

относится к разделам, посвященным дефектам кристаллической решетки, аморфным твердым телам, сверхпроводимости, диэлектрическим и магнитным явлениям, а также магнитному резонансу.

Предлагаемый здесь текст в основном повторяет текст второго издания, выпущенного в 1974 г. издательством «Сондерс». Однако благодаря предоставленной мне издательством «Кембридж юниверсити пресс» возможности опубликовать эту книгу в настоящий момент, я смог исправить ряд ошибок, добавить более подробные объяснения (например, в конце разд. 1.5), изменить некоторые рисунки и обновить списки литературы, помещенные после предисловия и в конце каждой главы. В книге, конечно, сохранена принятая во втором издании система единиц измерения СИ. Необходимо, однако, упомянуть два имеющихся в книге исключения из системы СИ: межатомные расстояния измеряются в ангстремах (Å), а энергия в расчете на один электрон — в *электронвольтах*. Похоже, что кристаллографы пока не собираются выражать постоянные кристаллической решетки в нанометрах; проще пользоваться множителем 10^{-10} для перехода от ангстремов к метрам. Использование электронвольта вместо $1{,}6 \cdot 10^{-19}$ Дж также упрощает рассмотрение многих явлений, сопровождающихся преобразованием энергии. Выбор системы единиц, несомненно, важен, когда требуется довести до числа решение какой-либо из задач для самостоятельной работы.

Задачи (всего их 125) приводятся в конце каждой главы. Многие из них требуют численных расчетов. Это сделано для того, чтобы читатель обратил внимание главным образом на относительный порядок различных величин и их взаимосвязь, а не на число десятичных знаков в результате. Задачи умышленно подобраны так, чтобы они сильно различались по сложности и объему вычислений; мне несколько раз говорили, что для текста такого уровня некоторые из задач слишком сложны. Такие задачи, несомненно, будут хорошим испытанием для тех, кто, изучив данную книгу, хочет перейти к одному из современных курсов физики твердого тела, перечень которых приведен в общем списке литературы после предисловия.

Как и в предыдущих изданиях этой книги, число приведенных в сносках ссылок на литературу в ней намного больше, чем обычно принято в пособиях для студентов. Кроме того, в конце каждой главы приводится список книг, из которого читатель может сам выбрать оптимальную для себя литературу для *дополнительного* чтения. Работы, цитированные в сносках, подобраны с таким расчетом, чтобы они могли помочь читателю в изучении литературы и через несколько лет после первого знакомства с этой книгой.

Данная книга была написана с целью отразить *идеи,* лежащие в основе физики твердого тела, а не с целью дать компиляцию из фактов и цифр. Соответственно в ней всего несколько таблиц, содержащих численные значения различных характеристик, в то время как, например, в пятом издании известного учебника Киттеля содержится около 60 таблиц. Читатель, которому потребуются количественные значения физических характеристик твердых тел, может обратиться к множеству источников; обширные сведения приведены в Справочнике Американского института физики (American Institute of Physics Handbook), последнее издание которого вышло в 1972 г., и в Справочнике по химии и физике (Handbook of Chemistry and Physics), который выходит ежегодно. Кроме того, как указано в общем списке литературы на с. 11, в последние годы появились новые тома таблиц Ландольта — Бернштейна (Landolt — Börnstein Tables), в которых приводятся сведения о некоторых полупроводниковых материалах. Работа по накоплению и уточнению информации о численных значениях характеристик твердых тел, конечно же, продолжается непрерывно.

В течение тех лет, когда писалась эта книга и подготавливались последующие издания, многие помогали мне: одни вносили предложения, улучшающие текст, другие разрабатывали задачи, а третьи подбирали иллюстративный материал. Я попрежнему благодарен всем тем, кому я персонально выразил признательность в предисловиях к первому и второму изданиям. В настоящем втором издании книги, выпущенном издательством «Кембридж юниверсити пресс», мне хочется выразить особую признательность Л. Е. Мюрру из Учебного центра в Орегоне за фотографии, благодаря которым появился ряд новых интересных и полезных рисунков в гл. 1, и Г. К. Генишу из Пенсильванского университета за фотографию, использованную в качестве рис. 1.2.

Дж. Блейкмор

Бивертон, Орегон
март 1985

Общая литература *

Физика твердого тела (начальный и средний уровень подготовки)

Bube R. H. Electronics in Solids: An Introductory Survey, Academic Press, 1981.

Bube R. H. Electronic Properties of Crystalline Solids, Academic Press, 1973.

Dekker A. J. Solid State Physics, Prentice-Hall, 1957. [Эта книга с 1957 г. не переиздавалась, однако содержит интересное рассмотрение некоторых вопросов, не затрагиваемых в других книгах.]

Problems in Solid State Physics, ed. H. J. Goldsmid, Pion, 1968.

Harrison W. A. Electronic Structure and the Properties of Solid, Freeman, 1980. [Имеется перевод: *Харрисон У. А.* Электронная структура и свойства твердых тел.— М.: Мир, 1983.]

Kittel C. Introduction to Solid State Physics, 5th, ed., Wiley, 1976. [Имеется перевод: *Киттель Ч.* Введение в физику твердого тела.— М.: Наука, 1978.]

McKelvey J. P. Solid State and Semiconductor Physics, Krieger, 1982.

Rosenberg H. M. The Solid State, 2nd ed., Oxford Univ. Press. 1979.

* *Каганов М. И.* Электроны, магноны, фононы.— М.: Наука, 1979.

Физика твердого тела (для более углубленного изучения)

Ashcroft N. W., Mermin N. D. Solid State Physics, Holt, 1976. [Имеется перевод: *Ашкрофт Н., Мермин Н.* Физика твердого тела — М.: Мир, 1979.]

Callaway J. Quantum Theory of the Solid State, v. 1, 2, Academic Press, 1974.

Goodstein D. L. States of Matter, Prentice-Hall, 1975.

Harrison W. A. Solid State Theory, Dover, 1980.

Haug A. Theoretical Solid State Physics, Pergamon Press, v. 1, 2, 1972.

Jones W., March N. H. Theoretical Solid State Physics, v. 1, 2, Wiley, 1973.

Kittel C. Quantum Theory of Solids, Wiley, 1963. [Имеется перевод: *Киттель Ч.* Квантовая теория твердых тел.— М.: Наука, 1967.]

Solid State Physics, eds. R. Kubo, T. Nagamiya, McGraw-Hill, 1969.

Solid State Theory, ed. P. T. Landsberg, Methods and Applications, Wiley, 1969.

Peierls R. E. Quantum Theory of Solids, Oxford, 1965. [С тех пор не переиздавалась, но продолжает оставаться классическим учебником.] [Имеется перевод 1-го издания: *Пайерлс Р.* Квантовая теория твердых тел.— М.: Гостехиздат, 1956.]

Seitz F. Modern Theory of Solids, McGraw-Hill, 1940. [Более не переиздавалась, но отнюдь не потеряла своего значения.] [Имеется перевод: *Зейтц Ф.* Современная теория твердого тела.— М.— Л.: Гостехиздат, 1949.]

Ziman J. Principles of the Theory of Solids, 2nd ed., Cambridge Univ. Press, 1972. [Имеется перевод: *Займан Дж.* Принципы теории твердого тела.— М.: Мир, 1974.]

---

* Здесь и далее звездочками отмечены работы, добавленные при переводе.— *Прим. ред.*

Электронные свойства твердых тел

Bar-Lev A. Semiconductors and Electronic Devices, Prentice-Hall, 1979.
VLSI Electronics: Microstructure Science, ed. N. G. Einspruch, v. I — v. 8, Academic Press, 1981—1984 (будет следовать продолжение).
Elliott R. J., Gibson A. F. An Introduction to Solid State Physics and its Applications, Barnes and Noble, 1974.
Grove A. S. Physics and Technology of Semiconductor Devices, Wiley, 1967.
Milnes A. G. Semiconductor Devices and Integrated Circuits, Van Nostrand, 1980.
Moss T. S., Barrett G. J., Ellis B. Semiconductor Optoelectronics Butterworths, 1973.
Streetman B. G. Solid State Electronic Devices, Prentice-Hall, 2nd ed., 1980.
Sze S. M. Physics of Semiconductor Devices, 2nd ed., Wiley, 1981. [Имеется перевод: *Зи С. М.* Физика полупроводниковых приборов.— М.: Мир, 1984.]
Wang F. F. Y. Introduction to Solid State Electronics, North-Holland, 1980.
* *Слетер Дж.* Диэлектрики, полупроводники, металлы.— М.: Мир, 1969.
* *Лифшиц И. М., Азбель М. Я., Каганов М. И.* Электронная теория металлов.— М.: Наука, 1971.
* *Бонч-Бруевич В. Л., Калашников С. Г.* Физика полупроводников.— М.: Наука, 1977.

Квантовые явления

Bockhoff F. J. Elements of Quantum Theory, Addison-Wesley, 2nd ed., 1976.
French A. P., Taylor E. F. An Introduction to Quantum Physics, Norton, 1978.
Ridley B. K. Quantum Processes in Semiconductors, Oxford. Univ. Press, 1981. [Имеется перевод: *Ридли Б. К.* Квантовые процессы в полупроводниках.— М.: Мир, 1986.]
Problems in Quantum Mechanics, ed. D. ter Haar, 3rd ed., Pion, 1975.
* *Ландау Л. Д., Лифшиц Е. М.* Квантовая механика.— М.: Наука, 1964.

Статистическая физика

Fujitta S. Statistical and Thermal Physics, Krieger, 1984.
Kittel C. Elementary Statistical Physics, Wiley, 1958. [Имеется перевод: *Киттель Ч.* Элементарная статистическая физика.— М.: ИЛ, 1960.]
Kittel C., Kroemer H. Thermal Physics, Freeman, 2nd ed., 1980.
Mohling F. Statistical Mechanics, Wiley-Halsted, 1982.
Reichl L. E. A Modern Course in Statistical Physics, Univ. Texas, 1980.
Tolman R. C. The Principles of Statistical Mechanics, Dover, 1979.
* *Ландау Л. Д., Лифшиц Е. М.* Статистическая физика.— М.: Наука, 1964.

Волновые явления

Brillouin L. Wave Properties and Group Velocity, Academic Press, 1960.
Brillouin L. Wave Propagation in Periodic Structures. Dover, 1972. [Имеется перевод: *Бриллюэн Л.* Распространение волн в периодических структурах.— М.: ИЛ, 1959.]
Main I. G. Vibrations and Waves in Physics, Cambridge Univ. Press, 1978.
Squire C. F. Waves in Physical Systems, Prentice-Hall, 1971.

Численные значения физических величин

American Institute of Physics Handbook, McGraw-Hill, 3rd ed., 1972.
Handbook of Chemistry and Physics, CRC Press, 66th ed., 1985.
Handbuch der Physik (главный редактор серии из 54 томов S. Flügge), Springer-Verlag, 1956—1974.
Landolt—Börnstein Tables, Springer-Verlag [Серия начинается до 1950-х годов, последние тома из этой серии включают в себя данные по физике твердого тела.]

1

Кристаллическая структура и форма твердых тел

Твердые тела можно классифицировать в соответствии с различными критериями. Среди более важных из них — разделение твердых тел на *кристаллические* и *аморфные*. За период с середины 1940 г. по конец 1960 г. в физике твердого тела уделялось значительно больше внимания кристаллам, чем аморфным телам, изучение которых представляет собой более трудную задачу.

Аморфные тела проявляют главным образом ближний порядок в связях между ближайшими соседями, а не дальний порядок, присущий периодической атомной решетке; примерами этого являются статистические полимеры, угольная сажа, аллотропные формы таких элементов, как селен и сурьма, а также стекла. Стекло можно представить себе как переохлажденную жидкость, вязкость которой слишком велика для того, чтобы расположение атомов могло принять более упорядоченную форму. Поскольку степень упорядоченности аморфных тел сильно зависит от условий их приготовления, можно предположить, что изучение аморфных тел в отличие от кристаллов относится скорее к области искусства, чем к науке. Похоже, однако, что начавшееся в 1960-х годах интенсивное изучение стекловидных твердых тел, таких, как аморфный кремний (интерес к нему вызван его электронными свойствами), привело к созданию основ для получения в большей степени количественной информации как об электронных, так и о структурных свойствах некристаллических материалов.

В фундаментальной теории твердого тела принято начинать с рассмотрения моделей абсолютно чистых монокристаллов бесконечно больших размеров.

Влияние примесей, дефектов, поверхностей и границ зерен учитывается затем как возмущения. Такой подход нередко оказывается плодотворным, даже когда изучаемое твердое тело имеет зерна микроскопического или субмикроскопического размера, при условии что дальний порядок существует на расстоя-

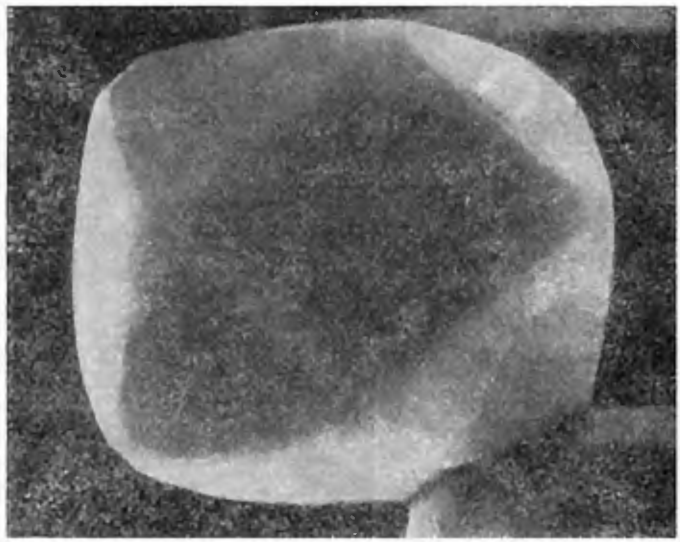


Рис. 1.1. Небольшой кристалл NiO с четко выраженными гранями. Снимок сделан с помощью сканирующего электронного микроскопа. (С любезного разрешения Л. Е. Марра, Орегонский учебный центр). При комнатной температуре антиферромагнитное упорядочение приводит к ромбоэдрическому искажению расположения атомов в NiO (изначально такого же, как у каменной соли).

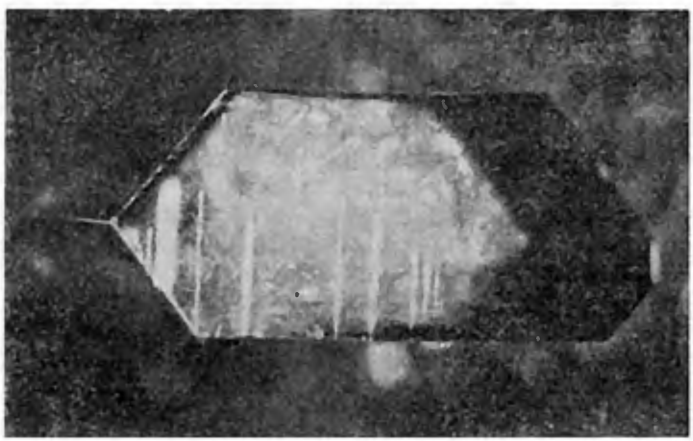


Рис. 1.2. Растущая поверхность кристалла тартрата кальция, выращиваемого в тартратовом геле, в который добавлен раствор хлорида кальция. (Из **книги:** *Henisch H. K.* Crystal Growth in Gels, Penn., State Univ. Press, 1970. [Имеется перевод: *Гениш Г.* Выращивание кристаллов в гелях.— М.: Мир, 1973.])

ниях, которые значительно превышают межатомные. Однако экспериментальные исследования удобнее всего все же проводить на больших монокристаллах, если они доступны, т. е. либо встречаются в природе, либо их получают искусственным путем [1]. На рис. 1.1 и 1.2 показаны примеры микроскопических и макроскопических синтетических кристаллов.

Большие природные кристаллы различных твердых тел были известны человеку тысячелетия назад. Типичными примерами таких кристаллов являются кварц (SiO_2), каменная соль (NaCl), сульфиды металлов, таких, как цинк и свинец, и, несомненно, драгоценные камни, такие, как рубин (Al_2O_3) и алмаз (C). Некоторые из этих природных кристаллов обладают поразительной степенью чистоты и совершенством строения, добиться которого в лабораторных условиях удалось только в последние годы [2]. В течение нескольких столетий слово «кристалл» применялось исключительно по отношению к кварцу; оно имеет греческое происхождение и означает «похожий на лед». В настоящее время *кристаллическим* твердым телом называют твердое тело, у которого расположение атомов периодически повторяется и поверхностные грани которого, если исследуемый образец — *монокристалл,* с большой степенью вероятности располагаются относительно друг друга под вполне определенными углами.

Сравнение двух монокристаллов одного и того же вещества показывает, что, как правило, соотношение между размерами поверхностных граней будет для них различным (т. е. эта характеристика меняется от кристалла к кристаллу). Однако углы между гранями кристаллов данного вещества всегда одни и те же. Это явление было замечено еще в XVI в. и заложило основы изучения кристаллографии в течение последующих трех столетий. Однако свое объяснение эти наблюдения получили лишь с развитием концепции об атомарном строении кристаллов. В 1912 г. Фридрих, Книппинг и Лауэ показали, что кристаллы ведут себя как трехмерная дифракционная решетка для рентгеновских лучей, и тем самым представление о регулярном периодическом расположении атомов в кристалле получило убедительное экспериментальное подтверждение. Сравнительно недавно стало возможным непосредственно наблюдать периодиче-

---

[1] Способы выращивания монокристаллов можно найти в литературе, приведенной в конце гл. 1.

[2] Действительно, искусственно выращенные алмазы до сих пор не могут сравниться по качеству с лучшими природными образцами. Что касается качества большинства других синтетических драгоценных камней, то тут, по-видимому, люди научились делать по крайней мере не хуже, чем природа.

ское расположение атомов с помощью эмиссионной микроскопии [3].

Какие бы свойства кристаллов (механические, тепловые, оптические, электронные или магнитные) мы ни изучали, большая часть полученных результатов независимо от того, изучаем ли мы природные кристаллы, синтетические монокристаллы (такие, как Ge, Si, Al_2O_3, KBr, Cu, Al) или поликристаллические образования, будет тесно связана с периодичностью расположения атомных ядер или со связанной с ней периодичностью электростатического потенциала. Существование периодического потенциала ведет к многочисленным следствиям, причем во многих случаях удается получить точные или приближенные решения. Этим следствиям будет посвящена большая часть данной книги.

В этой главе мы рассмотрим, как атомы связаны друг с другом и как требования симметрии ограничивают многообразие существующих кристаллографических классов. Трудно сказать, в какой последовательности следует рассматривать вопросы о связи между атомами и о симметрии в кристаллах, поскольку они тесно взаимосвязаны; поэтому мы рекомендуем читателю вначале бегло прочитать следующие два раздела, прежде чем приступить к детальному изучению каждого из них.

Затем в данной главе (в разд. 1.3) мы рассмотрим простые кристаллические решетки реальных кристаллов. В этом разделе основное внимание уделено изучению кристаллических структур различных химических элементов и наиболее известных неорганических бинарных соединений.

Разделы 1.4 («Дифракция в кристаллах») и 1.5 («Обратное пространство») тесно связаны друг с другом, и поэтому опять-таки полезно бегло прочитать оба раздела, прежде чем начинать подробное изучение каждого из них. Усвоив понятие обратной решетки, легче понять, что представляет собой дифракция на кристалле, и, наоборот, изучив дифракцию, легче понять, что такое обратная решетка.

Последний раздел этой главы (разд. 1.6) сводится практически к перечислению основных типов точечных и линейных дефектов в кристалле. Приведенные здесь ссылки на имеющуюся литературу предназначены для читателя, который хочет получить дополнительные сведения о дислокациях или о химической термодинамике взаимодействий в твердых телах, обусловленных дефектами.

---

[3] См., например, рис. 1.56, *а*, на котором показано полученное с помощью ионного микроскопа расположение атомов на поверхности кристалла иридия.

1.1. Типы межатомных связей

Все механизмы связи между атомами обусловлены силами электрического притяжения и отталкивания. Типы и силы связи зависят от конкретного строения электронных оболочек взаимодействующих атомов. Между близко расположенными атомами всегда существует небольшое притяжение, обусловленное слабыми вандерваальсовыми (или остаточными) связями, но его влиянием можно пренебречь, если одновременно выполняются условия, обеспечивающие существование ионной, ковалентной или металлической связи.

Наличие стабильной связи (между парой изолированных атомов или между атомами в большом трехмерном кристалле) предполагает, что пространственная конфигурация из положительно заряженных ионных остовов и внешних электронов обладает меньшей полной энергией, чем любая другая конфигурация (в том числе и та, в которой все атомы разнесены на бесконечно большие расстояния). Дефицит энергии этой конфигурации по сравнению с энергией изолированных атомов называется *энергией связи*. Величина этой энергии варьируется от 0,1 эВ/атом в твердых телах со слабым вандерваальсовым взаимодействием до 7 эВ/атом и более в некоторых ковалентных и ионных соединениях, а также в некоторых металлах[4]. Энергия связи равна разности потенциальных энергий системы изолированных атомов и связанной системы *за вычетом* кинетической энергии, которая, согласно принципу неопределенности, должна быть обусловлена локализацией ядер и электронов внешней оболочки.

Для ковалентных связей очень важную роль играют угловые направленности связей, в то время как некоторые другие типы связи главным образом стремятся сделать так, чтобы стало наибольшим координационное число (т. е. число ближайших соседей). Очевидно, что такие факторы имеют большое значение для получения наиболее выгодной трехмерной структуры. У некоторых твердых тел могут быть две и более абсолютно разных структур с практически одинаковой энергией, причем изменением температуры или гидростатического давления можно

[4] Джоуль — довольно большая единица энергии, если речь идет о событиях, касающихся одного атома. Поэтому в нашей книге энергия будет часто приводиться в электронвольтах на частицу или микроскопическую систему. (Из контекста будет ясно, к чему относится изменение энергии: к молекуле, атому или отдельному электрону.) Элементарный заряд, проходя через разность потенциалов в один вольт, приобретает потенциальную энергию $1{,}6022 \cdot 10^{-19}$ Дж (см. в конце книги таблицу наиболее употребительных констант). В химии энергию взаимодействия и энергию связи обычно измеряют в калориях на моль. Один электронвольт на молекулу эквивалентен 23 000 калориям на моль, или $9{,}65 \cdot 10^4$ джоулям на моль.

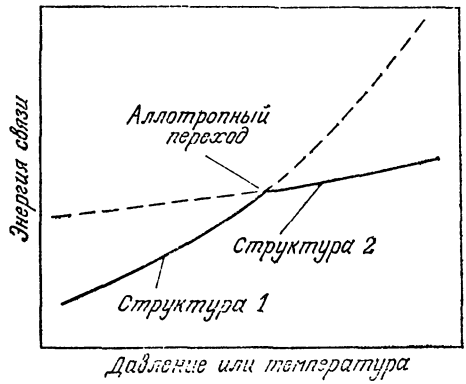


Рис. 1.3. Зависимость энергии связи от температуры или давления для твердого тела, в котором возможны две различные атомные структуры. Аллотропный переход *может* произойти при тех значениях температуры или давления, при которых одна структура заменяет другую как имеющая минимальную энергию.

вызвать превращение одной аллотропной формы твердого тела в другую так, как это представлено на рис. 1.3. Ниже в разделе «Ковалентная связь» будет показано, что при определенном режиме нагревания или охлаждения аллотропный переход в энергетически наиболее выгодную структуру можно задержать.

Вандерваальсово взаимодействие

Вандерваальсово взаимодействие, как уже отмечалось, всегда существует между близко расположенными атомами, но играет важную роль лишь в отсутствие более сильных механизмов связи. Это слабое взаимодействие (с характерной энергией 0,2 эВ/атом) имеет место между нейтральными атомами и между молекулами. Название взаимодействия связывается с именем Ван-дер-Ваальса, поскольку именно он предположил, что уравнение состояния с учетом слабого притяжения между молекулами газа описывает свойства реальных газов много лучше, чем уравнение состояния идеального газа. Однако природа этой универсальной силы притяжения была объяснена лишь в 1930 г., когда данной проблемой заинтересовался Лондон.

Лондон заметил, что, хотя постоянный электрический дипольный момент нейтральных атомов, как и многих молекул, равен нулю, эти атомы и молекулы притягиваются друг к другу электрическими силами. Он указал на то, что нулевые колебания, существующие в силу принципа неопределенности, создают у каждого нейтрального атома флуктуирующий дипольный мо-

мент с быстро меняющейся ориентацией и амплитудой. Поле, индуцируемое диполем, убывает пропорционально кубу расстояния. Таким образом, если ядра двух атомов находятся на расстоянии r друг от друга, диполь одного из атомов в каждый момент времени создает в точке, где находится другой диполь, мгновенное поле, пропорциональное $(1/r)^3$. При этом потенциальная энергия взаимодействия между диполями (приводящая к притяжению) дается выражением

$$E_{\text{притяж}} = -A/r^6. \tag{1.1}$$

Квантовомеханический расчет энергии этого диполь-дипольного взаимодействия показывает, что если бы можно было уменьшить r до 1 Å, то энергия притяжения $E_{\text{притяж}}$ достигла бы 10 эВ. Однако обусловленное перекрытием электронных оболочек отталкивание не позволяет сблизить атомы на столь малое расстояние.

С уменьшением расстояния между атомами электронные облака атомов начинают перекрываться, возникает механизм, который приводит к ослаблению притяжения. Это можно понять с помощью принципа Паули, согласно которому в одном и том же квантовом состоянии не могут находиться два и более электронов. Таким образом, перекрытие электронных облаков двух атомов с почти заполненными электронными оболочками возможно лишь при переходе некоторых электронов в более высокие квантовые состояния, для чего требуется дополнительная энергия.

Зависимость энергии отталкивания от расстояния между атомами можно описать либо степенным законом (для этого необходима столь сильная зависимость, как r^{-11}—r^{-12}), либо с помощью характерной длины. Последний способ обычно оказывается наиболее удовлетворительным, и, воспользовавшись им, полную энергию можно записать в виде

$$E = -\frac{A}{r^6} + B\exp\left(-\frac{r}{\rho}\right). \tag{1.2}$$

Эта зависимость изображена на рис. 1.4 сплошной кривой. Энергия образующейся связи и равновесное расстояние r_0 между связанными атомами определяются параметрами A, B и ρ. Поскольку характерная длина ρ мала по сравнению с межатомным расстоянием, в равновесной конфигурации, отвечающей минимуму энергии E, член, соответствующий отталкиванию, вносит весьма малый вклад в энергию связи[5].

До сих пор мы говорили о вандерваальсовом взаимодействии между парой изолированных от окружения атомов. В трехмер-

[5] См. помещенную в конце главы задачу 1.1 в качестве упражнения, показывающего правильность этого утверждения.

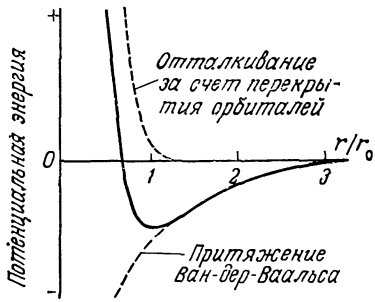


Рис. 1.4. Полная потенциальная энергия взаимодействия Ван-дер-Ваальса (сплошная кривая), полученная при сложении энергий притяжения и отталкивания. Устойчивая связь образуется при межъядерном расстоянии r_0.

ных телах эффекты диполь-дипольного притяжения и отталкивания, обусловленного перекрытием электронных оболочек различных соседних атомов, складываются, так что полная энергия связи по-прежнему определяется выражением (1.2). Поскольку ограничения на направления связей отсутствуют, твердые тела, в которых связи обусловлены силами Ван-дер-Ваальса, стремятся образовать плотноупакованные кристаллические структуры с максимально возможным числом ближайших соседей у каждого атома. [Таковы, например, кристаллы инертных газов Ne, Ar, Kr и Xe. Все они имеют гранецентрированную кубическую (г. ц. к.) решетку, в которой каждый атом окружен двенадцатью ближайшими соседями.] Поскольку вандерваальсово притяжение быстро спадает с расстоянием, все остальные атомы, кроме ближайших соседей, оказывают несущественное влияние.

Кристаллы инертных газов[6] являются прекрасными примерами твердых тел, в которых связи обусловлены *исключительно* силами Ван-дер-Ваальса, поскольку для конфигураций с целиком заполненными электронными оболочками исключается возможность существования других, более сильных механизмов связи. Однако значительно чаще встречаются кристаллы, в которых силы Ван-дер-Ваальса связывают между собой *молекулы* с насыщенными связями, в то время как внутри молекул действуют более сильные механизмы связи. Таковы, например, кристаллы многих насыщенных органических соединений, а также твердые H_2, N_2, O_2, F_2, Cl_2, Br_2 и I_2. На примере Cl_2, энергия возгонки которого составляет 0,2 эВ на молекулу, а энергия диссоциации — 2,5 эВ на молекулу, видно, что связи Ван-дер-Ваальса между двухатомными молекулами разрываются гораздо легче, чем ковалентная связь Cl—Cl.

[6] В гелии из-за сильных нулевых колебаний кристаллизация даже при абсолютном нуле температуры возможна только при приложении внешнего давления 30 атм.

Ковалентная связь

Ковалентная связь, которую иногда называют *валентной* или *гомеополярной*, образуется за счет взаимодействия между двумя электронами в условиях, когда эти электроны обобществлены парой соседних атомов. Как следствие такого обобществления в области между двумя атомами возникает высокая плотность электронного заряда[7]. Каждый атом может образовывать ограниченное число ковалентных связей (в зависимости от того, насколько число внешних электронов отличается от того, которое необходимо для полного заполнения электронной оболочки). Кроме того, имеются явно выраженные преимущественные направления связей. Так, углерод может образовывать четыре связи по направлениям вдоль ребер тетраэдра (под углами 109,5°) и характерная тетрагональная структура наблюдается в кристаллическом алмазе и многочисленных органических соединениях. Другие примеры характерных углов между соседними ковалентными связями — это угол 105° в пластической сере и 102,6° в теллуре.

Наиболее простым веществом с ковалентной связью является молекула водорода (H_2). В любом из двух изолированных атомов водорода электрон находится в состоянии 1*s*. Когда атомы сближаются, взаимодействие между ними расщепляет состояние 1*s* на два состояния с разными энергиями, как показано на рис. 1.5. При очень близком расстоянии двух ядер друг от друга полная энергия каждого из состояний увеличивается вследствие электростатического отталкивания между ядрами. Однако для 1*s*-состояния с четной (симметричной) орбитальной волновой функцией, которое имеет дополнительное обозначение[8] σ_g, на промежуточных расстояниях энергия уменьшается (т. е. проявляется тенденция к притяжению) [9].

---

[7] Напомним, что в квантовой механике мы не можем говорить о движении связанного электрона по определенной орбите, а оперируем лишь с волновой функцией ψ, квадрат которой пропорционален усредненной по времени вероятности обнаружить электрон в данной точке. Если ψ — *нормированная* волновая функция (так что интеграл от ψ^2 по всему пространству равен единице), то средняя плотность заряда в каждой точке равна $-e\psi^2$.

[8] В обозначениях σ_g и σ_u для двух орбитальных волновых функций нижние индексы происходят от немецких слов «gerade» (четный) и «ungerade» (нечетный).

[9] Главной особенностью связи с притяжением является наличие резонансной энергии, соответствующей обмену двумя электронами между двумя атомными орбиталями. Впервые на это свойство обратили внимание Гайтлер и Лондон [*Heitler W., London F.* — Z. Phys., 44, 455 (1927)]. Современную трактовку можно найти в книге Андерсона [*Anderson E. E.* Modern. Physics and Quantum Mechanics, W. B. Saunders, 1971, p. 390].

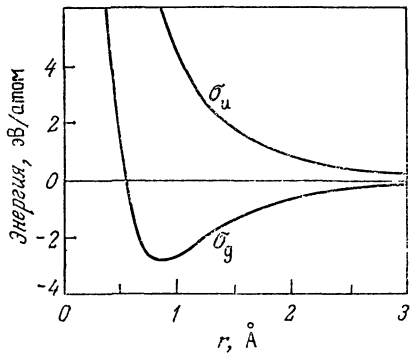


Рис. 1.5. Зависимость энергии молекулы водорода от межъядерного расстояния согласно Гайтлеру и Лондону. Показаны состояния σ_g (связывающее) и σ_u (антисвязывающее); в состоянии σ_g находятся два электрона с антипараллельными спинами.

В симметричном 1*s*-состоянии σ_g плотность электронного заряда $-e\psi^2$ должна быть сконцентрирована в области между двумя ядрами. Принцип Паули, требующий, чтобы полная волновая функция была антисимметричной, выполняется, если в 1*s*-состоянии σ_g находятся два электрона с антипараллельными спинами.

В случае же 1*s*-состояния σ_u два электрона, согласно принципу Паули, должны иметь параллельные спины. Однако из рис. 1.5 видно, что данное состояние является антисвязывающим (ядра отталкиваются) при любых расстояниях между ядрами. Для молекулы H_2 это не важно, поскольку оба ее электрона находятся в состоянии σ_g и образуют сильную связь.

Заметим, что парная связь между двумя атомами гелия существовать не может, поскольку при заполнении как σ_g-, так и σ_a-состояний полная энергия увеличивается. Интересно, что молекулярный ион He_2^+ *является* стабильным.

Квантовомеханическая задача об образовании ковалентных связей значительно усложняется при переходе к многоэлектронным атомам, однако качественно картина остается той же, что и для связи H—H. Во всех случаях сближение атомов ограничивается кулоновским отталкиванием ядер, а в более тяжелых атомах при этом добавляется еще отталкивание за счет перекрытия заполненных внутренних электронных оболочек.

К веществам с ковалентной связью относятся:

1) большинство органических соединений;

2) твердые и жидкие вещества, у которых связи образуются между парами атомов галогенов (а также между парами атомов водорода, азота и кислорода);

3) элементы VI группы (например, спиральные цепочки теллура), элементы V группы (например, деформированные шести-

угольники мышьяка) и элементы IV группы (такие, как алмаз, Si, Ge, α-Sn);

4) соединения, подчиняющиеся правилу 8-*N* (такие, как InSb), когда образующие их элементы расположены не слишком далеко друг от друга по горизонтали в периодической таблице Менделеева.

Нередко в процессе кристаллизации твердых тел с ковалентной связью могут образовываться различные кристаллические структуры, энергия связи которых практически одинакова. Изменением температуры или давления можно заменить первоначальную структуру на энергетически наиболее выгодную (рис. 1.3); эта способность некоторых веществ известна как *аллотропия* или *полиморфизм*. Например, структура ZnS может быть кубической (цинковая обманка) или гексагональной (вюрцит). Расположение ближайших соседей в цинковой обманке и вюрците одинаково, а небольшое различие в энергиях этих двух структур определяется расположением атомов, следующих за ближайшими. Аналогично карбид кремния имеет целый ряд «политипов» с различной структурой — от чисто кубической до чисто гексагональной со слабо различающимися электронными свойствами.

Все многочисленные кристаллические модификации ZnS и SiC могут существовать при комнатной температуре без особого риска спонтанного перехода в энергетически наиболее выгодное состояние (условия, при которых образуются эти различные кристаллические формы, можно определить при низких температурах). В других веществах спонтанные переходы происходят довольно легко.

Например, при быстром охлаждении от точки плавления (218 °С) до комнатной температуры селен оказывается аморфным, однако нагревание до температуры 60—70 °С приводит к его кристаллизации, и при последующем охлаждении до комнатной температуры он остается кристаллическим. Другим хорошим примером вещества, способного к аллотропному превращению, является олово. При температурах ниже 17 °С стабильное состояние олова — это серый полуметалл (α-Sn) с тетраэдрической структурой алмаза, а при температурах выше 17 °С и, если приложено давление, даже ниже этой температуры оно переходит в гораздо более плотное металлическое состояние белого олова (β-Sn) с тетрагональной решеткой, в которой каждый атом имеет шесть ближайших соседей.

Структуры со смешанными ковалентными и вандерваальсовыми связями

Как уже отмечалось выше, сочетание ковалентного и вандерваальсова механизмов связи встречается в веществах, таких, как твердый водород, у которых за счет ковалентного взаимодействия каждая пара атомов связывается в молекулу, а вандерваальсовы связи создают «молекулярный кристалл». Этот же принцип справедлив и для большинства органических кристаллов.

Иное расположение ковалентных связей демонстрирует теллур (рис. 1.6), у которого в каждой спиральной цепочке соседние атомы связаны между собой ковалентными связями, а силы, действующие *между* цепочками, намного слабее, хотя, по-видимому, все же несколько больше, чем силы притяжения Ван-дер-

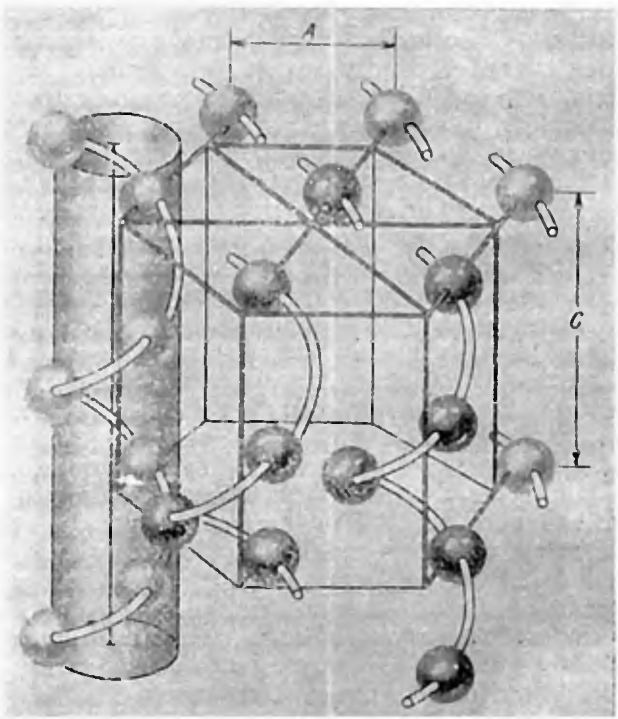


Рис. 1.6. Атомное строение теллура [*Blakemore J. et al.*— In: Progress in Semiconductors, v. 6, Wiley, 1962]. Каждый атом образует ковалентные связи с ближайшими соседями сверху и снизу в спиральной цепочке. Связи между цепочками слабые. Такая же структура у одной из аллотропных форм селена.

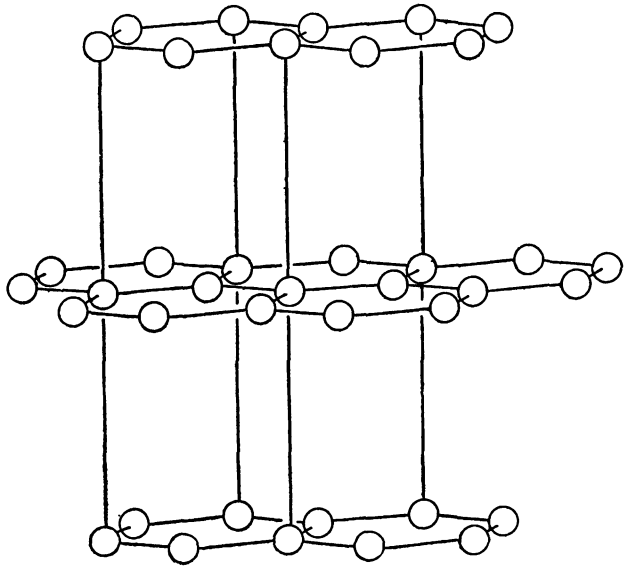


Рис. 1.7. Атомное строение графита. В пределах одного слоя каждый атом образует три сильные ковалентные связи (r_0=1,42 Å), чтобы сохранялась гексагональная структура. Связи между слоями (расстояние 3,4 Å) слабые, и, следовательно, слои могут легко скользить относительно друг друга.

Ваальса. Из-за этого у теллура низкая прочность структуры, а его механические, тепловые и электронные свойства анизотропны.

Аналогично в графите (рис. 1.7) атомы углерода в каждом слое расположены в вершинах правильных шестиугольников, так что три из четырех электронов внешней оболочки каждого атома участвуют в образовании ковалентных связей внутри слоя (четвертый электрон свободен). Расстояние между слоями велико, поскольку притяжение между ними осуществляется только за счет вандерваальсова взаимодействия. Таким образом, слои могут легко скользить относительно друг друга, что позволяет использовать графит как «твердую смазку». То же самое справедливо и для MoS_2.

Ионная связь

Ионный кристалл состоит из положительных и отрицательных ионов, расположенных так, что кулоновское отталкивание одноименно заряженных ионов оказывается меньше кулоновского притяжения ионов с противоположным зарядом. Типич-

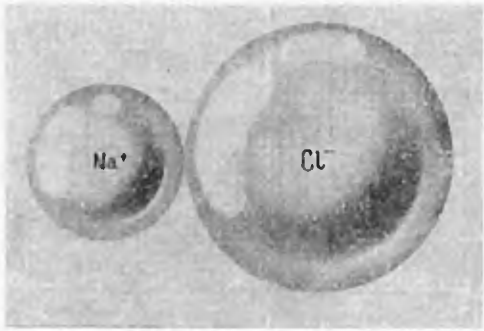


Рис. 1.8. Соприкосновение аниона Cl^- и катиона Na^+, изображенных в соответствии с наблюдаемыми относительными их размерами. В ионном кристалле относительные размеры играют важную роль при определении наиболее выгодной структуры.

ными представителями класса ионных кристаллов являются галогениды щелочных металлов, такие, как NaCl, который кристаллизуется (почти) в виде Na^+Cl^-. Поскольку потенциал ионизации I_e щелочного металла невелик, электрон от Na практически переходит к Cl (для превращения атома Na в катион Na^+ с конфигурацией заполненной электронной оболочки необходимо совершить работу eI_e), в то время как галоген имеет большое сродство E_a к электрону. (Энергия E_a выделяется, когда атом Cl получает электрон и становится анионом Cl^- также с заполненной электронной оболочкой.) Баланс энергий, связанный с возникновением ионной связи в отдельной молекуле галогенида щелочного металла, рассмотрен в задаче 1.2.

Когда ионы Na^+ и Cl^- приближены друг к другу в отсутствие каких-либо других атомов (как это изображено на рис. 1.8), энергия их кулоновского притяжения при межъядерном расстоянии r, отсчитанная от нулевого значения энергии при удалении их друг от друга на бесконечное расстояние, равна

$$E_{\text{кул}} = -e^2/4\pi\varepsilon_0 r, \qquad (1.3)$$

поскольку распределение электронного заряда в заполненной электронной оболочке сферически-симметрично. Благодаря отталкиванию, возникающему вследствие принципа Паули при перекрытии заполненных оболочек аниона и катиона, расстояние, на которое могут сблизиться ионы, ограниченно. Как уже отмечалось при рассмотрении вандерваальсовых взаимодействий, энергия отталкивания быстро изменяется с межъядер-

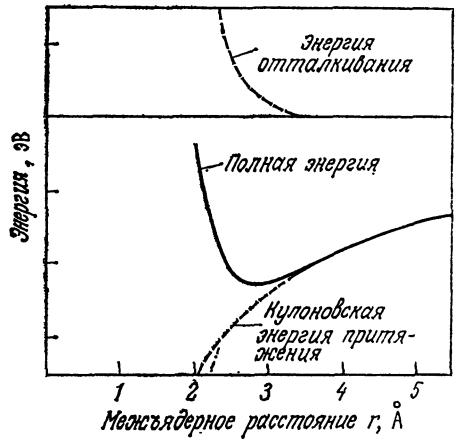


Рис. 1.9. Энергия молекулы Na^+Cl^- по сравнению с энергией отдельных ионов в соответствии с расчетами по формуле (1.6). Характерная длина для энергии отталкивания предполагается равной $\rho=0{,}345$ Å, при этом минимум полной энергии приходится на расстояние $r_0=2{,}82$ Å.

ным расстоянием, и ее можно записать в виде следующих двух приближенных выражений:

$$E_{\text{отт}} = A/r^n \quad (n \approx 12) \tag{1.4}$$

или

$$E_{\text{отт}} = B \exp(-r/\rho). \tag{1.5}$$

Ни одно из этих выражений не соответствует сложному квантовомеханическому процессу, который приводит к отталкиванию.

Из выражения (1.5), дающего представление об изменении энергии отталкивания с межъядерным расстоянием, очевидно, что устойчивой будет связь с таким расстоянием между ионами Na^+ и Cl^-, при котором энергия

$$E_i = E_{\text{кул}} + E_{\text{отт}} = -e^2/4\pi\varepsilon_0 r + B \exp(-r/\rho) \tag{1.6}$$

минимальна. Этот минимум энергии показан на рис. 1.9. Поскольку в формуле (1.6) член, соответствующий отталкиванию, изменяется с r значительно более существенно, чем кулоновский, энергия связи оказывается лишь немного меньше $e^2/4\pi\varepsilon_0 r_0$, в то время как возвращающие силы, действующие на атом при r, отличном от r_0, определяются величинами B и ρ.

Отмеченные выше правила образования связи между отдельными ионами Na^+Cl^- при учете некоторых дополнительных

геометрических соображений оказываются вполне справедливы[10] и для кристалла NaCl. В разд. 1.3 мы вернемся к вопросу о структуре хлорида натрия и рассмотрим его с точки зрения геометрических свойств и симметрии [используя для этого рис. 1.33, *а,* а пока исследуем все четыре схемы на рис. 1.10, чтобы понять, как возникает конкретное пространственное расположение связанных ионов. В кристалле NaCl каждый катион (т. е. каждый ион натрия) имеет в качестве ближайших соседей шесть анионов (и наоборот, как можно видеть из рис. 1.10, *б*)]. Взаимодействие с ближайшими соседями состоит из кулоновского притяжения и отталкивания, обусловленного перекрытием электронных оболочек. На рис. 1.10, *в* и *г*, на которых катионы и анионы изображены в соответствии с их размерами, видно, что ни катионы между собой, ни даже огромные анионы между собой не соприкасаются. Таким образом, взаимодействие иона со своими соседями, кроме ближайших, является исключительно кулоновским. Поскольку у кристалла энергия связи должна быть положительной, сумма всех слагаемых (как положительных, так и отрицательных), обусловленных кулоновским взаимодействием, должна с избытком компенсировать энергию отталкивания, обусловленную перекрытием волновых функций данного атома и его шести ближайших соседей. (Положительность энергии связи означает, что энергия E_i *меньше,* чем у бесконечно удаленных ионов). Несложный расчет показывает (см. задачу 1.2), что энергия, необходимая для разделения кристалла с ионной связью на отдельные ионы, больше, чем энергия, необходимая для разделения его на изолированные нейтральные атомы.

Энергетически наиболее выгодная структура кристалла определяется соотношением между радиусами аниона и катиона. (Напомним, что анион Cl^- значительно больше катиона Na^+.) В случае когда это соотношение радиусов допускает контакт аниона с анионом, структура, по-видимому, заменится на другую ионную структуру. Таким образом, структура кристалла NaCl предпочтительнее структуры CsCl, только если $r_-/r_+ \geqslant 1{,}41$, а когда отношение r_-/r_+ становится очень большим, энергетически наиболее выгодной оказывается структура цинковой обманки. Подобные соображения (существование большого числа контактов анионов с катионами при отсутствии анион-анионных контактов) определяют выбор наиболее предпочтительной структуры таких ионных кристаллов, как галогениды щелочно-

[10] В следующем разделе будет показано, что электрон не полностью переходит от одного атома к другому даже в химических соединениях с наибольшей ионностью связи. Однако рассмотрение галогенидов щелочных материалов как чисто ионных соединений вполне оправданно.

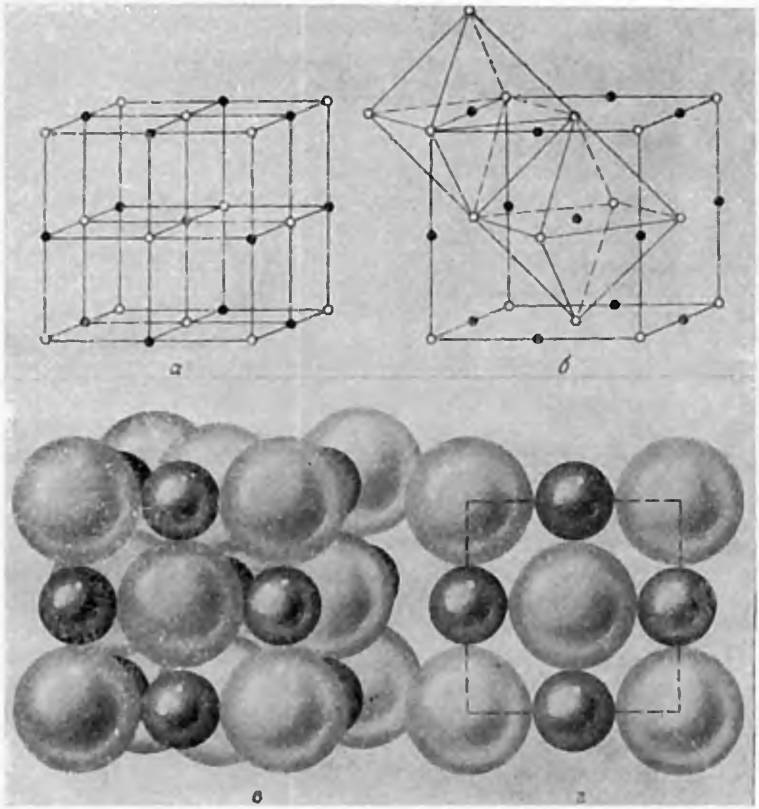


Рис. 1.10. Четыре способа изображения кристаллической решетки NaCl. Черными кружками изображены катионы, белыми — анионы [*Evans R. C.* Introduction to Crystal Chemistry, Cambridge Univ. Press, 1964]. *а* — условное изображение, показывающее расположение кристаллографических плоскостей; *б* — октаэдрическая конфигурация, в которой каждый ион имеет шесть ближайших соседей; *в* — элементарный куб, в котором ионы изображены в соответствии с их размерами, так что катионы и анионы касаются друг друга; *г* — сечение вдоль грани куба, которое показывает, что анион-анионный и катион-катионный контакты *отсутствуют*.

земельных элементов, в которых отношение радиусов анионов и катионов не равно единице.

Возвращаясь к кристаллу NaCl как к наиболее типичному примеру кристалла с ионной связью, заметим, что катион натрия окружен шестью ближайшими соседями Cl^- на расстоянии r_0, двенадцатью следующими за ближайшими соседями анионами Na^+ на расстоянии $\sqrt{2}r_0$, восемью следующими за ними катионами Cl^- на расстоянии $\sqrt{3}r_0$ и т. д. Таким образом, полная

энергия кулоновского притяжения, приходящаяся на *одну ионную пару,* равна сумме бесконечного ряда

$$E_{\text{кул}} = -(e^2/4\pi\varepsilon_0 r_0)\left\{6 - \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}} - \ldots\right\} =$$
$$= -1{,}748\,(e^2/4\pi\varepsilon_0 r_0) = -\alpha\,(e^2/4\pi\varepsilon_0 r_0). \qquad (1.7)$$

Множитель α называется *постоянной Маделунга* (Маделунг, 1918), и его величина для данной решетки определяется ее геометрией. (Для сравнения с решеткой NaCl заметим, что решетка цинковой обманки имеет $\alpha = 1{,}638$, а структура типа CsCl—$\alpha = 1{,}763$.)

Определение постоянной Маделунга с помощью выражения (1.7) требует суммирования бесконечного ряда, который является знакопеременным и сходится, очевидно, довольно медленно. Такая медленная сходимость этого ряда для большинства кристаллических структур представляет собой довольно сложную проблему. Методы точного суммирования таких рядов были разработаны Эвальдом (1921) и Эвьеном [11]. В основе этих методов лежит разбиение пространства вокруг одного иона на зоны, заключенные между последовательными многогранниками. Поверхности многогранников выбираются таким образом, чтобы полный заряд каждой зоны был равен нулю, а заряд иона, находящегося на границе между двумя зонами, равномерно распределялся между ними. Метод Эвьена позволяет построить новый ряд, который теперь уже оказывается быстро сходящимся. Этот метод применяется в задаче 1.3, но в более упрощенном виде для случая двух измерений.

Из выражений (1.5) и (1.7) находим, что полная энергия образования одной молекулы ионного кристалла из ионов, удаленных друг от друга на бесконечное расстояние, записывается в виде

$$E_i = -\frac{\alpha e^2}{4\pi\varepsilon_0 r} + C\exp\left(-\frac{r}{\rho}\right), \qquad (1.8)$$

где значения C и ρ могут быть получены только с помощью полного квантовомеханического расчета. Однако одну из этих констант можно исключить, если учесть тот факт, что при равновесном значении расстояния r_0 между ближайшими соседями энергия имеет минимум. Дифференцируя выражение (1.8), находим

$$\frac{dE_i}{dr} = \frac{\alpha e^2}{4\pi\varepsilon_0 r^2} - \frac{C}{\rho}\exp\left(-\frac{r}{\rho}\right),$$

а из условия $(dE/dr)_{r_0} = 0$ следует, что

$$C = \frac{\alpha\rho e^2}{4\pi\varepsilon_0 r_0^2}\exp\left(\frac{r_0}{\rho}\right). \qquad (1.9)$$

[11] *Evjen H. M.*— Phys. Rev., 39, 680 (1932).

Таким образом, при любом расстоянии r между ионами энергия связи равна

$$E_i = -\frac{\alpha e^2}{4\pi\varepsilon_0 r}\left[1-\left(\frac{r\rho}{r_0^2}\right)\exp\left(\frac{r_0-r}{\rho}\right)\right]. \qquad (1.10)$$

При равновесном же значении расстояния r мы имеем

$$E_{\text{равн}} = -\frac{\alpha e^2}{4\pi\varepsilon_0 r_0}\left[1-\left(\frac{\rho}{r_0}\right)\right], \qquad (1.11)$$

а поскольку константа ρ по сравнению с r_0 невелика (составляет всего несколько процентов r_0), энергия связи определяется главным образом энергией Маделунга.

Оставшуюся неизвестную константу ρ можно исключить из выражений (1.10) и (1.11), если воспользоваться имеющимися экспериментальными данными об объемной сжимаемости χ. Такая возможность возникает благодаря тому, что сжимаемость определяется второй производной энергии по расстоянию при равновесном его значении, на которое в значительной степени влияет характерная длина потенциала близкодействующих сил отталкивания. Таким образом, рассмотрим бесконечно малое изменение объема кристалла dv (на одну молекулу) при давлении p. В структуре хлорида натрия объем одной пары ионов $v=2r^3$, так что его изменение $dv=6r^2dr$. Совершаемая при этом изменении работа равна $dE=-pdv=-6r^2dr$, откуда мы находим следующее выражение для давления:

$$p = -\frac{1}{6r^2}\frac{dE}{dr},$$

производная которого запишется в виде

$$\frac{dp}{dr} = -\frac{1}{6r^2}\frac{d^2E}{dr^2}+\frac{1}{3r^3}\frac{dE}{dr}. \qquad (1.12)$$

Напомним, что $(dE/dr)_{r_0}=0$, т. е. при равновесном расстоянии между ионами в правой части выражения (1.12) второе слагаемое обращается в нуль.

Сжимаемость χ связана с зависимостью давления от объема следующим образом:

$$\chi = -\frac{1}{v}\frac{dv}{dp} = -\frac{3}{r}\frac{dr}{dp}. \qquad (1.13)$$

Отсюда мы можем выразить dp/dr через χ и полученное значение сравнить с (1.12). Таким образом, при равновесном расстоянии r_0 мы имеем

$$(d^2E/dr^2)_{r_0} = 18r_0/\chi. \qquad (1.14)$$

Можно получить и другое выражение для $(d^2E/dr^2)_0$ в виде функции от характерной длины ρ, если дважды продифферен-

цировать выражение (1.10). Сравнивая результат с формулой (1.14), находим

$$\rho/r_0 = [2 + (72\pi\varepsilon_0 r_0^4/\alpha e^2 \chi)]^{-1}. \qquad (1.15)$$

Таким образом, энергию связи можно вычислить, если известны равновесное значение периода решетки, постоянная Маделунга и сжимаемость, что и используется при решении задачи 1.4.

Данные о свойствах ионных кристаллов собраны в работах целого ряда авторов [12]. В частности, Мотт и Герни сообщают, что для всех двадцати кристаллов галогенидов щелочных металлов (будь то структура типа NaCl или CsCl) вполне удовлетворительным является значение $\rho = 0{,}345$ Å, но что для различных соединений значение множителя C в выражении для энергии отталкивания $C \exp(-r/\rho)$ меняется в широких пределах. Это обусловлено тем, что положительные ионы щелочных металлов и отрицательные ионы галогенов с полностью заполненными электронными оболочками ведут себя по существу как несжимаемые сферы, касающиеся друг друга при равновесных расстояниях. В табл. 1.1 приведены значения радиусов различных ионов, которые Мотт и Герни принимают за «основные». При этом если энергию отталкивания записать в виде

$$C \exp(-r/\rho) = C' \exp[(r_a + r_c - r)/\rho], \qquad (1.16)$$

то значения C' будут примерно одни и те же для всех двадцати бинарных соединений, что указывает на почти одинаковую деформируемость всех сферических ионов.

Полуэмпирическая шкала Полинга [13] дает несколько большие значения радиусов ионов щелочных металлов и галогенов, поскольку эта шкала устроена так, что сумма радиусов ионов ближайших соседей (в чисто ионном соединении с координационным числом 6) должна равняться равновесному межъядерному расстоянию.

То, насколько удовлетворительными являются такие полуэмпирические схемы, определяется тем, насколько хорошо связь позволяет обеспечить полный переход электрона от катиона к аниону. Здесь следует заметить, что *полный* переход электрона не реализуется *ни в одном* из кристаллов.

---

[12] *Mott N. F., Gurney R. W.* Electronic Processes in Ionic Crystals, Oxford, 2nd ed., 1948 [имеется перевод: *Мотт Н. Ф., Герни Р. В.* Электронные процессы в ионных кристаллах.— М.: ИЛ, 1950], а также *Tosi M. P.*— In: Solid State Physics, v. 16, Academic Press, 1964.

[13] *Pauling L.* The Nature of the Chemical Bond, Cornell Univ. Press, 3rd ed., 1960. [Имеется перевод 1-го издания: *Полинг Л.* Природа химической связи.— М.: ИЛ, 1947.]

Таблица 1.1. **Радиусы ионов в галогенидах щелочных металлов**

Li*+	$r_c = 0{,}475$ Å	F^-	$r_a = 1{,}110$ Å
Na+	$r_c = 0{,}875$ Å	Cl^-	$r_a = 1{,}475$ Å
K+	$r_c = 1{,}185$ Å	Br^-	$r_a = 1{,}600$ Å
Rb+	$r_c = 1{,}320$ Å	I^-	$r_a = 1{,}785$ Å
Cs+	$r_c = 1{,}455$ Å		

Смешанные ионно-ковалентные связи

Для осуществления чисто ионной связи в соединении необходимо наличие двух компонент: существенно электроположительной компоненты (которая может быть легко ионизована, чтобы образовать катион) и существенно электроотрицательной компоненты (имеющей по возможности большое сродство к электрону, чтобы образовать анион). Этим требованиям довольно хорошо удовлетворяют галогениды щелочных металлов, в которых существуют все условия для перехода электрона от катиона к аниону.

Однако в соединениях элементов с не столь явно выраженными электроположительными и электроотрицательными свойствами степень перехода заряда от катиона к аниону значительно меньше 100 %. Например, благородные металлы имеют бóльшую энергию ионизации, чем щелочные, и поэтому связь в галогенидах серебра в меньшей степени относится к ионной, чем в соответствующих галогенидах щелочных металлов. В действительности если рассматривать соединения элементов, у которых разность электроотрицательностей становится все меньше и меньше, то мы будем наблюдать непрерывный переход от чисто ионных связей к чисто ковалентным.

При наличии частичного обобществления электронов оптимальную связь можно рассматривать как результат *резонанса* между ионной и ковалентной конфигурациями зарядов. При этом усредненная по времени волновая функция электрона, участвующего в образовании такой связи, запишется в виде

$$\psi = \psi_{\text{ков}} + \lambda\psi_{\text{ион}}, \tag{1.17}$$

где $\psi_{\text{ков}}$ и $\psi_{\text{ион}}$ — нормированные волновые функции для чисто ковалентной и чисто ионной связей соответственно, а λ — параметр, определяющий степень ионности:

$$\text{Ионность, \%} = \frac{100\lambda^2}{1+\lambda^2}. \tag{1.18}$$

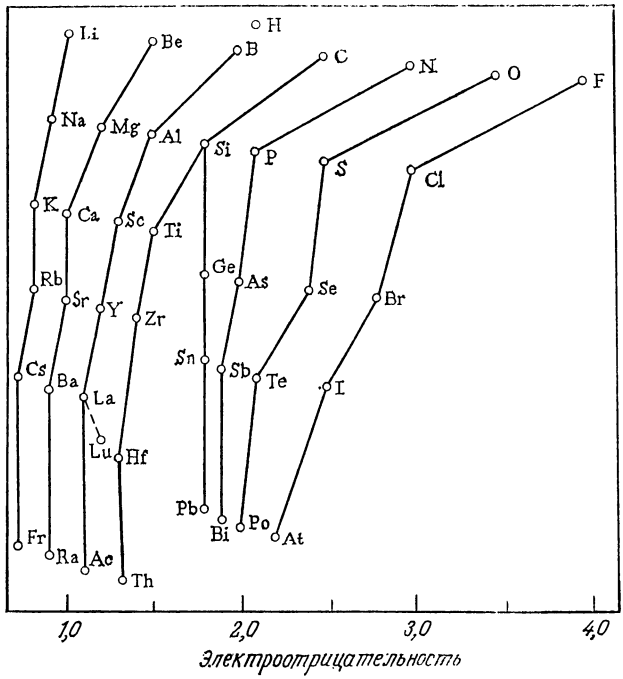


Рис. 1.11. Коэффициенты электроотрицательности некоторых элементов по Полингу (1960). Масштаб по оси абсцисс выбран таким образом, чтобы для резонансной энергии двухкомпонентных соединений выполнялось уравнение (1.21).

Значение λ определяется из квантовомеханических расчетов с применением вариационных методов, основывающихся на предположении, что наиболее стабильной (и, следовательно, равновесной) конфигурации соответствует такое значение λ, для которого энергия[14]

$$E = \frac{\int \psi \mathcal{H} \psi d\tau}{\int \psi^2 d\tau} \qquad (1.19)$$

проходит через отрицательный минимум.

Смешанный характер связи приводит к появлению добавки к энергии связи:

$$\Delta = E_{\text{ков}} - E, \qquad (1.20$$

которую нередко называют ионно-ковалентной резонансной энергией. Для того чтобы определить Δ, необходимо вычислить

[14] Выражение (1.19) определяет среднюю энергию системы с гамильтонианом $\mathcal{H}$ и волновой функцией ψ. Квантовомеханический расчет совсем не нужен для того, чтобы понять, что резонанс повышает прочность связи.

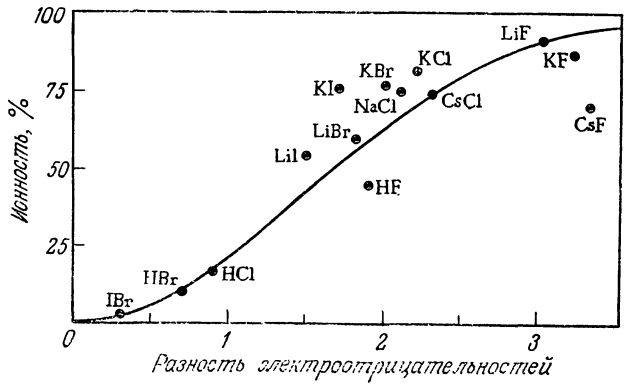


Рис. 1.12. Кривая следует уравнению (1.22), предложенному Полингом (1960). Точками отмечены значения ионности, определенные из дипольных моментов. Соответствующие разности электроотрицательностей взяты со шкалы на рис. 1.11.

значение энергии $E_{\text{ков}}$ для гипотетического случая чисто ковалентной связи между атомами данного соединения. Величину $E_{\text{ков}}$ для атомов A и B обычно определяют как среднее геометрическое энергий связи $A — A$ и $B — B$.

Для полуколичественных оценок величин λ и Δ (напомним, что Δ увеличивается с ростом λ) Полинг[13] на основе эмпирических данных составил шкалу коэффициентов электроотрицательности различных элементов. Эта шкала представлена на рис. 1.11. Значения x выбираются таким образом, чтобы удовлетворялось соотношение

$$1{,}3(x_A - x_B)^2 = \Delta_{AB}, \qquad (1.21)$$

где Δ_{AB} измеряется в электронвольтах. Выбор Полингом подобных единиц измерения электроотрицательности продиктован еще и тем, что в этом случае разность $x_A - x_B$ численно равна дипольному моменту связи, выраженному в дебаях [1 дебай $= 10^{-18}$ ед. СГСЭ $\cdot$ см $= 3{,}33 \cdot 10^{-30}$ Кл $\cdot$ м].

Ионность определяется выражениями (1.17) и (1.18) через величину λ. Согласно Полингу, ионность в процентах (определяемая из дипольных моментов) с хорошей точностью связана с разностью коэффициентов электроотрицательности следующим соотношением:

$$\text{Ионность, \%} = 100\left\{1 - \exp\left[-\left(\frac{x_A - x_B}{2}\right)^2\right]\right\}. \qquad (1.22)$$

На рис. 1.12 изображена определяемая этим выражением зависимость, которая сравнивается также со значениями ионности (полученными из дипольных моментов) для различных связей.

Малые поправки для ионных и частично полярных кристаллов

В дополнение к преобладающим механизмам связи, разумеется, необходимо учитывать вандерваальсово взаимодействие. Однако его вклад в энергию связи частично или полностью полярных кристаллов редко превышает 2—3 %.

Кроме того, необходимо учесть и поправки к энергии решетки за счет нулевых колебаний, т. е. за счет того, что при локализации атомы (вследствие принципа неопределенности) приобретают кинетическую энергию. Эта энергия, которую нужно вычесть из энергии притяжения, равна [15] $9\hbar\omega_D/4$ и, как правило, не превышает 0,03 эВ/атом, т. е. представляет собой относительно малую поправку к энергии связи.

Водородная связь

Атом водорода, имеющий лишь один электрон, может образовать ковалентную связь только с одним атомом. Однако он может участвовать также в образовании дополнительной электростатической связи со вторым атомом, обладающим сильной электроотрицательностью. Таким атомом может быть атом фтора, кислорода и в меньшей степени азота. Эта дополнительная связь называется *водородной связью*. Она осуществляется между двумя атомами или структурами. Энергия одной водородной связи может изменяться в пределах 0,1—0,5 эВ.

Водородная связь обеспечивает сцепление между молекулами H_2O (рис. 1.13) в кристаллах обыкновенного льда, которые имеют структуру типа вюрцита, причем расстояние между атомами кислорода соседних молекул составляет 2,76 Å. Это расстояние более чем в два раза превышает межъядерное расстояние O—H (0,96Å) в изолированной молекуле воды. Как показали нейтронографические исследования при образовании водородной связи, когда атом водорода оттягивается к соседней молекуле, связь O—H удлиняется всего до 1,01 Å. Таким образом, лед состоит из *почти* нормальных молекул воды. Существует много различных вариантов расположения молекул воды в кристаллах льда; равновесная конфигурация определяется температурой и давлением; при высоких давлениях известны многочисленные аллотропные модификации льда.

При охлаждении до очень низкой температуры лед переходит в одну из многочисленных модификаций, которая может быть и не самой упорядоченной. В конечном счете кристалл льда сохраняет остаточную энтропию. Для N молекул число

---

[15] Здесь ω_D — максимальная частота в дебаевском спектре колебаний. Подробно этот вопрос будет рассмотрен в гл. 2.

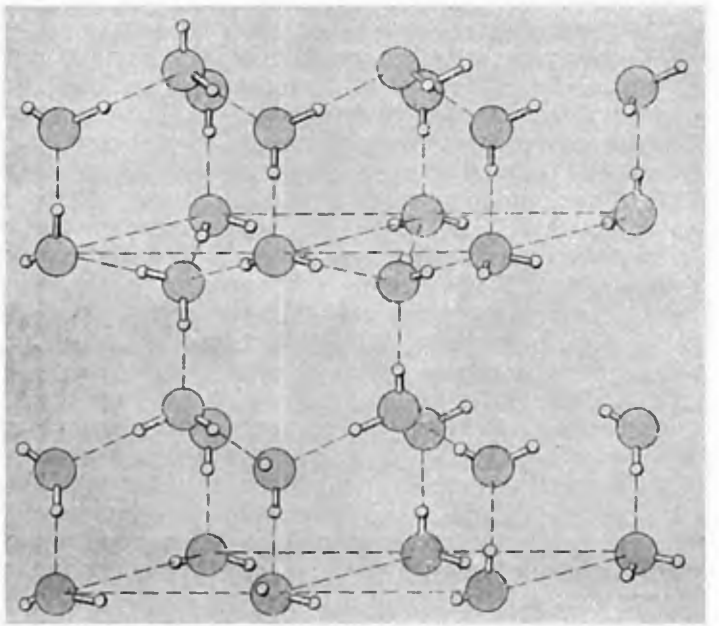


Рис. 1.13. В кристаллической структуре льда каждый атом кислорода (изображенный большим кружком) образует четыре водородные связи, направленные вдоль ребер тетраэдра. Расстояние O—O равно 2,76 Å. Каждый протон (маленький кружок) удален от атома кислорода своей молекулы на 1,01 Å и от ядра атома кислорода соседней молекулы на 1,75 Å. Для сравнения укажем, что в изолированной молекуле воды расстояние O—H равно 0,96 Å [*Pauling L.* The Nature of the Chemical Bond, Cornell University Press, 1960].

возможных конфигураций равно $(^3/_2)^N$, и поэтому молярная энтропия должна составлять примерно $k_0 N_A \ln(^3/_2) \approx$ $\approx 3{,}37$ Дж/(моль·К). Этот результат хорошо согласуется с экспериментально измеренным значением остаточной энтропии 3,42 Дж/(моль·К).

Существует множество подтверждений того, что часть водородных связей продолжает существовать и в воде. Наличием водородных связей объясняется необыкновенно высокая для этого соединения с молекулярной массой 18 температура кипения и теплота испарения. Поразительные диэлектрические свойства воды и льда также обусловлены существованием водородных связей. Отклик системы на внешнее низкочастотное электрическое поле проявляется главным образом в том, что полярные молекулы H_2O начинают вращаться. При этом два

протона перескакивают на другие позиции и образуют новые водородные связи. В воде это вращение возникает довольно легко и обусловливает диэлектрическую постоянную при низких частотах $\varkappa \approx 80$. Значение этой постоянной для льда меньше и существенно зависит от температуры, поскольку вращение молекул льда при низких температурах замедляется.

Водородные связи важны не только для взаимодействия молекул H_2O. Они играют важную роль и в полимеризации таких соединений, как HF, HCN и NH_4F. Водородные связи ответственны за сегнетоэлектрические свойства твердых тел, таких, как KH_2PO_4, и в гл. 5 мы увидим, что температура сегнетоэлектрического перехода существенно меняется при замене обыкновенного водорода дейтерием.

Учет водородных связей имеет большое значение для понимания свойств многих органических соединений и важных биологических веществ. В качестве очень простого примера, приведем соединение C_2H_6O, которое может существовать в виде двух хорошо известных изомеров: диметилового эфира $(CH_3)_2O$ и этилового спирта CH_3CH_2OH. В последнем водородные связи образуются между водородом одной гидроксильной группы и кислородом соседней молекулы, но конфигурация молекулы эфира препятствует образованию водородной связи. Вследствие этого точка кипения и теплота испарения этилового спирта намного выше, чем у диметилового эфира.

В биологических веществах наиболее яркие водородные связи — это те, что связывают между собой части макромолекул в белках и нуклеиновых кислотах. Полинг и Кори[16] показали, что белок имеет спиральную структуру (α-спираль), жесткость которой обусловлена водородными связями между близлежащими полипептидными группами. Полинг показал, что в каждой группе C=O...H—N (водородная связь между кислородом и водородом обозначена точками) четыре атома находятся почти на одной прямой. Это «основа», благодаря которой вся структура представляет собой единую жесткую конфигурацию. Еще более существенной оказалась роль водородных связей в предложенной Криком и Уотсоном[17] модели двойной спирали для дезоксирибонуклеиновой кислоты (ДНК). Здесь связь

[16] *Pauling L., Corey R. B., Branson H. R.*— Proc. Nath. Acad. Sci., U. S., 37, 205 (1951). См. также: *Corey R. B., Pauling L.*— Proc. Roy. Soc., B141, 10 (1953).

[17] Первое сообщение об этом содержится в статье: *Watson J. D., Crick F. H. C.*— Nature, 171, 737 (1953). Интересно и оригинально эта проблема изложена в книге: *Watson J. D.* The Double Helix, Athaneum, 1968. [Имеется перевод: *Уотсон Дж. Д.* Двойная спираль.— М.: Мир, 1969.] Обзор современного состояния теории нуклеиновых кислот дан Спенсером. [*Spenser J. H.* The Physics and Chemistry of DNA and RNA, W. B. Saunders, 1972.]

между двумя антипараллельными цепочками нуклеотидов обеспечивается *исключительно* водородными связями между отдельными парами азотистых оснований (аденином и тимином, гуанином и цитозином). Ряд водородных связей может «расцепиться» при копировании генов во время деления клетки, однако «сцепленная» полная молекула ДНК имеет конечную жесткость и кристаллизуется в виде структур, которые дают четкую картину дифракции рентгеновских лучей.

Металлическая связь

Металлические структуры характеризуются тем, что они имеют довольно редкое расположение атомов (межъядерные расстояния большие) и большое число ближайших соседей у каждого атома в кристаллической решетке. В табл. 1.2 указаны три типичные металлические структуры.

Каждый атом участвует в образовании столь большого числа связей, что все они не могут быть одновременно локализованы, поэтому связь должна осуществляться за счет резонанса внешних электронов каждого атома для всех возможных мод колебаний. Во многих металлах для образования такой связи достаточно взять по одному электрону у каждого атома. Именно это происходит в литии, у которого на внешней оболочке (следующей за заполненной K-оболочкой) имеется всего один электрон. Кристалл лития следует рассматривать как решетку ионов Li^+ (шаров радиусом 0,68 Å), окруженных *электронным газом,* плотность которого составляет один электрон на атом.

Слабость каждой отдельной связи проявляется в том, что межъядерные расстояния в металлах больше, чем в двухатомных молекулах. Так, согласно табл. 1.2, расстояние Li—Li в кристалле лития составляет 3,04 Å, в то время как в молекуле Li_2 с ковалентной связью оно равно всего 2,67 Å. Однако полная энергия связи в металлических кристаллах больше, чем

Таблица 1.2. **Некоторые характеристики трех наиболее распространенных структур металлов**

Структура	Металл	Число ближайших соседей	r_0
О. ц. к.	Литий	8 атомов на расстоянии r_0 6 атомов на расстоянии $2r_0/\sqrt{3}$	3,04 Å
Г. ц. к. (к. п. у.)	Медь	12 атомов на расстоянии r_0	2,56 Å
Г. п. у.	Цинк	12 атомов на расстоянии r_0	2,66 Å

Таблица 1.3. **Краткая классификация типов связи в твердых телах**

Тип связи	Типичные примеры				Некоторые характерные свойства
	Вещество	Кристалическая структура	Энергия связи, эВ/молекула	Расстояние между ближайшими соседями, Å	
Вандерваальсова (молекулярные кристаллы)	Аргон	Г. ц. к.	0,1	3,76	Низкие температуры плавления и кипения
	Хлор	Тетрагональная	0,3	4,34	Высокая сжимаемость
	Водород	Г. п. у.	0,01	3,75	Высокие диэлектрические свойства Прозрачен для фотонов с энергией в далекой УФ-области
Ковалентная	Кремний	Кубическая (алмаз)	3,7	2,35	Низкая сжимаемость и высокая прочность. Во многих случаях высокая температура плавления
	InSb	Кубическая (цинковая обманка)	3,4	2,80	Диэлектрики или полупроводники
	Mg_2Sn	Кубическая (флюорит)	1,0	2,92	Сильное поглощение фотонов с энергиями выше края собственного поглощения, прозрачны (в отсутствие примесей) для бо́льших длин волн

Продолжение табл. 1.3

Тип связи	Типичные примеры				Некоторые характерные свойства
	Вещество	Кристаллическая структура	Энергия связи, эВ/молекула	Расстояние между ближайшими соседями, Å	
Ионная	KCl	Кубическая (каменная соль)	7,3	3,14	Довольно пластичны. Часто диссоциируют при нагревании. Диэлектрики при низких температурах
	AgBr	Кубическая (каменная соль)	5,4	2,88	Дефекты решетки обусловливают ионную проводимость при более высоких температурах
	BaF_2	Кубическая (флюорит)	17,3	2,69	Обладают поглощением остаточных лучей в ИК-области спектра и собственным поглощением
Водородная	Лед	Гексагональная	0,5	1,75	Множество аллотропных форм Не проводят электрического тока Высокие диэлектрические свойства Оптическая прозрачность
Металлическая	Натрий	О. ц. к.	1,1	3,70	Большое межатомное расстояние и высокое координационное число
	Серебро	Г. ц. к.	3,0	2,88	Хорошая электропроводность
	Никель	Г. ц. к.	4,4	2,48	Отсутствие прозрачности и высокая отражательная способность в ИК- и видимой области спектра Прозрачны в УФ-области спектра

в отдельных молекулах, поскольку общее количество связей в кристалле металла значительно возрастает (хотя каждая из них и становится слабее). Таким образом, энергия связи в расчете на один атом увеличивается от 0,6 эВ у Li_2 до 1,8 эВ у кристаллического лития.

В гл. 3 и 4 мы неоднократно будем пользоваться квантовомеханическим представлением о металлах как о совокупности положительно заряженных «ионных остовов» с большими промежутками между ионами, заполненными «электронным газом» таким образом, чтобы на макроскопическом уровне система оставалась электронейтральной. Здесь мы ограничимся следующими замечаниями:

1. Волновые функции электронов, составляющих этот газ, сильно перекрываются; волновые функции полностью делокализованы (и при наличии периодического ионного потенциала представляют собой блоховские функции).

2. Не все электроны, участвующие в образовании связи (валентные электроны), имеют одну и ту же энергию. Связь в металлических кристаллах обусловливается тем, что *средняя* энергия валентного электрона *меньше,* чем у изолированных атомов. На рис. 3.39 в разд. 3.4 (зонная теория твердых тел) это проиллюстрировано для металлического натрия.

3. Для более строгого рассмотрения металлической связи необходимо учесть электростатическое отталкивание электронных остовов (экранированных электронным газом), вандерваальсово притяжение между этими ионами и их отталкивание за счет перекрытия электронных оболочек (последние два взаимодействия очень слабы), связь, обусловленную незаполненностью внутренних оболочек (имеющую большое значение для переходных элементов типа железа), а также корреляцию электронов внутри электронного газа. Развитие физики в 1950-е и 1960-е годы достигло такого уровня, при котором некоторые из перечисленных взаимодействий вполне реально рассчитать.

Невозможно составить единую таблицу, в которой были бы отражены все существенные аспекты влияния преобладающего механизма связи на свойства твердых тел. Частично такое сравнение приводится в табл. 1.3, в которой кратко перечислены те свойства, которые будут рассмотрены в последующих главах.

1.2. Операции симметрии

Трансляционная симметрия

Идеальный монокристалл представляет собой бесконечное повторение в трехмерном пространстве идентичных блоков одинаковой ориентации. Каждый блок, называемый *базисом*, может

представлять собой атом, молекулу либо группу атомов или молекул. Базис — это количество вещества, содержащееся в *элементарной ячейке*, которая имеет форму трехмерного параллелепипеда. Перемещая этот параллелепипед на определенные дискретные расстояния во всех трех направлениях, можно заполнить все пространство.

Трансляционная симметрия является наиболее очевидной для кристаллических твердых тел. Согласно требованиям этой симметрии можно выбрать три вектора трансляции **a**, **b** и **c** таким образом, что оператор трансляции

$$\mathbf{T} = n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c} \qquad (1.23)$$

(где n_1, n_2 и n_3 — произвольные целые числа) будет соединять два положения в кристалле, имеющие одинаковые атомные окружения. Векторы трансляции **a**, **b** и **c** направлены вдоль трех смежных ребер элементарной ячейки, имеющей форму параллелепипеда. Однако трансляционная симметрия означает не только то, что расположения атомов на противоположных границах элементарной ячейки должны быть одинаковы, но и то, что локальное расположение атомов вблизи *любой* точки **r** (совпадающей или несовпадающей с положением атома) должно быть тем же самым, что и вблизи любого множества точек **r′**, связанных с **r** соотношением

$$\mathbf{r}' = \mathbf{r} + \mathbf{T}. \qquad (1.24)$$

Множество операторов трансляции **T** определяет *пространственную решетку* или *решетку Браве*. Понятие пространственной решетки является чисто геометрическим. Реальная кристаллическая решетка получается, когда базисы заполняют все пространство вокруг каждой геометрической точки решетки Браве.

Для гипотетического двумерного кристалла можно составить пять различных решеток Браве. На них мы проиллюстрируем некоторые особенности для случая двумерного пространства. Это полезно сделать, прежде чем мы перейдем к рассмотрению реальных трехмерных структур, в которых число возможных решеток Браве равно четырнадцати.

Лишь для определенных типов решеток можно выбрать векторы **a**, **b** и **c** одинаковой длины и лишь для некоторых простых решеток они являются взаимно перпендикулярными. Решетка, составленная из точек **r′**, и векторы трансляции **a**, **b** и **c** называются *примитивными*, если *каждая* точка, эквивалентная **r**, принадлежит совокупности точек **r′**. *Примитивным базисом* называется минимальное число атомов или молекул, которое достаточно для определения данной кристаллической структуры; это есть количество вещества в примитивной (наименьшей)

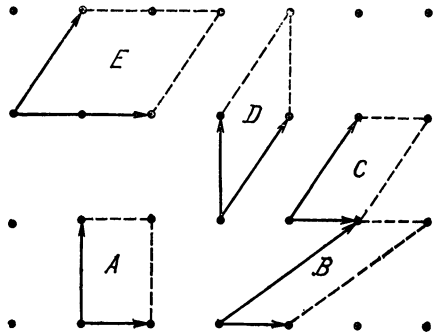


Рис. 1.14. Несколько способов задания элементарной ячейки двумерной прямоугольной решетки с помощью векторов трансляции. Ячейки *A*, *B*, *C*, *D* являются примитивными, а ячейка *E* непримитивная. Для этой структуры очевиден выбор *A*. У структур с более низкой симметрией выбор оптимального набора примитивных векторов может быть и не столь очевидным.

элементарной ячейке. Для некоторых типов кристаллических структур выбор примитивных векторов можно произвести несколькими способами (выбор векторов в большей *непримитивной* элементарной ячейке можно осуществить бесконечным числом способов). На рис. 1.14 показана наиболее логичная для плоской прямоугольной решетки комбинация двух взаимно перпендикулярных примитивных векторов. Здесь же приведены другие возможные, но менее удобные способы выбора примитивных векторов, а также один из способов выбора пары непримитивных векторов.

Базис

Пространственная решетка, или решетка Браве, еще не определяет расположение атомов, а описывает лишь геометрическое расположение точек в пространстве. Для описания кристаллической *структуры* необходимо, помимо решетки, задать также и симметрию *базиса*, т. е. атомов, расположенных в каждой точке кристаллической решетки. Базис — это совокупность координат атомов, расстояний между ними и направлений (углов) связей, которая повторяется в окрестности каждой точки решетки одинаковым образом, так что учитывается любой атом в кристалле.

Для задания базиса примитивной решетки нужно меньше атомов, чем при использовании непримитивных векторов, а для некоторых простых элементарных кристаллов (таких, как аргон или натрий) базис состоит лишь из одного атома. Другие элементы образуют кристаллические структуры, в которых при-

митивный базис содержит несколько атомов (например, базис из двух атомов у кремния и из четырех атомов у галлия). В соединениях базис должен содержать по крайней мере одну молекулу вещества. Отметим, что в некоторых сложных органических соединениях один базис включает в себя несколько тысяч атомов.

Внешнее строение кристалла (углы между гранями) определяется решеткой Браве. Однако симметрию базиса по отношению к операциям, таким, как вращение и отражение, нельзя отделить от симметрии решетки Браве. Лишь для нескольких типов решеток можно получить из атомов образования любой данной формы и симметрии, и высшая симметрия кристаллической решетки определяет высшую симметрию, которой может обладать ее базис. В связи с этим базису ставится в соответствие одна из конечного числа *точечных групп*. Каждая точечная группа характеризуется полным набором всех вращений, отражений и комбинаций вращений и отражений относительно некоторой точки, после выполнения которых кристалл переходит сам в себя. Как мы увидим в дальнейшем, все возможные образования в гипотетических двумерных кристаллах можно описать с помощью десяти точечных групп, в то время как трехмерные кристаллы имеют 32 точечные группы.

Элементарная ячейка

Говоря о кристалле как о совокупности повторяющихся в трехмерном пространстве одинаковых блоков, мы представляем себе в качестве одного блока либо атомный базис, либо параллелепипед, составленный из векторов **a**, **b** и **c**. Базис — это количество вещества, содержащееся[18] в объеме *элементарной ячейки*

$$V = \mathbf{a} \times \mathbf{b} \cdot \mathbf{c}. \tag{1.25}$$

Форма и размеры элементарной ячейки, а также распределение вещества внутри нее дают полное кристаллографическое описание кристалла. Будет ли это описание удобным в каждом конкретном случае, зависит от того, как мы выберем элементарную ячейку.

Это стоит отметить, поскольку выбор элементарной ячейки произволен в той же степени, что и выбор векторов трансляции. Однако независимо от того, выберем ли мы ячейку простейшей

[18] Мы считаем, что атом, расположенный *внутри* элементарной ячейки, принадлежит ей полностью, а атом на поверхности ячейки принадлежит ей только частично. Таким образом, от каждого из восьми атомов, расположенных в углах куба, учитывается лишь одна восьмая часть (см. рис. 1.15).

примитивной или нет, получаемая кристаллографическая информация автоматически будет полной. Иногда сознательный выбор непримитивной ячейки помогает подчеркнуть некоторые особенности симметрии кристалла.

Для гипотетической двумерной прямоугольной решетки, изображенной на рис. 1.14, таких тонкостей не существует. Ясно, что примитивная решетка A всегда является наиболее удобной. Однако при исследовании гранецентрированной кубической (г. ц. к.) решетки (рис. 1.15, *а*) простейшей будет *примитивная* ромбоэдрическая ячейка с тремя векторами трансляции длиной $(L/\sqrt{2})$, направленными под углами 60° друг к другу. Высокая кубическая симметрия этой решетки выявляется, если выбрать в качестве элементарной непримитивную ячейку, а именно *элементарный куб* с тремя взаимно перпендикулярными векторами трансляции длиной L. Поскольку объем элементарного куба равен L^3, а объем ромбоэдрической ячейки — $L^3/4$, в элементарном кубе содержится то же количество вещества, что и в четырех примитивных базисах.

Подобным же образом и для объемноцентрированной кубической (о. ц. к.) решетки выбор непримитивной элементарной ячейки в виде куба со стороной L гораздо лучше выявляет кубическую симметрию, чем примитивная ромбоэдрическая ячейка, которой соответствуют примитивные векторы длиной

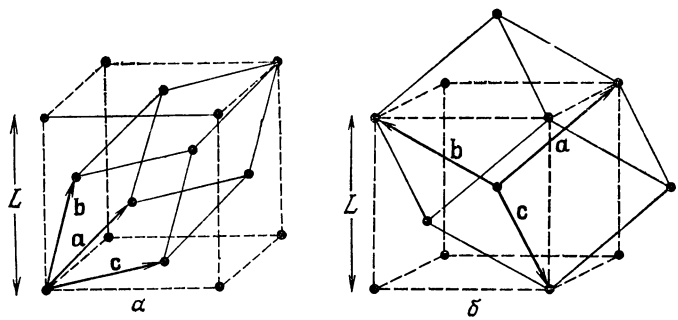


Рис. 1.15. *а* — г. ц. к.-решетка; *б* — о. ц. к.-решетка. На обоих рисунках штриховыми линиями показана общепринятая элементарная кубическая ячейка, а сплошными линиями изображена одна из возможных примитивных ячеек ромбоэдрической формы. Если $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$ и $\hat{\mathbf{z}}$ — единичные векторы, направленные вдоль трех взаимно перпендикулярных осей, то примитивную ячейку г.ц.к-решетки образуют векторы $\mathbf{a}=\frac{1}{2}L(\hat{\mathbf{x}}+\hat{\mathbf{y}})$, $\mathbf{b}=\frac{1}{2}L(\hat{\mathbf{y}}+\hat{\mathbf{z}})$, $\mathbf{c}=\frac{1}{2}L(\hat{\mathbf{z}}+\hat{\mathbf{x}})$. Эти векторы каждый длиной $L/\sqrt{2}$ образуют друг с другом углы 60°. Примитивная о. ц. к.-ячейка на рис. 1.15, *б* образуется векторами $\mathbf{a}=\frac{1}{2}L(\hat{\mathbf{x}}+\hat{\mathbf{y}}-\hat{\mathbf{z}})$, $\mathbf{b}=\frac{1}{2}L(\hat{\mathbf{y}}+\hat{\mathbf{z}}-\hat{\mathbf{x}})$, $\mathbf{c}=\frac{1}{2}L(\hat{\mathbf{z}}+\hat{\mathbf{x}}-\hat{\mathbf{y}})$. Эти векторы каждый длиной $L\sqrt{3}/2$ составляют друг с другом углы 109°.

$L\sqrt{3}/2$. Объем примитивной ячейки равен $L^3/2$, т. е. половине объема куба, который на рис. 1.15, *б* показан штриховыми линиями. Если бы мы выбрали другую примитивную комбинацию векторов в о. ц. к.-решетке и составили бы примитивную элементарную ячейку другой формы, то объем ее, несомненно, остался бы равным $L^3/2$. Можно получить неограниченное число различных возможных элементарных ячеек с объемом, большим или равным L^3, но кубическая ячейка на рис. 1.15, *б* является наиболее удобной.

Симметрия решетки

На первый взгляд может показаться странным, почему любое твердое вещество не может существовать в виде многочисленных кристаллических структур, соответствующих всем возможным типам решеток Браве. Утверждение о том, что в действительности одна из аллотропных модификаций оказывается предпочтительней любой другой *возможной* модификации, в большой степени определяется энергетической стабильностью, о которой мы уже упоминали при рассмотрении межатомных связей. Однако из условий симметрии кристаллов следует, что для любого вещества число возможных аллотропных форм ограниченно, причем нередко возможна лишь единственная структура. Это обусловлено требованием того, что точечная группа (симметрия базиса) должна обязательно согласовываться с симметрией самой решетки Браве.

Операции симметрии решетки удобно изучать с помощью мощного и замечательного аппарата теории групп[19], причем для детального изучения кристаллографии желательно использовать обозначения и методы этого аппарата. Однако нам необязательно привлекать эти алгебраические обозначения для того, чтобы получить понятие о физически реализуемых операциях симметрии. По замыслу этой книги, знание элементов кристаллографии должно служить только основой для изучения тех свойств твердых тел, которые связаны с существованием периодической структуры. В литературе, приведенной в конце этой главы, содержится более исчерпывающее рассмотрение вопросов кристаллографии и симметрии решетки.

---

[19] О применении методов теории групп для изучения кристаллических структур см. в книгах: *Knox R. S., Gold A.* Symmetry in the Solid State, W. A. Benjamin, 1964 [имеется перевод: *Нокс Р., Голд А.* Симметрия в твердом теле.— М.: Наука, 1970]; *Koster G. F.* Space Groups and Their Representations, Academic Press, 1964. В этих книгах рассматриваются также электронные и другие свойства кристаллов, возникающие как следствие симметрии решетки, причем используется терминология теории группы

В дополнение к трансляционной симметрии рассмотрим, каким образом решетка может оставаться инвариантной (не измениться по сравнению с первоначальной) по отношению к любой из следующих операций:

отражение от плоскости;
вращение вокруг оси (1, 2, 3, 4 или 6-го порядка);
инверсия в точке;
скольжение (отражение в плоскости с одновременным переносом параллельно плоскости);
винтовой поворот (вращение вокруг оси с одновременным переносом вдоль этой же оси).

Последние две операции, которые требуют выполнения комбинированных действий, называются *сложными операциями*. Инверсию в точке можно также рассматривать как сложную операцию, поскольку она эквивалентна вращению на 180° с последующим отражением в плоскости, перпендикулярной оси вращения. Воображаемая плоскость, от которой происходит отражение, или ось вращения не обязательно проходит через точки решетки, что четко демонстрируется на примерах для двумерных случаев на рис. 1.18.

Принято считать, что кристаллическая структура имеет центр инверсии, если расположение атомов относительно произвольной точки **r** (расстояние **r** измеряется относительно центра инверсии) совпадает с расположением атомов относительно противоположной точки — **r**. Таких кристаллов в действительности гораздо меньше, чем можно предположить с первого взгляда. В тех кристаллах, в которых *существует* симметрия относительно инверсии, центр инверсии либо совпадает с одним из атомов, либо находится в точке, равноудаленной от двух атомов. Как легко убедиться из следующих ниже рисунков, центр инверсии может существовать в *любой* решетке Браве и его отсутствие в большинстве кристаллов связано лишь с геометрическими особенностями *базиса*.

Плоскость же отражения (плоскость зеркальной симметрии) отсутствует в ряде случаев, так как она несовместима с некоторыми из основных решеток Браве, независимо от вида базиса. В дальнейшем мы покажем, что плоскость отражения либо совпадает с одной из граней элементарной ячейки, либо делит элементарную ячейку пополам.

Говорят, что кристалл обладает осью вращения n-го порядка, если после поворота на угол $2\pi/n$ расположение атомов не изменяется. Выше уже упоминались операции вращения для случаев $n=1$ (это универсальная и тривиальная операция, которая на языке теории групп называется единичным элементом или единичной операцией), а также для $n=2$, 3, 4 и 6. Возникает

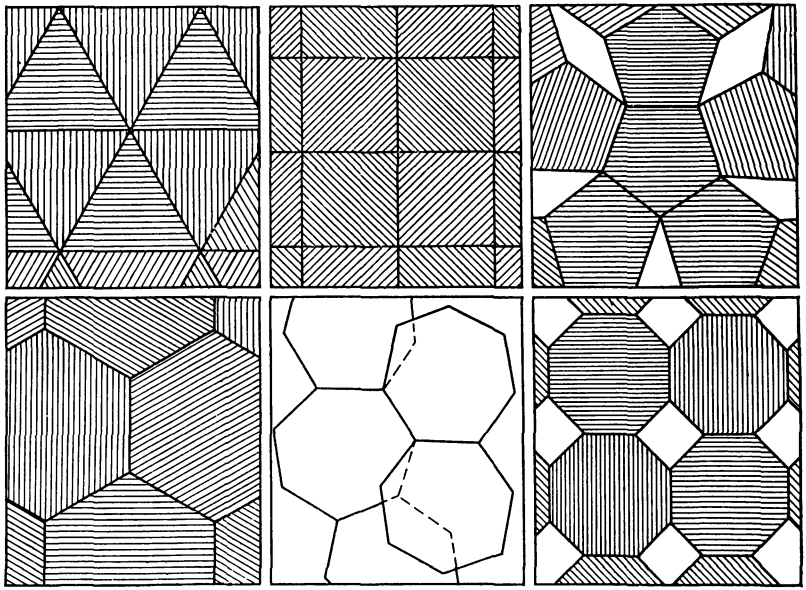


Рис. 1.16. Из всех правильных многоугольников только треугольники, квадраты и шестиугольники могут образовать плотноупакованную решетку, демонстрируя таким образом свойства как трансляционной, так и вращательной симметрии.

вопрос, почему в симметрии кристаллов повороты на углы, не равные 180, 120, 90 и 60°, недопустимы, ведь в природе, естественно, встречаются и другие виды симметрии, например пятиконечная морская звезда. Дело в том, что наличие трансляционной симметрии в кристаллах допускает существование лишь осей 1, 2, 3, 4 и 6-го порядка. Это показано на рис. 1.16. Правильные треугольники, квадраты и шестиугольники, полностью стыкуясь друг с другом, образуют плотную упаковку; попытка же заполнить всю область правильными многоугольниками любой другой вращательной симметрии приводит к тому, что либо остаются незаполненные участки, либо многоугольники перекрываются *.

Эти утверждения можно проверить на простой геометрической модели (рис. 1.17) двумерного «кристалла», имеющего в горизонтальном направлении постоянную решетки a. Тогда в ряду A расстояние между 1-м и m-м атомами равно $(m-1)\mathbf{a}$. Теперь допустим, что симметрией данной решетки разрешен

* В последние годы активно изучаются так называемые квазикристаллы — системы, в которых существует ось 5-го порядка.— *Прим. ред.*

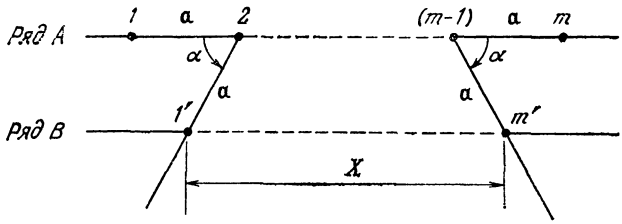


Рис. 1.17. Двумерная модель, показывающая, какие углы вращения совместимы также с трансляционной симметрией кристалла.

поворот на угол α. При вращении вокруг атома 2 в направлении против часовой стрелки атом 1 переместится в положение, занятое прежде атомом $1'$. Аналогично при вращении по часовой стрелке вокруг $(m-1)$-го атома атом m переместится в положение, занятое ранее атомом m'. Очевидно, что атомы $1'$ и m' принадлежат ряду B, так что расстояние $\mathbf{X}$ между ними должно быть кратно $\mathbf{a}$, если поворот на угол α разрешен симметрией решетки. Предположим, что $\mathbf{X}=p\mathbf{a}$, где p — неизвестное целое число. Разность двух неизвестных целых чисел m и p можно выразить через α. Из рис. 1.17 с помощью очень простой тригонометрии получаем

$$\mathbf{X} = p\mathbf{a} = (m-3)\,\mathbf{a} + 2\mathbf{a}\cos\alpha, \tag{1.26}$$

откуда

$$\cos\alpha = (3+p-m)/2. \tag{1.27}$$

При условии что m и p целые числа, это уравнение имеет всего пять решений, которые представлены в табл. 1.4.

В дальнейшем при рассмотрении различных двумерных и трехмерных решеток будет показано, что инвариантность по отношению к сложным операциям, включающим, например, вращение с одновременной инверсией, скольжение или же винто-

Таблица 1.4. **Решения уравнения (1.27) для вращений, разрешенных в периодической решетке**

$p-m$	$\cos\alpha$	α	Порядок оси вращения
—1	1	0	1-й
—2	1/2	$\pi/3$	6-й
—3	0	$\pi/2$	4-й
—4	—1/2	$2\pi/3$	3-й
—5	—1	π	2-й

вой поворот может встречаться даже чаще, чем инвариантность по отношению к «простым» операциям. Прежде чем приступить к изучению реальных трехмерных решеток, целесообразно рассмотреть основные понятия на примере более простых гипотетических двумерных решеток, точечных и *плоских групп. Плоская группа* — это двумерный эквивалент *пространственной группы* в трех измерениях; т. е. это совокупность всех операций симметрии, при выполнении которых идеальный двумерный кристалл переходит сам в себя.

Плоские решетки и их симметрия

Существует всего пять различных способов объединения линейных цепочек в двумерную решетку Браве. Эти решетки вместе с присущими им элементами симметрии изображены на рис. 1.18. Точечная симметрия, показанная в каждом случае,— это симметрия *решетки*; симметрия образующегося кристалла будет ниже, если симметрия базиса, соответствующего каждой точке решетки, ниже симметрии решетки. Как следует из рисунка, центр инверсии может существовать в каждой из решеток (см. задачу 1.6).

Для каждого из пяти типов решеток оси вращения и (или) плоскости отражения проходят через точки решетки. Однако в элементарной ячейке существуют и другие точки с таким же или более низким порядком симметрии по отношению к вращениям и отражениям.

Наиболее общим типом двумерной решетки является *косоугольная* решетка — первая из изображенных на рис. 1.18. Возвращаясь к рис. 1.14, вспомним, что для прямоугольной решетки элементарную ячейку (примитивную или непримитивную) можно выбрать бесконечным числом способов, но в данном случае логичнее всего выбрать примитивную ячейку, определенную единичными векторами **a** и **b** на рис. 1.18. В наиболее общем случае, когда **a**, **b** и α не связаны между собой каким-то конкретным соотношением, существуют только центр инверсии, о котором уже говорилось выше, и оси вращения второго порядка, положения которых указаны на рисунке. Линия отражения или же оси вращения более высокого порядка возможны либо при определенных соотношениях между **a** и **b**, либо при вполне определенных значениях угла α.

К таким решеткам специального вида относится *примитивная прямоугольная* решетка, для которой $\alpha = 90°$. Поскольку на отношение длин **b**/**a** не накладывается никаких ограничений, порядок оси вращательной симметрии по-прежнему не больше двух, но теперь возникают два набора взаимно перпендикулярных линий зеркального отражения. В другой решетке специ-

ального вида — *центрированной прямоугольной решетке* — угол α не прямой, а равен $\arccos(a/2b)$. Даже если примитивной элементарной ячейкой для этой решетки является косоугольный параллелограмм, из рис. 1.18 можно видеть, что непримитивную ячейку можно выбрать в виде центрированного прямоугольника. *Центрированная прямоугольная* решетка имеет линии отражения и оси вращения второго порядка, присущие непримитивной прямоугольной ячейке, и, кроме того, оси вращения второго порядка, характерные для примитивной косоугольной элементарной ячейки.

Еще одна решетка специального вида — это *квадратная решетка*, которая имеет оси симметрии четвертого порядка, проходящие через середины сторон. Оси вращения более высокого порядка проходят через точки пересечения линий отражения, образующих друг с другом углы 45°.

Последняя из специальных кристаллических решеток, которую можно получить, накладывая ограничения как на углы, так и на соотношения между сторонами параллелограмма, является *гексагональной* решеткой. Ее можно представить себе как частный случай центрированной прямоугольной решетки, у которой $a=b$ и $\alpha=60°$. Однако, рассматривая гексагональную решетку с этой точки зрения, можно не заметить элементов симметрии, свойственных самому гексагональному расположению. Как показано на рис. 1.18, лучше всего выбрать элементарную примитивную ячейку в форме ромба с углами 120 и 60°. Три такие примитивные ячейки образуют «гексагональную» элементарную ячейку. Линии отражения пересекаются под углом 30° друг к другу в точках, через которые проходят оси вращения высшего порядка, а именно шестого порядка.

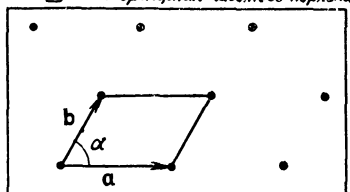
Двумерные точечные группы и пространственные группы

Во всех двумерных решетках Браве на рис. 1.18 точками обозначены элементарные группы атомов или «базисы». Любой базис должен состоять из одинаково расположенных атомов и иметь одинаковые ориентации в пространстве, симметрия базиса должна соответствовать симметрии решетки.

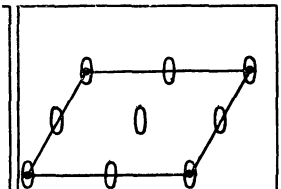
Существует лишь десять способов совместить симметрию базиса с симметрией двумерных решеток Браве. Эти десять *точечных групп* изображены на рис. 1.19. Цифры показывают число положений внутри базиса, эквивалентных по отношению к операциям вращения. Буква m означает, что базис обладает линией зеркальной симметрии. Заметим, что если существуют

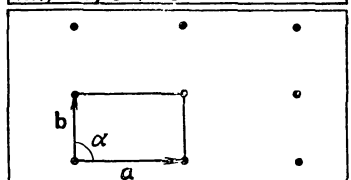
Рис. 1.18. Пять двумерных решеток и их симметрии.

– ось вращения второго порядка
– ось вращения третьего порядка
– ось вращения четвертого порядка
– ось вращения шестого порядка
– плоскость зеркальной симметрии
– ортогональные зеркальные плоскости
– зеркальные плоскости, следующие под углом 45°

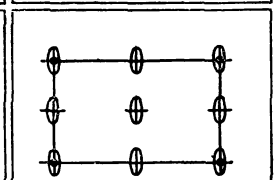


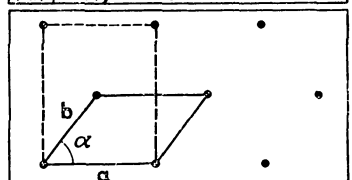
Косоугольная решетка
$|a| \neq |b|$, $\alpha \neq 90°$





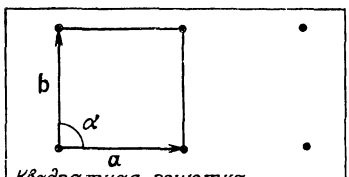
Прямоугольная решетка
$|a| \neq |b|$, $\alpha = 90°$



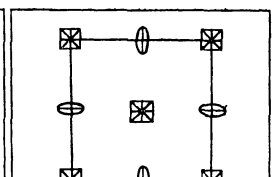


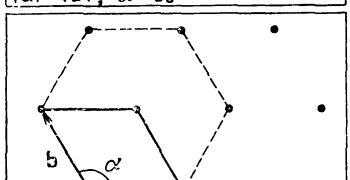
Центрированная прямоугольная решетка $\alpha = \arccos(a/2b)$

Симметрия любой прямоугольной решетки плюс минимальная симметрия любой косоугольной решетки

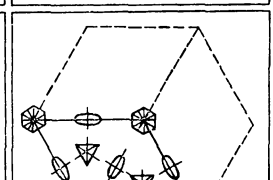


Квадратная решетка
$|a| = |b|$, $\alpha = 90°$





Гексагональная решетка
$|a| = |b|$, $\alpha = 120°$



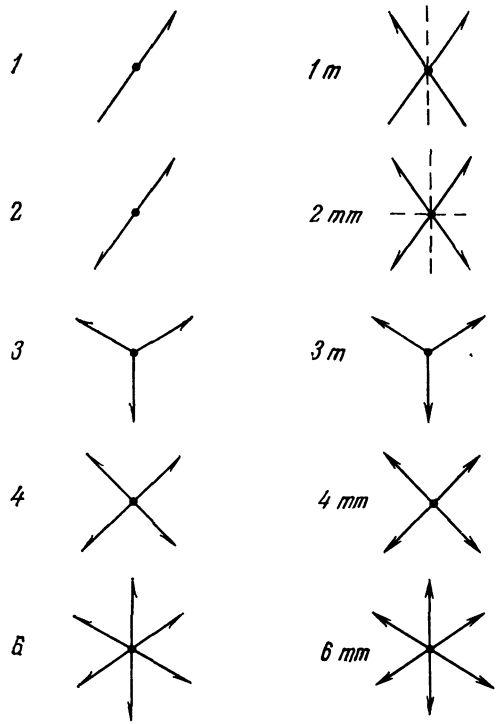


Рис. 1.19. Десять типов точечных групп, которые могут быть совместимы с симметриями плоской решетки.

Таблица 1.5. **Двумерные кристаллические системы, решетки и плоские группы**

Кристаллическая система	Точечные группы, совместимые с системой	Решетки Браве, входящие в систему	Число плоских групп, совместимое с решеткой
Косоугольная	1, 2	*P* (примитивная)	2
Прямоугольная	1*m*, 2*mm*	*P* (примитивная)	5
		C (центрированная)	2
Квадратная	4*m*, 4*mm*	*P* (примитивная)	3
Гексагональная	3,3*m*, 6,6*mm*	*P* (примитивная)	5

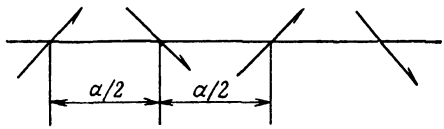


Рис. 1.20. Операция скольжения (отражение от горизонтальной линии, проходящей через середину стрелок, плюс перенос на $a/2$) не изменяет решетки. Вторичное применение операции скольжения приводит к тому, что стрелка возвращается в положение, эквивалентное ее первоначальному. Симметрии операции скольжения в двумерном пространстве имеют большое значение для трех примитивных прямоугольных и одной квадратной структуры.

ось вращения 2-го порядка и одна линия зеркальной симметрии, то обязательно *должна* существовать и другая ось зеркальной симметрии, перпендикулярная первой; отсюда следуют обозначения *2mm*, *4mm* и *6mm*.

Сравнение элементов симметрии десяти точечных групп с элементами симметрии пяти решеток Браве приводит первоначально к существованию только 13 двумерных *плоских групп* или *структур*, так как каждая из точечных групп совместима с симметрией лишь одной двумерной кристаллической системы. Из табл. 1.5 видно, что прямоугольная система — единственная из двумерных систем, которая включает в себя более чем одну решетку Браве (хотя, как мы увидим ниже, большинство трехмерных систем содержат по нескольку типов решеток). Существование еще четырех плоских групп связано с возможностью добиться инвариантности относительно операции скольжения (состоящей, как показано на рис. 1.20, из отражения с одновременным переносом на половину длины основного вектора решетки). Таким образом, всего существует 17 плоских групп. Их строение рассматривал Нассбаум [20] и Бюргер [21]. Нассбаум показал также, что воображаемая двумерная кристаллическая структура подчиняется тем же закономерностям, что и повторяющийся рисунок обоев. В табл. 1.5 показано распределение 17 плоских групп по четырем кристаллическим системам.

Трехмерные кристаллические системы, точечные группы и пространственные группы

В трехмерном пространстве наиболее общим типом решетки является *триклинная* решетка, у которой простейшая примитивная ячейка имеет вид параллелепипеда с отличными от прямых

[20] *Nussbaum A.* Semiconductor Device Physics, Prentice-Hall, 1962, ch. 5.
[21] *Buerger M. J.* Contemporary Crystallography, McGraw-Hill, 1970, ch. 1.

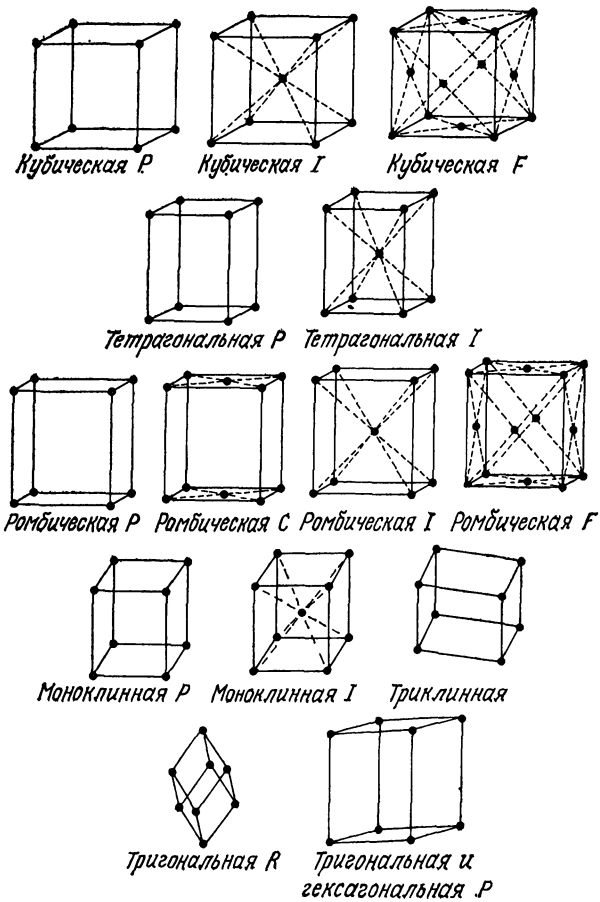


Рис. 1.21. Наиболее естественные элементарные ячейки для четырнадцати возможных трехмерных решеток Браве. Для семи решеток используются непримитивные ячейки (*Kittel C.* Introduction to Solid State Physics, Wiley, 1971. [Имеется перевод: *Киттель Ч.* Введение в физику твердого тела.— М.: Наука, 1978.])

углами и примитивными векторами a, b и c разной длины. При выполнении некоторых специальных соотношений между сторонами и углами ячейки получаются 13 «особых типов решеток», т. е. всего 14 трехмерных решеток Браве. Эти решетки представлены на рис. 1.21 либо примитивными ячейками, либо непримитивными, если последние лучше выявляют свойства симметрии. Из табл. 1.6 видно, как 14 решеток можно разде-

Таблица 1.6. **Семь трехмерных кристаллических систем и четырнадцать решеток Браве**

Система	Ограничения, накладываемые на размеры и углы в стандартных элементарных ячейках	Решетки Браве, совместимые с системой
Триклинная	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma$	*P* (примитивная)
Моноклинная	$a \neq b \neq c$ $\alpha = \gamma = 90° \neq \beta$	*P* (примитивная) *I* (объемно-центрированная)
Ромбическая	$a \neq b \neq c$ $\alpha = \beta = \gamma = 90°$	*P* (примитивная) *C* (базоцентрированная) *I* (объемно-центрированная) *F* (гранецентрированная)
Тетрагональная	$a = b \neq c$ $\alpha = \beta = \gamma = 90°$	*P* (примитивная) *I* (объемно-центрированная)
Кубическая	$a = b = c$ $\alpha = \beta = \gamma = 90°$	*P* (примитивная или простая кубическая) *I* (объемно-центрированная) *F* (гранецентрированная)
Тригональная или ромбоэдрическая	$a = b = c$ $120° > \alpha = \beta = \gamma \neq 90°$	*R* (ромбоэдрическая примитивная)
Гексагональная	$a = b \neq c$ $\alpha = \beta = 90°, \gamma = 120°$	*P* (примитивная ромбоэдрическая)

лить на семь кристаллических систем *. Элементарные ячейки для этих семи систем получаются последовательными нарушениями кубической симметрии, как показано на рис. 1.22.

Подробное обсуждение симметрий, свойственных семи кристаллическим системам и тридцати двум трехмерным *точечным группам*, заняло бы непозволительно много места в данной книге [22]. Для наших целей вполне достаточно лишь заметить,

* Для разделения кристаллов по конфигурации их элементарной ячейки в кристаллографии наряду с термином «система» используется также термин «сингония».— *Прим. ред.*

[22] Для более подробного изучения см. библиографию в конце гл. 1 (в частности, книгу Мегау). Некоторые аспекты этой проблемы рассматриваются с весьма интересных позиций в гл. 5 книги: *Nussbaum A.* Semiconductor Device Physics, Prentice-Hall, 1962.

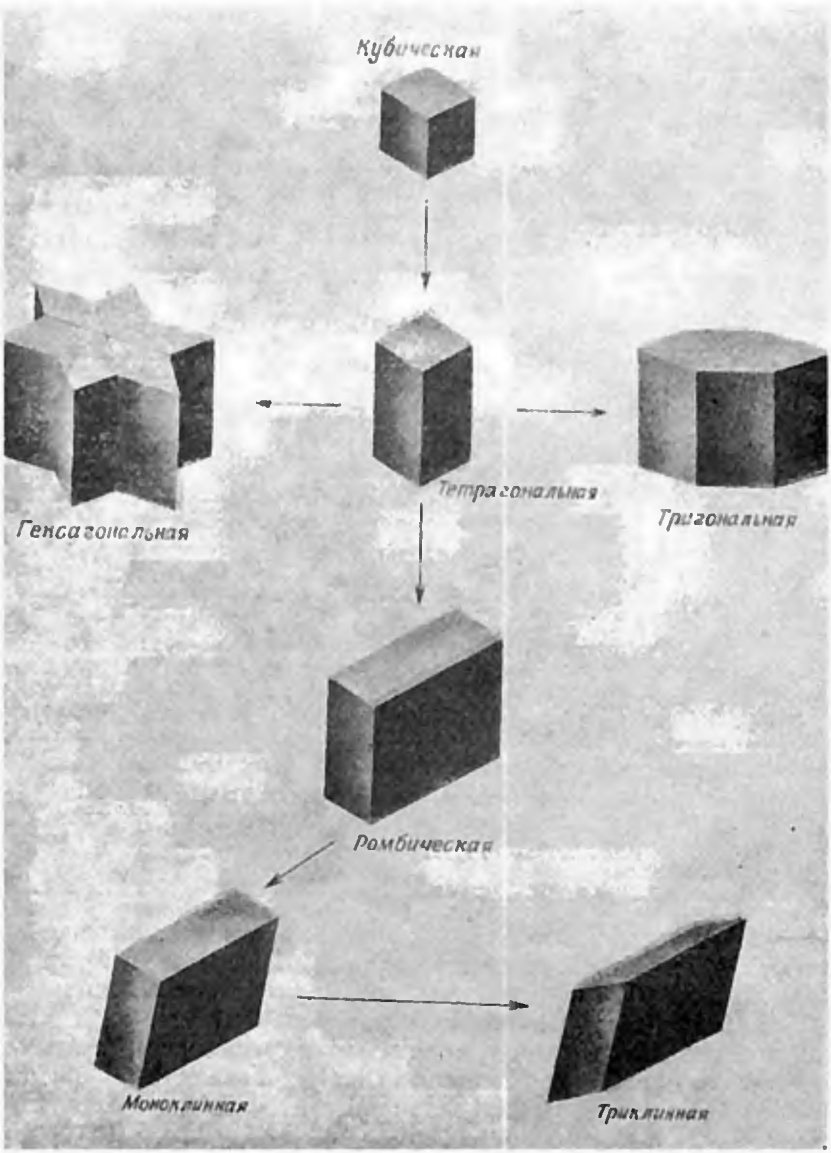


Рис. 1.22. Семь трехмерных кристаллических систем, полученных последовательными искажениями кубической симметрии [*Nussbaum A.* Semiconductor Device Physics, Prentice-Hall, 1962].

что каждая из 32 точечных групп симметрии оказывается совместима с одной и только с одной из семи кристаллических систем. Из 32 точечных групп две связаны с триклинной решеткой, три — с моноклинной, три — с ромбической, семь — с тетрагональной, пять — с кубической, пять — с тригональной, а остальные семь — с гексагональной. Непосредственное сравнение этих 32 точечных групп с соответствующими решетками Браве приводит к существованию 73 «простых» трехмерных *пространственных групп*. Но это еще не все; при добавлении сложных операций симметрии возникает еще 157 дополнительных *сложных пространственных групп*. Таким образом, всего существует 230 трехмерных пространственных групп или «структур». Несмотря на такое многообразие, большинство важных и хорошо изученных кристаллов характеризуется небольшим числом относительно простых структур.

Сто пятьдесят семь дополнительных пространственных групп возникает при рассмотрении сложных элементов симметрии, определяемых операциями *скольжения* и *винтового поворота*. Операция скольжения уже обсуждалась нами в связи с определением числа двумерных плоских групп и иллюстрируется рис. 1.20. Операция *винтового поворота*, имеющая смысл только в трехмерном пространстве, представляет собой трансляцию, сопровождающуюся поворотом на такой угол (60, 90, 120 или 180°), чтобы сохранилась инвариантность структуры. Эта операция продемонстрирована на простом примере кристаллической структуры теллура на рис. 1.6. На первый взгляд решетка теллура представляется гексагональной, однако при более внимательном рассмотрении видно, что существование винтовой оси вдоль каждой спиральной цепочки нарушает принцип гексагональной симметрии. Структура остается инвариантной (т. е. кристалл переходит в самого себя) при повороте на 120° и последующей трансляции на одну треть вектора **c**. Такая же комбинация — трансляция с поворотом на 120° является необходимым элементом симметрии синтетического кристаллического полимера изотактического полистирола, который имеет спиральную цепочку вида $(CH_2CH)_n$ с присоединенными фенильными группами, по три группы на каждый виток спирали[23].

[23] Выше мы уже упоминали о спиральных структурах в связи с водородными связями в двойной спирали ДНК. Имеет ли молекула ДНК винтовую симметрию? На этот вопрос можно дать следующий ответ: внешне имеет, но если говорить о деталях, то для молекулы ДНК нарушается первое условие этого раздела, согласно которому кристалл должен быть бесконечным повторением *одинаковых* блоков. Синтетические полинуклеотиды могут быть составлены из бесконечной последовательности идентичных азотных оснований, но в природной молекуле ДНК существует сложное апериодическое чередование четырех оснований, несущих генетическую информацию. Таким образом, «кристалл» ДНК или другого важного биологического вещества

Необходимо напомнить, что в кристаллографии термины «структура» или «пространственная группа» не эквивалентны кристаллу, а обозначают совокупности математических точек, каждой из которых поставлена в соответствие группа атомов с данной симметрией. Два кристалла могут иметь одну и ту же структуру, даже если базис одного из них состоит всего из двух или трех атомов, а базис другого содержит несколько тысяч атомов. В целом картины дифракции рентгеновских лучей на обоих кристаллах будут выглядеть одинаково, однако изучение более тонких деталей, а именно отношений амплитуд различных дифрагированных линий, позволит получить информацию об атомах, составляющих базис.

1.3. Структуры реальных кристаллов

Индексы Миллера и кристаллографические направления

Часто возникает необходимость задать определенную кристаллографическую плоскость или определенное направление в реальном трехмерном кристалле. Плоскости кристалла задаются, как правило, *индексами Миллера*. Эти индексы определяют следующим образом.

Пусть мы имеем общий случай трехмерной решетки, в которой длины примитивных векторов трансляции **a**, **b** и **c** произвольны, а углы между парами этих векторов не обязательно равны 90°. На рис. 1.23 показана часть такой решетки, где точки O, A, B и C представляют собой узлы решетки и выбраны таким образом, что векторы **OA**, **OB** и **OC** совпадают по направлению с примитивными векторами. Расстояния $|\mathbf{OA}|$, $|\mathbf{OB}|$ и $|\mathbf{OC}|$ в целое число раз n_1, n_2 и n_3 превосходят длины соответствующих примитивных векторов элементарной ячейки. Каким образом с помощью целых чисел n_1, n_2 и n_3 можно указать положение плоскости ABC?

Для задания плоскости ABC вполне было бы достаточно указать проекции ($n_1|\mathbf{a}|$, $n_2|\mathbf{b}|$, $n_3|\mathbf{c}|$), однако такое обозначение слишком громоздко и потому не используется. Не используется и более краткое обозначение плоскости, а именно (n_1, n_2, n_3). Вместо этого по причинам, которые станут очевидными после того, как мы рассмотрим свойства обратного пространства и обратных решеток, плоскость ABC удобно обозначать

состоит из похожих, но не полностью идентичных частей. Эта небольшая неоднородность не мешает дифракции рентгеновских лучей на кристалле ДНК (см. примечание 17).

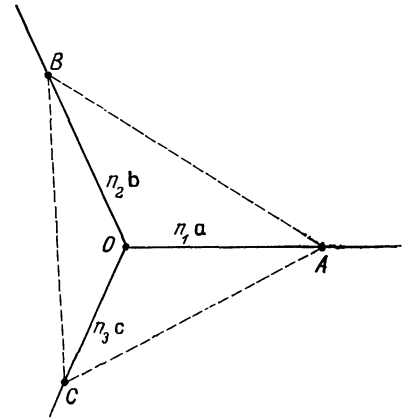


Рис. 1.23. Плоскость, проходящая через точки решетки A, B и C; a, b и c — примитивные векторы, а n_1, n_2 и n_3 — произвольные целые числа.

как (hkl), где целые числа h, k, l представляют собой набор целых чисел, которые выражают отношение

$$h : k : l = n_1^{-1} : n_2^{-1} : n_3^{-1}. \qquad (1.28)$$

Говорят, что плоскость, которая удовлетворяет отношению (1.28), имеет *индексы Миллера* (hkl). Обычно в качестве h, k и l выбирают наименьшие целые числа, удовлетворяющие соотношению (1.28), однако, как мы увидим в дальнейшем, для извлечения дополнительной информации можно поступить и иным способом.

Использование индексов Миллера проще всего продемонстрировать на примере кубической решетки (рис. 1.24). Таким же образом обозначают плоскости и в структурах с более низкой симметрией. Плоскость, обозначаемая ($hk0$), пересекает ось c (направление **OC** на рис. 1.23) на бесконечно большом расстоянии; очевидно, такая плоскость параллельна оси **c**. Аналогично плоскость, параллельная осям **b** и **c**, обозначается как ($h00$). В соответствии с принципом использования наименьшего из подходящих целых чисел любая плоскость ($h00$), если о ней не нужно сообщать дополнительной информации, обозначается как (100).

В случае когда рассматриваемая плоскость пересекает ось a в области отрицательных значений координат, в соответствии с терминологией Миллера используют обозначение ($\bar{h}kl$). На рис. 1.24 показаны плоскости (100) и ($\bar{1}00$), которые связаны друг с другом трансляционной симметрией. Символом $\{hkl\}$ обозначаются все плоскости, эквивалентные (hkl) в силу симметрии.

На рис. 1.24 изображена плоскость (200), где отличие первого индекса от единицы указывает на то, что эта плоскость

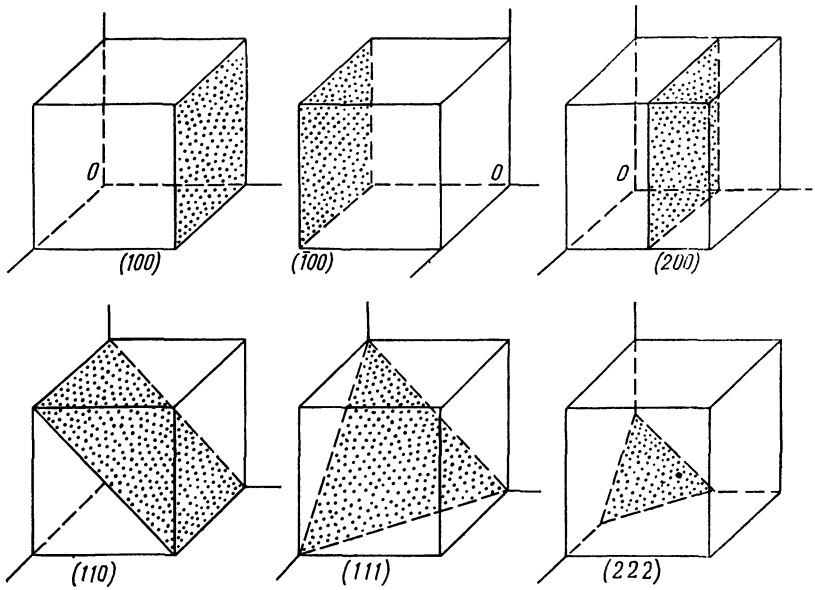


Рис. 1.24. Некоторые наиболее часто встречающиеся *плоскости* в кубической решетке и их индексы Миллера.

отсекает на оси **a** отрезок длиной (1/2)**a**. Аналогичным образом связаны на том же рисунке плоскости (222) и (111).

В гексагональных и тригональных кристаллах положения кристаллографических плоскостей по-прежнему можно задавать тремя индексами Миллера. Однако для этих кристаллографических классов более принято использовать набор *четырех* индексов Миллера, представляющих собой обратные значения координат точек пересечения плоскости с осями $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ и с осью **c** (рис. 1.25). Таким образом, плоскость можно обозначать как $(10\bar{1}0)$ или (0001), где четвертый индекс связан с пересечением с осью **c**. Соотношения между компланарными векторами $\mathbf{a}_1$, $\mathbf{a}_2$ и $\mathbf{a}_3$ накладывают условия на первые три индекса Миллера; их сумма должна быть равна нулю.

Пример плоскости $(\bar{1}010)$ на рис. 1.25 показывает, каким образом индексы, обозначающие плоскость, связаны с выбранными направлениями векторов $\mathbf{a}_1$, $\mathbf{a}_2$ и $\mathbf{a}_3$. Упражнение на применение этих обозначений содержится в задаче 1.11.

В кристаллах низкой симметрии введение индексов Миллера требует некоторого предварительного изучения кристалла, поскольку необходимо измерить координаты точек пересечения

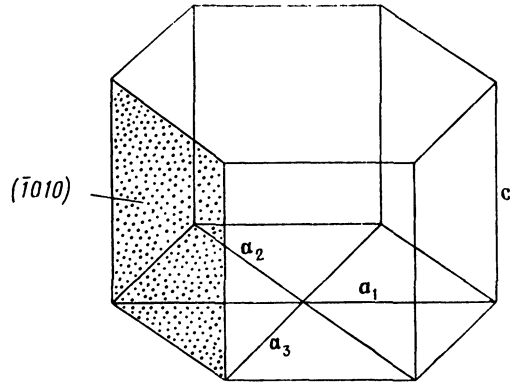


Рис. 1.25. Наиболее естественная элементарная ячейка для гексагональных и тригональных кристаллов. Показаны четыре вектора (из которых любые два из первых трех независимы), используемые в обозначениях через четырехзначные индексы Миллера.

плоскости с координатными осями и выразить их в длинах соответствующих единичных векторов. Привыкнув к такой системе обозначений, можно добиться существенного упрощения рассмотрения. Без индексов Миллера описание диагональных плоскостей в ромбической, моноклинной и триклинной решетках было бы значительно сложнее.

Интересно отметить, что для указания *кристаллографических направлений* традиционно используется иной подход. Направление в кристалле обозначают тремя целыми числами в квадратных скобках: [*hkl*]. Это обозначение настолько похоже на индексы Миллера для кристаллографических плоскостей, что невольно возникает предположение о том, что направление [*hkl*] ортогонально плоскости (*hkl*). К сожалению, кристаллографические направления определяются таким образом, что это предположение оказывается верным лишь для кубических кристаллов и для нескольких других отдельных случаев.

Кристаллографическое направление [*hkl*] определяется как набор наименьших целых чисел, относящихся друг к другу как косинусы углов между данным направлением и осями элементарной ячейки. Иными словами, числа *h*, *k* и *l* пропорциональны длинам векторов, направленных вдоль осей элементарной ячейки, которые в сумме составляют вектор, ориентированный в данном направлении.

Определение положения точки в элементарной ячейке

Положение точки в элементарной ячейке определяется тремя числами: u, v и w. Каждое из них описывает расстояние от данной точки до начала координат, выраженное в единицах длины ребер элементарной ячейки. Таким образом, если положение точки в одном углу элементарной ячейки обозначить как 000, то точка в противоположном углу обозначится как 111, а промежуточным точкам внутри или на поверхности ячейки будут соответствовать дробные числа. В этих обозначениях точка $^1/_2{}^1/_2 \cdot 0$ находится в центре грани, а точка $^1/_2{}^1/_2{}^1/_2$ — в центре ячейки. В дальнейшем мы будем часто пользоваться этими обозначениями при рассмотрении структур реальных кристаллов.

Простые плотноупакованные структуры

Если атомы одноатомного кристалла считать относительно несжимаемыми шарами, то естественно ожидать, что они будут плотно прилегать друг к другу и образовывать наиболее компактную структуру, если только особенности механизма связи не приведут к установлению иного типа симметрии с более редким расположением атомов. Те, кто наблюдал игру в бильярд, знают, что каждый из положенных плотно друг к другу шаров на плоскости касается шести соседей (рис. 1.26). Из таких плоских упаковок с гексагональной симметрией можно двумя простыми способами сложить наиболее компактный одноатомный кристалл. В обоих случаях каждый атом будет иметь двенадцать ближайших соседей — шесть в той плоскости, где он находится, три в плоскости над этим атомом и три в плоскости под ним.

На рис. 1.27 и 1.28 показаны два способа последовательной укладки слоев. В каждой из последовательностей второй слой B следует уложить над первым слоем A и сместить его так, чтобы каждый атом из слоя B попадал в одно из углублений первого слоя. Очевидно, что слой B будет плотно прилегать к слою A. Также очевидно, что третий слой, идентичный слою A, можно плотно уложить сверху на слой B. Продолжая последовательность слоев $ABABAB$, получаем гексагональную плотноупакованную (г. п. у.) структуру.

Элементарная ячейка г. п. у.-структуры изображена на рис. 1.29, причем размеры атомов здесь уменьшены для того, чтобы показать их геометрическое расположение в решетке. Структура имеет *простую гексагональную* пространственную решетку с базисом из двух атомов, один из которых расположен в точке 000, а другой — в точке $^2/_3 : {}^1/_3 : {}^1/_2$. Из элементар-

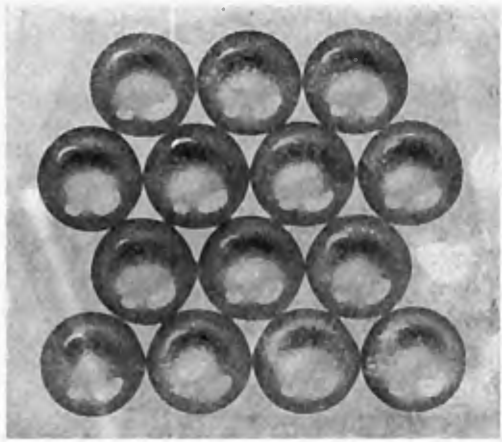


Рис. 1.26. Расположение шаров в слое с плотной упаковкой.

ных геометрических соображений ясно, что отношение высоты ячейки к расстоянию между ближайшими соседями равно

$$c/a = (8/3)^{1/2} = 1{,}633. \qquad (1.29)$$

Оказалось, что на самом деле для большинства реальных кристаллов с гексагональной структурой отношение c/a больше 1,633, т. е. в гексагональных структурах плотность упаковки не является максимальной. Тем не менее, если отклонение отношения c/a от идеального не превышает 10 %, такие кристаллы можно все еще рассматривать как г. п. у.-структуры.

Примерами элементов с кристаллической г. п. у.-структурой являются элементы второй группы Периодической системы: Be, Mg, Zn и Cd. В определенном интервале температур некоторые переходные элементы также образуют устойчивые г. п. у.-структуры.

В правой части рис. 1.28 показан другой способ плотной упаковки шаров. В этом случае третий слой *C* располагается над слоем *B* таким образом, что его атомы попадают в те углубления слоя *B*, которые расположены прямо над углублениями слоя *A*. Затем добавляется четвертый слой, идентичный слою *A*, и последовательность *ABCABCABC* повторяется. В результате образуется *кубическая* решетка. Гранецентрированную кубическую (г. ц. к.) структуру имеют кристаллы благородных газов и некоторые металлы, такие, как Cu, Ag, Au, Ni, Al и Pb. Таким образом, двум различным хорошо известным металлическим структурам с высокой степенью симметрии соответствуют два вида плотной упаковки.

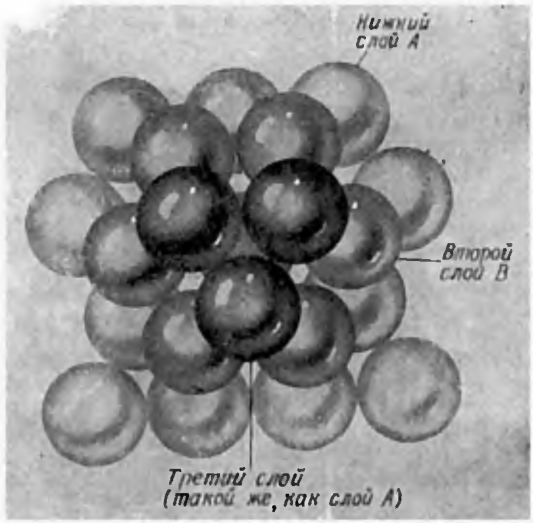


Рис. 1.27. Последовательность слоев в гексагональной плотноупакованной структуре.

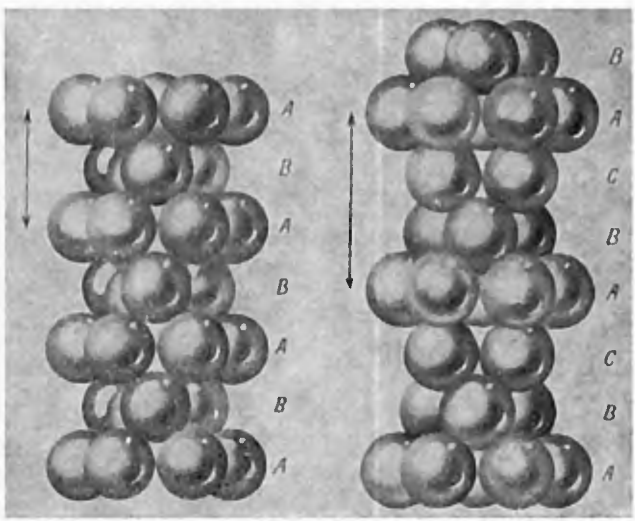


Рис. 1.28. Способы укладки слоев из плотноупакованных шаров. Слева — гексагональная плотноупакованная (г. п. у.) структура, справа — кубическая плотноупакованная (г. ц. к.) структура [*Pauling L.* The Nature of the Chemical Bond, Cornell, 1960].

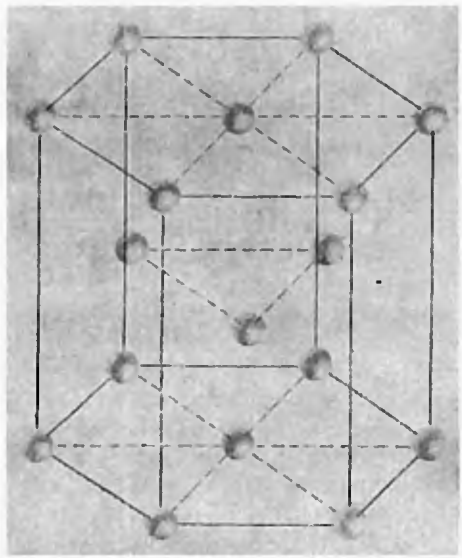


Рис. 1.29. Общепринятая элементарная ячейка г. п. у.-структуры. Каждый атом окружен двенадцатью ближайшими соседями, шесть в той плоскости, в которой находится этот атом, и по три атома в верхней и нижней плоскостях. Ячейка, показанная на рисунке, эквивалентна трем примитивным ячейкам.

Для сравнения структуры, изображенной в правой части рис. 1.28, с элементарной ячейкой г. ц. к.-решетки на рис. 1.15, *а* нужно выбрать правильное направление. Нижний левый атом на рис. 1.15, *а* необходимо поместить в плоскости *A*; затем диагональ элементарной ячейки следует направить таким образом, чтобы она проходила через плоскости *B* и *C* и заканчивалась опять в плоскости *A*, в которой должен располагаться верхний правый угол элементарной ячейки (рис. 1.15, *а*). Помимо осей *четвертого* порядка, очевидных для квадратов и кубов, кубические кристаллы обладают осями *третьего* порядка, направленными вдоль пространственных диагоналей. Симметрия относительно пространственных диагоналей представляет собой минимальную симметрию любого кубического кристалла.

Другие элементарные решетки с кубической симметрией

На первый взгляд может показаться удивительным, что наиболее компактной из кубических структур является г. ц. к.-структура. Если основываться на чисто умозрительных заключениях, то естественно предположить, что в объемно-центрированной кубической (о. ц. к.) структуре расположение одинаковых ша-

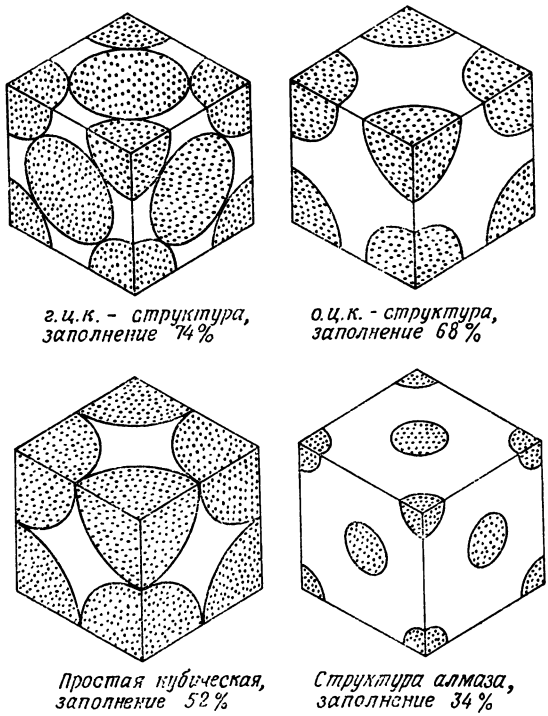


Рис. 1.30. Заполнение объема соприкасающимися шарами одинакового размера для четырех наиболее часто встречающихся кубических структур (*Wert C., Thomson R.* Physics of Solids, McGraw-Hill, 2nd ed., 1970. [Имеется перевод: *Уэрт Ч., Томсон Р.* Физика твердого тела.— М.: Мир, 1969.])

ров должно быть по крайней мере столь же компактным, как и в любой другой структуре. Ошибочность этого предположения становится ясной из рассмотрения рис. 1.30 и подтверждается геометрическим расчетом в задаче 1.12. Исходя из коэффициентов заполнения для различных кубических решеток, представленных на рис. 1.30, следует напомнить, что коэффициент заполнения для г. ц. к.-структуры, равный 74 %, является тем же самым, что и для идеальной г. п. у.-структуры, поскольку в обоих случаях мы имеем дело с максимальной упаковкой.

Каждый атом в объемно-центрированной кубической (о. ц. к.) структуре, показанной на рис. 1.15, *б* (в табл. 1.6 и на рис. 1.21 эта структура обозначена как «кубическая *I*»), имеет восемь ближайших соседей и соприкасается с ними только вдоль направлений пространственных диагоналей куба, что приводит к менее плотной упаковке. О. ц. к.-структуру имеют многие металлы, в том числе щелочные и некоторые из переходных.

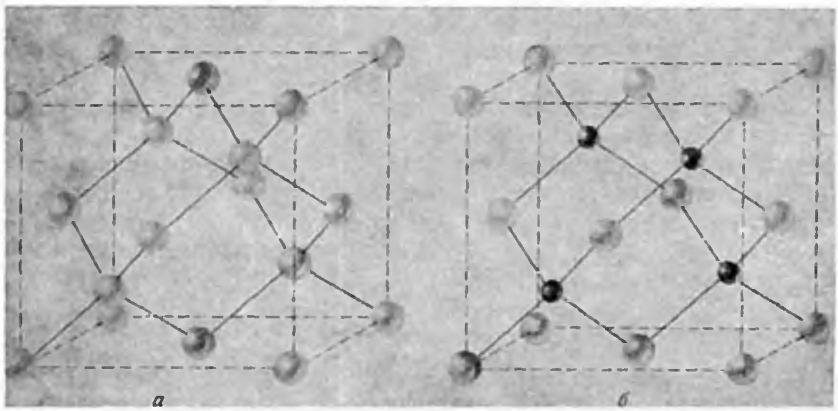


Рис. 1.31. а — расположение атомов в алмазе; б — расположение атомов в цинковой обманке (сфалерите). Обе структуры имеют трансляционную симметрию, присущую г. ц. к.-решетке с базисом, состоящим из двух атомов, расположенных в точках 000 и $^1/_4$ $^1/_4$ $^1/_4$. Таким образом, каждый атом окружен только четырьмя ближайшими соседями, расположенными в вершинах тетраэдра. В цинковой обманке все четыре ближайших соседа — это атомы другого сорта.

Еще менее плотная упаковка присуща *простой кубической* структуре (обычно обозначаемой либо как п. к., либо как кубическая Р). У одноатомных кристаллов простая кубическая решетка встречается достаточно редко. Из всех химических элементов она существует только у полония, и то эта структура оказывается устойчивой лишь в определенном интервале температур. Простую кубическую структуру имеют главным образом кристаллы с многоатомным базисом и простой кубической примитивной ячейкой [такие, как CsCl (рис. 1.33, *б*) и перовскит (рис. 1.36)].

Рассмотрение решетки алмаза с точки зрения плотности упаковки не имеет решающего значения (рис. 1.31, *а*). Для веществ с подобной структурой преобладающим фактором оказывается существование четырех симметрично расположенных валентных связей. Помимо алмаза, к таким веществам относятся кремний, германий и серое олово. Алмаз имеет трансляционную симметрию г. ц. к.-решетки, причем каждой узловой точке г. ц. к.-решетки Браве соответствует базис из двух атомов, расположенных в точках 000 и $^1/_4{}^1/_4{}^1/_4$.

Структуру алмаза можно представить как две г. ц. к.-решетки, вставленные одна в другую и смещенные относительно друг друга вдоль диагонали куба на четверть ее длины. Такой способ рассмотрения структуры алмаза наглядно демонстрируется сравнением рис. 1.31, *а* и *б*. Светлые атомы в цинковой

обманке (рис. 1.31, *б*) принадлежат одной г. ц. к.-решетке, а четыре темных атома относятся к другой г. ц. к.-решетке, сдвинутой относительно первой на четверть диагонали куба. Ниже мы встретимся с другими решетками, которые удобно рассматривать как вставленные одна в другую простые подрешетки. Поскольку обе подрешетки алмаза состоят из одинаковых атомов, каждый атом окружен четырьмя ближайшими соседями — точно такими же атомами, расположенными в вершинах правильного тетраэдра,— и еще двенадцатью атомами во второй координационной сфере.

Структуры некоторых простых соединений

Многие соединения с определенной долей ионности связей и с не слишком большим отношением размеров аниона и катиона имеют тетраэдрическую конфигурацию ближайших соседей. На рис. 1.31, *б* изображена структура цинковой обманки (или сфалерита) — соединения ZnS (элементов II и VI групп). Такую же структуру имеют и соединения элементов III и V, а также I и VII групп. Сфалерит отличается от алмаза только тем, что две его г. ц. к.-подрешетки состоят из различных атомов; поэтому ближайшими соседями атомов цинка оказываются четыре атома серы и, наоборот, ближайшими соседями атома серы оказываются четыре атома цинка. Двенадцать ближайших соседей атома цинка во второй координационной сфере — также атомы цинка, поскольку во второй координационной сфере г. ц. к.-решетки содержится именно двенадцать соседей.

Уже в структуре алмаза отсутствовали многие свойства симметрии, характерные для примитивного куба; еще меньше элементов симметрии в структуре сфалерита. Поскольку центр инверсии у такой структуры отсутствует, плоскости (111) и $(\bar{1}\bar{1}\bar{1})$ не эквивалентны. В ZnS плоскость (111) составлена только из атомов цинка, в то время как плоскость $(\bar{1}\bar{1}\bar{1})$ — из атомов серы.

Ряд твердых тел, которые кристаллизуются так же, как и сфалерит, обладает и другой аллотропной формой — в виде гексагональной структуры, известной под названием структуры вюрцита. Из рис. 1.32 мы видим, что в вюрците *ближайшие* соседи расположены так же, как в сфалерите. Образование гексагональной структуры с энергией связи, лишь немного отличающейся от энергии связи кубического кристалла, обусловлено изменением расположения атомов во второй координационной сфере.

Структуры, промежуточные между сфалеритом (кубическая структура) и вюрцитом (гексагональная структура), об-

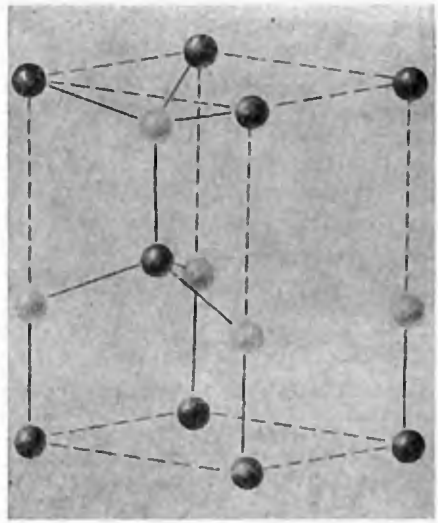


Рис. 1.32. Гексагональная элементарная ячейка для структуры вюрцита.

наружены в различных *политипах* карбида кремния[24]. Существование этих политипов обусловлено разнообразием способов периодической укладки последовательностей атомных плоскостей, перпендикулярных оси **c**. Так, хорошо изученный политип $15R$[25], имеющий ромбоэдрическую симметрию, образуется благодаря периодическому повторению последовательности *ABCACBCABACABCB* из пятнадцати плоскостей. Примитивный базис этого политипа состоит из двадцати атомов, а в элементарной ячейке имеется пять неэквивалентных положений для атомов углерода или кремния. Электронные свойства политипов SiC зависят от того, какая симметрия преобладает в структуре: кубическая или гексагональная.

Соединения с сильно выраженной ионной связью и одинаковыми числами анионов и катионов кристаллизуются в решетках типа NaCl и CsCl (рис. 1.33). Как уже отмечалось выше (см. пояснения к рис. 1.10), выбор энергетически наиболее благоприятной кристаллической структуры определяется соотношением радиусов аниона и катиона и в такой структуре должна быть исключена возможность контакта аниона с анионом.

Точно так же как кубическая структура ZnS состоит из г. ц. к.-решетки цинка и г. ц. к.-решетки серы, вставленных друг в друга таким образом, что они смещены одна относительно другой на *одну четверть* вдоль диагонали куба, так и кубиче-

[24] *Krishna P., Verma A. R.*— Proc. Roy. Soc., **A272**, 490 (1963); см. также: Silicon Carbide — 1968, eds. H. K. Henish, R. Roy, Pergamon Press, 1969.

[25] *Patrick L. et al.*— Phys. Rev., **132**, 2023 (1963).

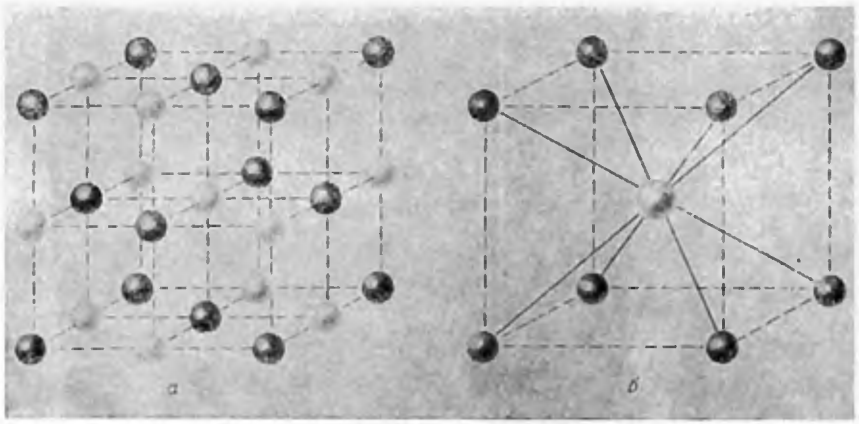


Рис. 1.33. Две кристиллические структуры, характерные для ионных соединений, с равным количеством анионов и катионов. *а* — структура хлорида натрия (каменной соли); *б* — структура хлорида цезия.

ская структура NaCl состоит из г. ц. к.-решетки натрия и г. ц. к.-решетки хлора, сдвинутых относительно друг друга на *половину* диагонали куба. Иными словами, NaCl имеет трансляционную симметрию г. ц. к.-решетки с базисом, состоящим из атома Na, расположенного в точке 000, и атома хлора, расположенного в точке $^1/_2\,^1/_2\,^1/_2$. Кроме большого числа гидридов и галогенидов щелочных металлов структура каменной соли свойственна галогенидам серебра, халькогенидам свинца (таким, как PbS, PbSe, PbTe и т. д.) и многочисленным соединениям элементов II—VI групп.

Структура хлорида цезия (рис. 1.33, *б*) внешне похожа на структуру о. ц. к.-решетки, но в действительности ее пространственная решетка является простой кубической с двухатомным базисом, состоящим из атома Cs в точке 000 и атома Cl в точке $^1/_2\,^1/_2\,^1/_2$. Эту структуру называют также структурой β-латуни, поскольку такая структура свойственна не только некоторым кристаллам с большой степенью ионности, но также фазам в системе медь—цинк и ряду других металлических сплавов.

Кубическую структуру часто имеют соединения, в которые химические элементы не входят в одинаковой пропорции. Примеры таких соединений показаны на рис. 1.34—1.36. Структура закиси меди (рис. 1.34) довольно необычна: атомы кислорода образуют о. ц. к.-решетку, а атомы меди находятся на серединах отрезков, соединяющих центр куба с его вершинами. Упражнением для изучения этой структуры является задача 1.13.

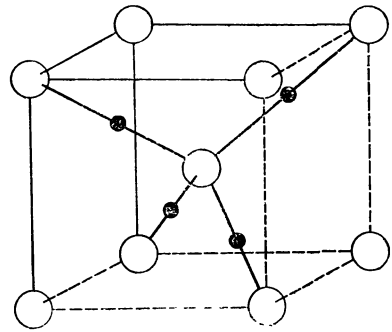


Рис. 1.34. Необычная кубическая структура закиси меди (Cu_2O).

Из рис. 1.34 видно, что каждый атом кислорода окружен атомами меди, расположенными в вершинах тетраэдра, так что диагонали куба оказываются осями вращательной симметрии третьего порядка. Как уже отмечалось выше, это свидетельствует о наличии кубической симметрии данного кристалла. Кубической симметрией обладают и изображенные на рис. 1.35 кристаллы SiF_4 и $MoAl_{12}$; каждый из этих кристаллов имеет о. ц. к.-структуру.

Из приведенных примеров видно, что наличие у кристалла кубической симметрии определяется не столько простотой или сложностью его химической формулы, сколько направлениями связей внутри молекулы. Перовскит (рис. 1.36) — это другой

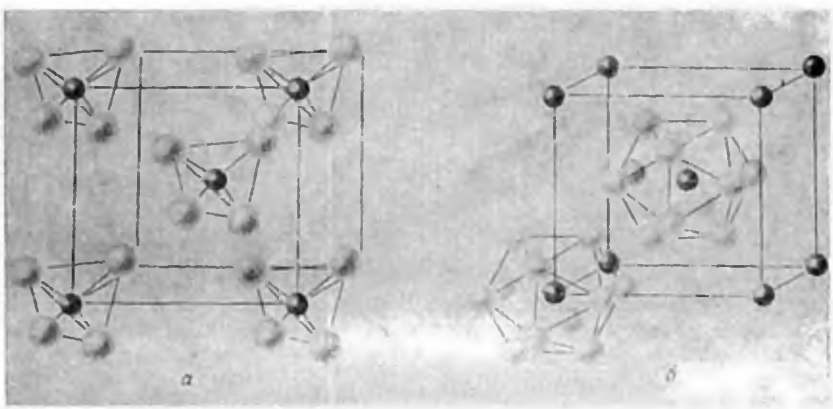


Рис. 1.35. Примеры соединений с пространственной о. ц. к.-решеткой. *а* — структура тетрафторида кремния — хороший пример реализации кубической симметрии при наличии тетраэдрической симметрии вдоль направления объемных диагоналей куба; *б* — структура $MoAl_{12}$ и WAl_{12}, в которой каждая группа из двенадцати атомов алюминия образует правильный икосаэдр [*Pauling L.* The Nature of the Chemical Bond, Cornell, 1960]

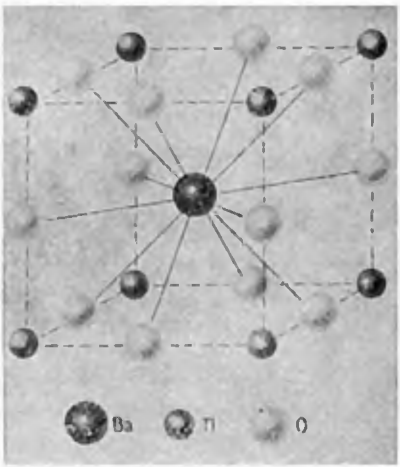


Рис. 1.36. Структура перовскита на примере $BaTiO_3$. При высоких температурах структура является кубической (показано на рисунке), а при охлаждении ниже температуры сегнетоэлектрического перехода она немного изменяется и приобретает черты тетрагональной структуры.

пример вещества с относительно сложной молекулой, пространственное строение которой *может* соответствовать симметрии кубической решетки. Сам перовскит (титанат кальция) обладает кубической симметрией при любой температуре; однако в соединениях типа $BaTiO_3$, $PbTiO_3$ и $KNbO_3$ при охлаждении ниже температуры сегнетоэлектрического перехода[26] происходит незначительное смещение к тетрагональной форме.

В качестве последнего примера рассмотрим структуру молекулы K_2PtCl_4 (рис. 1.37), которая иллюстрирует то, что при такой молекулярной форме кристалл должен иметь тетрагональную, а не кубическую структуру. Наиболее существенными элементами симметрии этого соединения являются оси вращения четвертого порядка, параллельные оси **c**.

Кристаллы с более сложной и менее симметричной структурой анализируются с помощью дифракции рентгеновских лучей и других методов. Заинтересованный читатель может найти примеры структур, соответствующих большинству 230 теоретически возможных пространственных групп *, в литературе, при-

[26] Сегнетоэлектрические свойства, а также различные следствия наличия в решетке постоянного электрического дипольного момента мы рассмотрим в гл. 5. Смещение к тетрагональной структуре заключается в том, что вся подрешетка кислорода сдвигается вверх (или вниз) относительно подрешетки бария и титана. В действительности при охлаждении существенно ниже температуры сегнетоэлектрического перехода кристалл титаната бария претерпевает последовательные искажения сначала в направлении к моноклинной, а затем к ромбоэдрической форме [*Merz W. J.*— Phys. Rev., 76, 1221 (1949).]

* Эти пространственные группы были исследованы в 1890 г. Е. С. Федоровым и поэтому называются федоровскими группами — *Прим. ред*

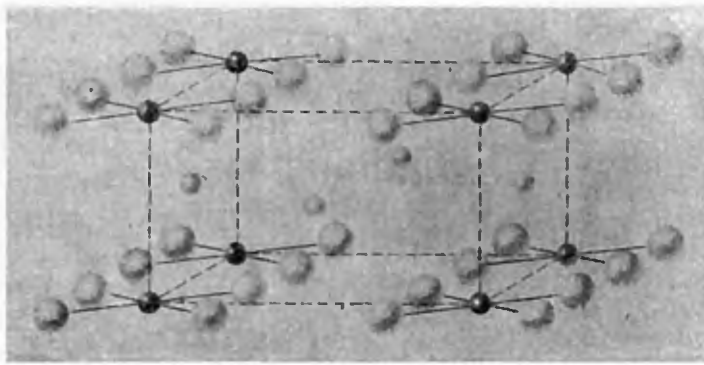


Рис. 1.37. Элементарная ячейка кристалла K_2PtCl_4. Каждый атом платины (маленькие темные кружки) окружен четырьмя атомами хлора (большие кружки), расположенными в вершинах квадрата. Кружки среднего размера — атомы калия. Пространственная решетка примитивная, тетрагональная [*Pauling L.* The Nature of the Chemical Bond, Cornell, 1960].

веденной в конце данной главы. (Свойства всех пространственных групп описаны с помощью теории групп в книге под редакцией Зака.) Кристаллические структуры многих химических элементов имеют довольно сложное строение. Примерами таких сложных структур являются графит и теллур (см. рис. 1.6 и 1.7). Типичные структуры химических соединений, состоящих из атомов двух, трех или четырех сортов, на самом деле обладают значительно *более* высокой симметрией.

1.4. Дифракция в кристаллах

Дифракция как метод исследования

Полезную информацию о структуре некоторых кристаллических твердых тел можно получить из макроскопических и микроскопических исследований внешней морфологии образца. Кроме того, с помощью методов, основанных на отражении видимого света от поверхности образца, можно определить кристаллографическую *ориентацию* монокристалла с известной структурой. Однако разрешающей способности видимого света и ультрафиолетового излучения совершенно недостаточно для определения положений атомов и молекул в кристалле, а также расстояний между ними. Обычно эти расстояния порядка 1 Å (10^{-10} м).

Микроскопическое исследование в общем случае подразумевает использование излучения с длиной волны, меньшей, чем

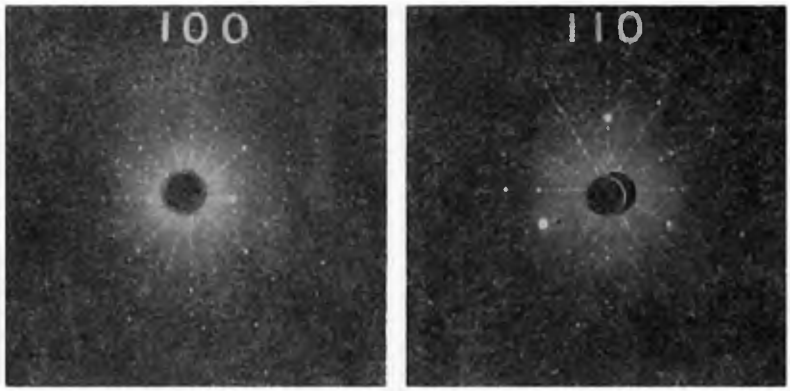


Рис. 1.38. Дифракционные пятна на лауэграммах монокристаллов пирита FeS_2 с ориентацией (100) и (110). (Фото любезно представлено Л. Е. Марром из Орегонского учебного центра).

расстояния, которые нужно определить, так что при этом можно пользоваться законами геометрической оптики. Для исследований кристаллической структуры наиболее общепринято использовать *дифракционные* методы, когда длина волны должна быть сравнима с размерами атомов. Фурье-анализ данных, полученных дифракционными методами, позволяет определить средние расстояния между рядами атомов и атомными плоскостями в кристалле, углы между ними, симметрию точечных групп, а также (при определенном навыке) координаты отдельных атомов. Как уже отмечалось во введении к настоящей главе, мы рекомендуем прочитать сначала разд. 1.4 и 1.5, а затем уже приступить к подробному изучению дифракции и обратного пространства. Эти оба вопроса тесно связаны между собой.

В 1912 г. Лауэ высказал предположение, что кристаллы можно рассматривать как трехмерную дифракционную решетку для рентгеновских лучей, длина волны которых сравнима с межатомными расстояниями, и что на основании дифракционной картины можно судить о правильном расположении атомов в кристалле. Метод дифракции рентгеновских лучей и по сей день остается основным источником информации о структуре кристаллов, хотя в настоящее время для этих же целей используется также дифракция электронов и нейтронов. На рис. 1.38 показана типичная лауэграмма монокристалла.

Для того чтобы длины волн всех трех типов квантовых частиц (фотонов, электронов и нейтронов), используемых в кристаллографии, соответствовали поставленной задаче, их энер-

гия должна быть различной. Для фотонов рентгеновского излучения мы имеем следующее соотношение:

$$\lambda = c/\nu = hc/E, \tag{1.30}$$

откуда следует, что длине волны 1 Å соответствует энергия $E \sim 12\,000$ эВ. Той же длине волны электронов будет соответствовать меньшая энергия. Согласно соотношению де Бройля,

$$\lambda = h/mv = h/(2mE)^{1/2}, \tag{1.31}$$

так что длине волны 1 Å в этом случае отвечает $E \sim 150$ эВ. Электрон с такой энергией и длиной волны движется со скоростью $v \sim 7 \cdot 10^6$ м/с.

Дифракция нейтронов на упорядоченной атомной структуре кристалла требует еще меньших значений энергии и скорости:

$$\lambda = h/M_n v = h/(2M_n E)^{1/2}. \tag{1.32}$$

Таким образом, нейтроны с длиной волны 1 Å движутся со скоростью всего лишь 4000 м/с и имеют кинетическую энергию 0,08 эВ. Эта энергия сравнима с энергией кванта колебаний кристаллической решетки, что ведет к сильному взаимодействию тепловых нейтронов с колебаниями решетки, как мы покажем ниже.

Использование трех типов излучения

Указанные выше три типа излучения применяют для получения различной, хотя и перекрывающейся информации относительно структуры твердых тел. Взаимодействие всех трех типов излучения с периодической решеткой подчиняется одним и тем же общим геометрическим законам, которые мы обсудим в следующем разделе.

Рентгеновское излучение с энергией фотонов 10—100 кэВ способно проникать глубоко внутрь кристалла, и поэтому его широко применяют в большом числе методов анализа неизвестных трехмерных структур. Однако, поскольку рентгеновские лучи рассеиваются главным образом электронами атомных оболочек, их использование для определения положений очень легких атомов типа водорода не приводит к столь же хорошим результатам, как для тяжелых элементов, причем протонный состав вещества не оказывает влияния на результаты таких исследований. В рентгеновской кристаллографии нас интересует рассеяние на каждом электроне, движущемся в фазе с падающей электромагнитной волной и излучающем *когерентно*

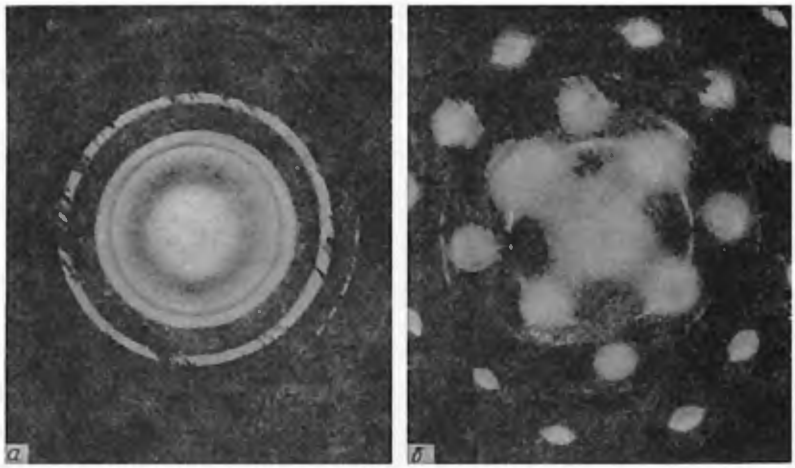


Рис. 1.39. Дифракционные картины рассеяния электронов на тонких пленках, напыленных на сколотую (001) поверхность монокристалла NaCl. *а* — рассеяние на мелкозернистом поликристаллическом Pd с г. ц. к.-решеткой; преимущественная ориентация зерен отсутствует, видны равномерено освещенные дифракционные кольца; *б* — рассеяние на ванадии с о. ц. к.-решеткой, напыленном на нагретую подложку; структура имеет в основном вид монокристаллической эпитаксиальной пленки с ориентацией (001), о чем свидетельствует наличие отдельных дифракционных пятен; дифракционные кольца обусловлены наличием поликристаллических областей со случайной ориентацией. (Фото любезно предоставлено Л. Е. Марром из Орегонского учебного центра).

без изменения длины волны[27]. Интерференция когерентно рассеянного излучения с усилением или ослаблением определяется геометрией атомных конфигураций в кристалле.

Наблюдение дифракции электронов на монокристаллах (Девиссон и Джермер, 1927) или кристаллических пленках (Томпсон, 1927) убедительно подтверждает как наличие периодической структуры у кристаллов, так и корпускулярно-волновые свойства электрона (рис. 1.39). Поскольку электрон является заряженной частицей, он испытывает сильное взаимодействие с веществом и, прежде чем подвергнется упругому или неупругому столкновению, успевает проникнуть в кристалл лишь на

[27] Рассеяние электроном электромагнитной волны без изменения частоты (т. е. без изменения энергии) впервые было описано Дж. Дж. Томсоном и называется иногда *томсоновским* рассеянием. Существует также *неупругое* рассеяние — эффект Комптона,— конкурирующее с интересующим нас когерентным рассеянием. Неупругое рассеяние приводит к образованию фона некогерентных фотонов с любыми углами рассеяния и энергиями, случайным образом смещенными в область более низких значений. Отделение этого фона от фотонов с начальной энергией, когерентно рассеянных под брэгговскими углами, представляет собой одну из главных проблем в структурном анализе.

несколько сотен ангстрем. Таким образом, для изучения структур большого объема метод дифракции электронов не может конкурировать с рентгеновскими методами. Однако его можно с успехом использовать в следующих двух случаях:

1) при изучении поверхностных слоев и состояния поверхности кристаллов; характер перехода от идеальной периодической решетки к незаполненному пространству вне кристалла зависит как от способа обработки поверхности, так и от степени ее чистоты;

2) при изучении тонких пленок; достаточно тонкую пленку можно рассматривать как два поверхностных слоя, пренебрегая объемом между ними; дифракция электронов оказалась весьма полезной для выявления отличий структурных свойств тонких слоев вещества от свойств «идеального» монокристалла.

Медленные нейтроны, у которых длина волны де Бройля сравнима с межатомными расстояниями, могут взаимодействовать с твердыми телами несколькими способами. В немагнитных кристаллах они взаимодействуют только с ядрами (поскольку нейтрон не имеет заряда, а масса его намного больше, чем у электронов атомных оболочек). Если происходящее при этом рассеяние упругое и когерентное, то оно подчиняется тем же геометрическим законам, что и рассеяние рентгеновских лучей и электронов. Дифракционные картины для нейтронов и рентгеновских лучей одинаковы, за исключением лишь того, что нейтроны более эффективно рассеиваются на легких атомах.

В присутствии нескольких изотопов данного элемента, случайно распределенных по всем воможным положениям атомов, помимо когерентного рассеяния, может также происходить упругое, но некогерентное рассеяние нейтронов. Однако по сравнению с этим больший интерес представляет процесс *неупругого* рассеяния нейтронов, обусловленный нарушением строгой периодичности расположения ядер. В гл. 2 мы вернемся к обсуждению неупругого рассеяния нейтронов, поскольку оно лежит в основе весьма важного метода исследования спектра колебаний решетки кристалла. Нейтрон может либо поглотить, либо испустить фонон[28], так что его энергия и направление

[28] По аналогии с названием *фотон* для кванта электромагнитного излучения, названия квантов различных элементарных возбуждений в кристалле также оканчиваются на «*он*». ***Фонон*** — квант колебаний решетки, т. е. периодического смещения атомов относительно положения равновесия, которому можно приписать энергию, импульс, а также скорость, длину волны и волновой вектор. *Магнон* — магнитная спиновая волна, т. е. возбуждение в ферромагнитных кристаллах спинов, строго параллельных друг другу в основном состоянии. *Плазмон* — коллективное возбуждение электронного газа, т. е. когерентные колебания всех электронов, обусловленные дальнодействующим характером кулоновских сил.

движения в соответствии с законами сохранения энергии и импульса изменяются. Энергия нейтрона с длиной волны, необходимой для дифракции на кристалле, сравнима с энергией фононов в твердых телах. Таким образом, изменения энергии и направления движения нейтрона после неупругого соударения значительны и легко измеряются.

Наличие у нейтрона магнитного момента приводит к дополнительному усложнению характера дифракции нейтронов в магнетиках. Сечение взаимодействия нейтронов с магнитоупорядоченными электронами в магнетиках сравнимо с сечением взаимодействия нейтронов с ядрами. Таким образом, упругое рассеяние нейтронов может служить источником информации о распределении магнитных моментов.

Условие дифракции Брэгга

Рассмотрим, какому геометрическому условию *необходимо* удовлетворить, чтобы кристалл мог эффективно рассеивать коллимированное монохроматическое излучение. Это условие должно удовлетворяться для любого типа излучения — рентгеновского, пучка электронов или нейтронов (а также для других частиц с корпускулярно-волновыми свойствами).

В 1913 г. У. Л. Брэгг предложил простую формулировку и получил простое математическое выражение необходимого условия дифракции волн на серии параллельных (атомных) плоскостей. Это условие нельзя считать строгим, поскольку Брэгг рассматривал отражение от каждой атомной плоскости в отдельности, используя законы геометрической оптики без учета эффектов физической оптики (дифракции Фраунгофера), ведущих к дифракционному усилению отраженной волны. Тем не менее следует заметить, что формула (1.33), полученная Брэггом, совпадает с результатами более точных рассмотрений Лауэ [выражение (1.55)] и Эвальда [выражение (1.52)]. Последний использовал в доказательстве понятие обратной решетки (см. разд. 1.5).

Следует сразу же обратить внимание на одно важное различие между дифракцией света на плоской дифракционной решетке и дифракцией рентгеновских лучей на кристалле. В первом случае угол падения не равен углу, под которым распространяется дифрагированное излучение. Существует взаимосвязь между этими двумя углами, длиной волны падающего света λ и расстоянием между соседними штрихами *плоской дифракционной решетки.* Согласно условию дифракции Брэгга, угол падения равен углу отражения (т. е. отражение является зеркальным). Из этого условия следует также, что максимальное отражение наступает при выполнении опреде-

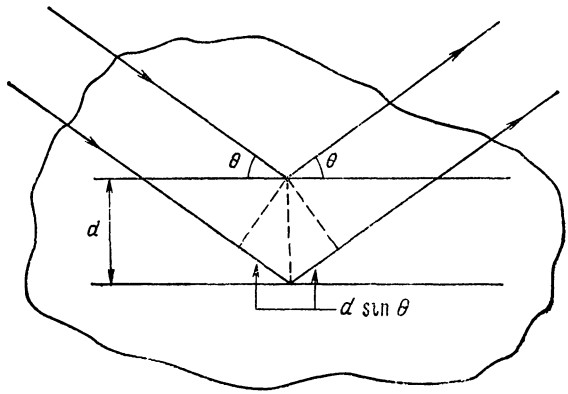


Рис. 1.40. Зеркальное отражение параллельного пучка рентгеновских лучей (нейтронов и т. д.) от двух последовательных атомных плоскостей в кристалле. Первая плоскость не обязательно должна совпадать с границей кристалла и может быть не параллельна ей; важно лишь, чтобы эта плоскость имела достаточно большую плотность атомов. Распределение атомов в плоскости несущественно. Для интерференционного усиления разность хода лучей должна быть равна целому числу длин волн.

ленных соотношений между углом падения, длиной волны и расстоянием между двумя параллельными плоскостями. При этом закон Брэгга совершенно не учитывает расположения атомов в отдельной отражающей плоскости.

Предположим, что на кристалл падает пучок параллельных монохроматических рентгеновских лучей (рис. 1.40), которые взаимодействуют со всеми атомами в объеме, соответствующем глубине их проникновения в кристалл. Рассмотрим взаимодействие рентгеновского излучения с семейством параллельных атомных плоскостей, составляющих с пучком угол θ. В результате взаимодействия с первой из этих плоскостей появляется зеркально отраженный под углом θ луч, который будет весьма слабым, если рентгеновские лучи проникают достаточно глубоко внутрь образца. Кроме того, поскольку первая плоскость имеет периодическую структуру, она, очевидно, действует на рентгеновское излучение как двумерная решетка, благодаря чему появятся слабые лучи, отраженные от первой плоскости под углами φ_m, соответствующими различным порядкам дифракции.

Для второй и всех последующих плоскостей также появятся компоненты под углами, соответствующими зеркальному отражению и дифракционному отражению различных порядков. Поскольку рентгеновские лучи могут проникать очень глубоко в кристалл и выявлять при этом отдельные слабые отклики

от нескольких тысяч следующих друг за другом плоскостей, очевидно, что вследствие интерференции с ослаблением для большинства углов падения рентгеновского излучения на множество рассматриваемых плоскостей не возникнет четко определенного отраженного пучка.

Однако *если* разность хода между лучами, отраженными от различных плоскостей, кратна длине волны падающего излучения, то будет иметь место интерференция с усилением. Из рис. 1.40 видно, что разность хода между лучами, отраженными от первой и второй плоскостей, равна $2d\sin\theta$. Таким образом, условие *эффективного* зеркального отражения (условие Брэгга) запишется в виде

$$2d\sin\theta = n\lambda, \tag{1.33}$$

где n — целое число. Как уже отмечалось выше, это соотношение зависит только от расстояния между плоскостями и не зависит от расположения атомов в каждой из плоскостей.

Естественно, возникает вопрос, можно ли записать условие, обеспечивающее интерференцию с усилением лучей, отраженных под одним из углов φ_m, соответствующих дифракционным максимумам m-го порядка при отражении от первой плоскости. В уравнение, определяющее это условие, должны входить как расстояние между плоскостями d, так и расстояние между рядами атомов на плоскости. Такое условие действительно существует, и его можно рассматривать как условие Брэгга * для зеркального отражения от некоторого *другого* семейства параллельных плоскостей, расположенных по отношению к исходным под определенным углом. Этот вопрос рассматривается в задаче 1.14.

Из сказанного выше ясно, что монохроматическое рентгеновское излучение, падающее на кристалл под произвольными углами, вообще говоря, отражаться не будет. Необходимо найти такую комбинацию длины волны и угла падения, для которой условие Брэгга выполняется, набор таких последовательных комбинаций позволяет получить информацию о пространственном расположении плоскостей и, следовательно, представление о структуре конкретного кристалла. Для однозначного определения расстояния d между плоскостями с индексами Миллера (hkl) нужно иметь как можно больше данных.

Нетрудно показать (что и требуется в задаче 1.15), что для *кубической* решетки с элементарной ячейкой размером a в каж-

* Это условие называется также условием Брэгга — Вульфа, поскольку независимо от Брэгга тоже в 1913 г. оно было установлено Г В Вульфом.— *Прим. ред.*

дом направлении расстояние между соседними плоскостями с индексами (hkl) равно

$$d = a(h^2 + k^2 + l^2)^{-1/2}. \quad (1.34)$$

Для того чтобы получить компактное выражение, эквивалентное (1.34), в более общем случае трехмерной пространственной решетки, необходимо использовать обозначения обратной решетки, о которых пойдет речь в разд. 1.5.

Как отмечалось выше, условие Брэгга является необходимым, но недостаточным условием эффективного зеркального отражения от кристалла. Качество дифракционной картины, получающейся при отражении от определенных плоскостей (hkl), зависит как от *атомного фактора рассеяния* (или *формфактора*) для атомов, из которых состоят эти плоскости, так и от того, разрешено ли отражение с точки зрения *структурного фактора*.

Атомный фактор рассеяния

Дифракция рентгеновских лучей на кристалле обусловлена взаимодействием (томсоновским рассеянием) фотонов рентгеновского излучения длиной волны λ с полем заряда всех внеядерных электронов. Нам теперь известно, что рассеянное излучение интерферирует с усилением лишь при выполнении условия Брэгга. Интенсивность дифрагированного излучения будет значительной только в том случае, если геометрический структурный фактор отличен от нуля, а значение атомного фактора рассеяния близко к максимальному для данного кристалла.

Согласно классической томсоновской теории когерентного рассеяния электромагнитной волны свободным электроном, эффективность рассеяния незначительно уменьшается с увеличением угла 2θ между падающей и отраженной волнами. Однако эффективность когерентного рассеяния рентгеновских лучей электронами кристалла спадает гораздо быстрее с увеличением угла 2θ. Это происходит из-за того, что электроны, связанные с каждым атомом, распределены в объеме, линейные размеры которого сравнимы с длиной волны рентгеновского излучения.

Для рентгеновских лучей с большой длиной волны и (или) малым углом рассеяния атом с порядковым номером Z действует как томсоновский точечный рассеиватель с зарядом Ze. В случае меньших длин волн и (или) бóльших углов рассеяние на атоме происходит так же, как на томсоновском источнике с эффективным зарядом fe. Число f называется *атомным фак-*

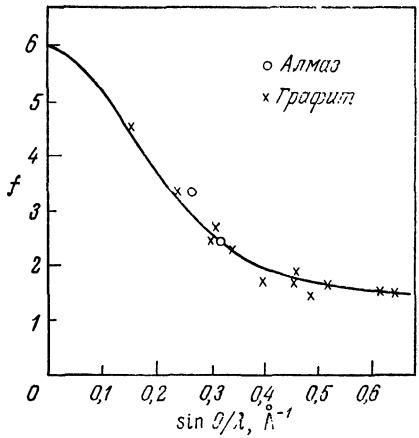


Рис. 1.41. Сравнение экспериментальных данных [*Bacon G. E.*— Acta Cryst., 5, 492 (1952)] для атомного фактора рассеяния рентгеновских лучей углеродом с кривой, вычисленной Мак-Уини [*McWeeny R.*— Acta Cryst., 4, 513 (1951)] для самосогласованного распределения электронов. Различные значения на оси абсцисс получены путем изменения угла Брэгга при постоянной длине волны λ. Экспериментальные точки соответствуют различным плоскостям (*hkl*) низкого порядка.

тором рассеяния или *формфактором* атома. С ростом величины $(\sin\theta)/\lambda$ число f все больше уменьшается по сравнению с Z.

Обозначим через $\psi_n(r)$ пространственную зависимость нормированной волновой функции n-го электрона, принадлежащего атому с порядковым номером Z. Тогда в случае рассеяния рентгеновских лучей с длиной волны λ можно написать следующее выражение для формфактора через угол 2θ[29]:

$$f = \sum_{n=1}^{Z} \int_0^{\infty} \frac{4\pi r^2 \psi_n^2(r) \sin[4\pi r(\sin\theta)/\lambda]\, dr}{4\pi r(\sin\theta)/\lambda}. \qquad (1.35)$$

При стремлении $(\sin\theta)/\lambda$ к нулю каждый член ряда стремится к единице, что обеспечивает выполнение в пределе малых углов требуемого условия $f = Z$.

Атомный фактор рассеяния можно вычислить для изолированного атома по известной плотности электронного заряда, найденной из распределения самосогласованного поля в методе Хартри — Фока. Хотя межатомное взаимодействие несколько меняет распределение заряда наиболее удаленных от ядра электронов, атомные факторы рассеяния в кристалле, как правило, достаточно близки к значениям, даваемым теорией Хартри — Фока. Это видно из рис. 1.41, на котором экспериментальные значения фактора рассеяния сравниваются с кривой, полученной на основании модели самосогласованного поля. Заметим, что такую кривую можно получить, изменяя угол

[29] См., например, гл. 1 в книге: *Warren B. E. X*-Ray Diffraction, Addison-Wesley, 1969.

Брэгга, а не длину волны рентгеновских лучей. Ниже мы покажем, что атом углерода с $Z=6$ когерентно рассеивает так же, как точечный заряд, равный шести электронным зарядам, длинноволновое излучение при малых углах. Поскольку масса заряженного рассеивателя входит в знаменатель выражения для томсоновского рассеяния, когерентным рассеянием рентгеновских лучей атомными ядрами можно пренебречь.

Структурный фактор

При написании условия Брэгга мы уже подчеркивали, что выражение (1.33) является лишь *необходимым* условием эффективного когерентного отражения рентгеновских лучей семейством параллельных плоскостей (hkl). Но, даже если угол падения равен углу отражения и для некоторого целого n выполняется условие Брэгга, интенсивность отраженного луча (если такой луч вообще существует) зависит от величины *структурного фактора* F_{hkl} для плоскостей (hkl), который может быть большим, малым или равным нулю. Структурный фактор зависит от формфакторов атомов разного типа, имеющихся в кристалле, и их положений в элементарной ячейке. Поскольку базис большинства кристаллов является многоатомным, вполне возможно, что усиленное брэгговское отражение от одной серии атомов (для некоторых hkl) находится в противофазе с когерентным рассеянием от других атомов базиса и, следовательно, должно ослабляться.

Выражение для структурного фактора многоатомного базиса (и, разумеется, само условие Брэгга) удобно записывать на языке обратного пространства. Однако, и не обращаясь к понятию обратного пространства, которое мы рассмотрим в следующем разделе, можно понять принцип, на основании которого определяется структурный фактор F_{hkl}. Рассмотрим кристалл с базисом из N атомов, и пусть f_j — формфакор j-го атома. Представим себе на мгновенье, что все N атомов лежат точно *на* плоскостях (hkl), разделенных расстоянием d_{hkl}. При этом наиболее интенсивным должно было бы быть брэгговское отражение первого порядка для рентгеновских лучей с длиной волны $\lambda=2d_{hkl}\sin\theta_{hkl}$. Волна, отраженная от любой плоскости, отстает по фазе точно на 2π от волны, отраженной от предыдущей плоскости. Амплитуда отраженной волны будет пропорциональна сумме формфакторов f_j всех N атомов каждого базиса.

Поскольку в реальных кристаллах некоторые атомы базиса находятся над соответствующими плоскостями (hkl), волны, рассеянные на отдельных атомах, будут *не* полностью сфазированы. Волна, отраженная от j-го атома базиса, расположен-

ного на Δd_j выше плоскости, будет опережать волну, отраженную от самой плоскости (hkl), на угол (в радианах)

$$\varphi_j = 2\pi\Delta d_j/d_{hkl}. \tag{1.36}$$

Результирующая амплитуда брэгговской отраженной волны равна векторной сумме волн, отраженных от каждого атома базиса, и пропорциональна, таким образом, *структурному фактору*

$$F_{hkl} = \sum_{j=1}^{N} f_j \exp i\varphi_j. \tag{1.37}$$

В качестве простого примера того, как два атома базиса могут (частично или полностью) скомпенсировать друг друга при брэгговском отражении, рассмотрим дифракцию на кристалле CsCl, элементарная решетка которого представлена на рис. 1.33, *б*, при отражении от плоскости (100). Атомы цезия образуют простую кубическую решетку, и очевидно, что направление рентгеновских лучей и длину волны можно выбрать таким образом, чтобы разность хода компонент, отраженных от соседних плоскостей (100), была равна λ (т. е. разность фаз была бы равна 2π радиан). Все отраженные при этом компоненты будут складываться в фазе. Однако ионы хлора в кристалле CsCl также образуют простую кубическую решетку, у которой плоскости (100) расположены строго посередине между плоскостями (100) подрешетки цезия. Если комбинация λ и θ соответствует брэгговскому отражению от плоскости (100), то компоненты, отраженные от плоскостей, образованных атомами хлора, будут отличаться по фазе от соответствующих компонент, отраженных от плоскостей подрешетки цезия, на π радиан. Возникающая при этом интерференция с ослаблением приводит к тому, что интенсивность брэгговского отражения от плоскости (100) останется конечной лишь благодаря тому, что атомные факторы рассеяния у цезия и хлора различны. Интерференционное гашение вызывает уменьшение интенсивности рассеяния рентгеновских лучей и при дифракции на других плоскостях (hkl) в этой структуре. Аналогичным образом за счет структурного фактора уменьшается интенсивность дифрагированных компонент во многих соединениях.

Возвращаясь к структуре CsCl, заметим, что если атомный фактор рассеяния у атома в центре элементарной ячейки такой же, как у атомов в ее вершинах, то брэгговское отражение от плоскости (100), равно как и от определенных плоскостей более высоких порядков, исчезает полностью. Структура CsCl с кубической ячейкой, в центре которой помещен такой же атом, как и в ее вершинах, представляет собой не что иное, как о. ц. к.-решетку, изображенную на рис. 1.15, *б*. Существует

правило, согласно которому *структурный фактор* о. ц. к.-решетки равен нулю, если сумма $h+k+l$ равна целому нечетному числу. В кристаллах с г. ц. к.-решеткой структурный фактор для брэгговского отражения от плоскости (hkl) не равен нулю, только если числа h, k и l либо все четные, либо все нечетные (см. задачу 1.16). Эти сложности с правилами для структурного фактора в г.ц.к.- и о.ц.к.-решетках обусловлены тем, что мы посчитали удобным характеризовать эти решетки непримитивными элементарными ячейками.

Экспериментальные методы рентгенографического исследования структуры кристаллов

Читателя, интересующегося подробным описанием экспериментальных методов исследования структуры кристаллов с помощью дифракции рентгеновских лучей и нейтронов, мы отсылаем к литературе, указанной в конце данной главы. В настоящее время разработано более десятка методов для изучения структурных особенностей твердых тел с помощью рентгеновской дифракции, причем большинство из них применимо также и в нейтронной дифракции, хотя используемое при этом экспериментальное оборудование будет совсем другим. Следует заметить, что из рентгенографических методов, которые кратко упоминаются в этом разделе, в нейтронографии почти совсем не используется метод Лауэ, в то время как широко применяются метод вращения кристалла и метод порошка, в которых используются коллимированные пучки монохроматических нейтронов. В следующих разделах мы рассмотрим экспериментальные методы, связанные с применением рентгеновских лучей.

При рассмотрении брэгговской дифракции пучка рентгеновских лучей на монокристалле следует заметить, что для получения интерференции с усилением и эффективного отражения в соответствии с условием (1.33) необходимо иметь определенную комбинацию величин d, θ и λ. Поэтому монохроматические рентгеновские лучи, падающие на кристалл под произвольным углом, обычно почти не отражаются. Нужно либо использовать рентгеновские лучи в большом диапазоне длин волн при фиксированном угле падения (метод Лауэ), либо иметь возможность менять угол, под которым монохроматические лучи входят в кристалл. Последнее условие выполняется автоматически при вращениях или колебаниях кристалла, а также в методе Вейзенберга. В методе Дебая — Шеррера используется порошкообразный образец, так что *любой* из углов падения может оказаться подходящим для некоторых кристаллитов.

Метод Лауэ весьма прост как с точки зрения осуществления его принципа, так и его экспериментального осуществления.

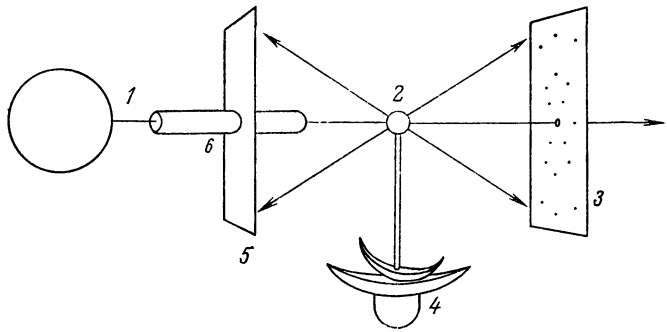


Рис. 1.42. Плоская камера Лауэ, использующая немонохроматическое рентгеновское излучение. Фотографии пятен Лауэ можно получать для рассеяния как вперед, так и назад. Образец может поворачиваться вокруг трех ортогональных осей; повторная фотосъемка производится для того, чтобы проверить правильность установленной ориентации. *1* — источник рентгеновского излучения; *2* — образец; *3* — положение фотопластинки при прямой фотосъемке; *4* — гониометрическое устройство; *5* — положение фотопластинки при обратной фотосъемке; *6* — коллиматор.

Узкий пучок рентгеновских лучей падает на монокристалл так, как показано на рис. 1.42. Рентгеновское излучение в этом пучке обладает широким диапазоном длин волн, и для любой длины волны, удовлетворяющей условию Брэгга, возникает дифракционный рефлекс. На рисунке показаны два положения фотопластинки для регистрации набора дифракционных «пятен», распределение которых характеризует структуру и ориентацию кристалла.

В случае когда структура кристалла заранее неизвестна, набор дифракционных пятен может с трудом поддаваться расшифровке, поскольку при отражении от одной и той же серии плоскостей в одну и ту же точку фотопластинки могут отразиться несколько волн разной длины, которым соответствуют различные порядки дифракционных максимумов. Поэтому метод Лауэ редко используется для исследования новых структур. Он находит применение главным образом для определения ориентации кристалла известной структуры. На рис. 1.42 показана схема закрепления исследуемого образца в трехосном гониометре. По известному распределению пятен (рис. 1.38), полученному после некоторой грубой предварительной установки кристалла, можно определить, на какой точно угол нужно повернуть кристалл, чтобы главная его плоскость идеально совпала с осью камеры Лауэ.

Само название *метода вращения кристалла* означает, что в нем угол падения θ изменяется во времени, в то время как

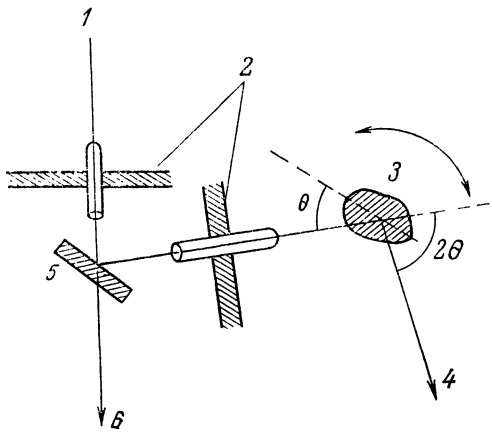


Рис. 1.43. Схема установки, в которой применяется метод вращения кристалла. В этом методе используется монохроматическое рентгеновское излучение, которое получается за счет брэгговских отражений на кристалле-монохроматоре. При вращении образца внутри камеры в ту или другую сторону на фотопленке, свернутой в цилиндр, ось которого совпадает с осью вращения кристалла, регистрируется серия дифракционных изображений. *1* — рентгеновский источник; *2* — коллиматоры; *3* — кристаллический образец; *4* — к фотопластинке или счетчику; *5* — кристалл-монохроматор; *6* — неотклоненный пучок (нежелательные длины волн).

рентгеновские лучи (или нейтроны) являются монохроматическими. На рис. 1.43 изображены существенные компоненты соответствующей экспериментальной установки. Из непрерывного спектра рентгеновских лучей нетрудно выделить монохроматическое излучение (обычно это линия K_α материала, из которого сделан катод рентгеновской трубки), причем в качестве монохроматора можно использовать любой подходящий монокристалл. При повороте исследуемого образца в ту или другую сторону всякий раз, когда выполняется условие Брэгга, на фотопленке, свернутой в цилиндр, коаксиальный с осью вращения кристалла, регистрируется распределение дифракционных пятен. Угол отклонения этих пятен относительно направления падающего пучка в плоскости, перпендикулярной оси цилиндра, равен 2θ, и, разумеется, большинство пятен также смещено в направлении оси цилиндра. К методам, основанным на принципе вращения кристалла и предназначенным для получения большей информации о его элементах симметрии, относятся также метод *качания* (метод *осциллирующего кристалла*), метод *Вейсенберга* и метод *прецессии*. Подробно все эти методы описаны в книге Вульфсона и Бюргера, указанной в литературе для данной главы.

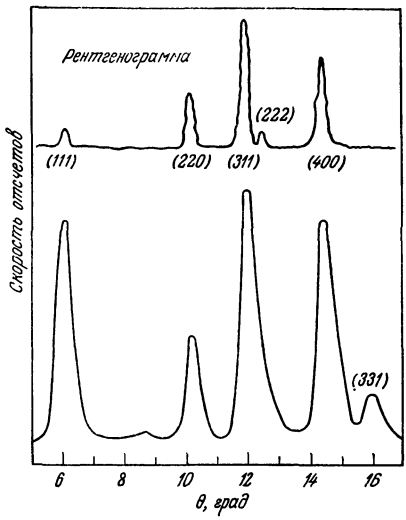


Рис. 1.44. Сравнение результатов рентгено- и нейтронографического исследований магнетита Fe_3O_4 [*Shull C. G. et al.*— Phys. Rev., 84, 912 (1951)].

Вместо того чтобы использовать монохроматическое рентгеновское излучение и изменение от времени угла θ, можно изготовить кристаллический образец таким образом, что в нем будут реализованы все углы падения на него рентгеновских лучей. Это осуществляется в методе *Дебая — Шеррера*, в котором в качестве кристаллического образца используется измельченный порошок со случайной ориентацией кристаллитов. Лучи, которые для того или иного кристаллита удовлетворяют условию Брэгга, выходят из образца по направлениям вдоль образующих концентрических конусов, ось которых совпадает с направлением падающего луча. При этом на фотопластинке регистрируется серия концентрических окружностей. Аналогично при дифракции электронов (см. рис. 1.39) сравниваются дифракционные картины от двух металлических пленок, в одной из которых ориентация кристаллитов является случайной, и на фотопластинке регистрируются кольца, а в другой имеется преимущественная ориентация и на фотопластинке регистрируются пятна. Фотометрирование картины интерференционных колец, полученной в порошковой рентгеновской камере, дает кривую с хорошо выраженными пиками при различных углах Брэгга.

В большинстве экспериментов по рентгеновскому исследованию методом вращения кристалла фотопластинку заменяют чувствительным счетчиком рентгеновских квантов, соединенным с записывающим устройством, которое вычерчивает кривую отсчетов, когда кристалл медленно поворачивается. (Аналогичные устройства для счета и записи используются и в лю-

бом эксперименте с дифракцией нейтронов.) Информацию, полученную по результатам дифракции рентгеновских лучей и нейтронов, можно затем сравнить для того, чтобы выявить линии, которые разрешаются, например, на рентгенограмме и не разрешаются на нейтронограмме. Мы уже отмечали, что для получения рефлексов от легких атомов, таких, как водород, лучше всего использовать нейтроны. Во многих соединениях структурный фактор при отражении от определенных плоскостей может быть равен нулю для рентгеновских лучей и отличен от нуля для нейтронов, и наоборот. Как показано на рис. 1.44, картина дифракции нейтронов в магнитном материале имеет дополнительные максимумы, которых нет на рентгенограмме; эти дополнительные максимумы обусловлены магнитным упорядочением электронов d-оболочек.

Дифракция в аморфных телах

До сих пор в данной главе речь шла в основном об идеальных периодических решетках. Разумеется, легче представить себе монокристалл, у которого на достаточно больших расстояниях имеются незначительные отклонения от периодичности, чем аморфное или микрокристаллическое твердое тело, однако мы не можем не рассматривать вещества, у которых вообще отсутствует дальний порядок.

В разд. 1.6 мы обсудим некоторые *точечные* и *линейные* дефекты, благодаря которым вместо идеальной периодичности совершенного кристалла мы имеем лишь подобие периодичности в реальном монокристалле. Здесь же мы сосредоточим внимание на веществах, которые считаются аморфными, но тем не менее дают хорошо воспроизводимую картину дифракции рентгеновских лучей. Интерес к аморфным твердым телам, в особенности к аморфным полупроводниковым материалам, сильно возрос начиная с середины 1960-х годов [30]. Увлечение аморфными халькогенидными стеклами возникло в основном из-за их необычных электронных свойств, поскольку была надежда на то, что благодаря этим свойствам стекла найдут широкое применение в переключающих устройствах [31].

Дифракция параллельного монохроматического пучка рентгеновских лучей на совокупности случайно распределенных изолированных атомов приводит к угловой дисперсии рассеяния,

[30] Amorphous Solids, Wiley-Interscience, eds. R W. Douglas, B. Ellis, 1972. См. также: *Mott N. F., Davis E. A.* Electronic Processes in Non-Crystalline Solids, Clarendon Press, Oxford, 1971. [Имеется перевод: *Мотт Н., Дэвис Э.* Электронные процессы в некристаллических веществах — М.: Мир, 1974.]

[31] *Henisch H. K.* Scientific American, **221**, No. 5, 30 (1969); Nature, **236**, 205 (1972).

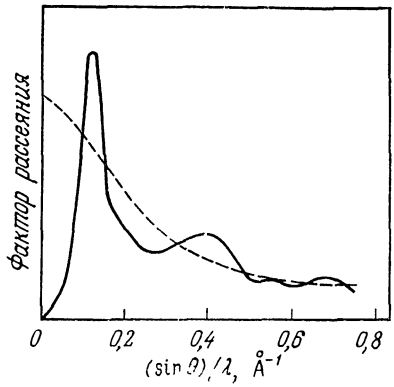


Рис. 1.45. Зависимость эффективности рассеяния рентгеновских лучей от угла дифракции в стеклообразном кремнеземе. Штриховая кривая соответствует среднему атомному фактору рассеяния для изолированных атомов кремния и кислорода. Эти данные, полученные Хеннингером и др. [*Henninger E. H. et al.*— J. Phys. Chem. Solids, 28, 423 (1967)], лишь незначительно отличаются от результатов Уоррена, полученных в 1938 г. Позднее эту кривую для стеклообразного кремния удалось продлить до значения $(\sin\theta)/\lambda = 1{,}6$ [*Mozzi R. L., Warren B. E.*— J. Appl. Cryst., 2, 164 (1969)].

которая определяется угловой зависимостью атомного фактора рассеяния, по форме напоминающей кривую на рис. 1.41. Дифракционная картина для жидкостей или аморфных твердых тел состоит из нескольких широких концентрических расплывчатых колец и внешне напоминает дифракционную картину для поликристаллического твердого тела. Первое время считалось, что существование таких колец свидетельствует о наличии кристаллической структуры. Однако это не верно.

Основываясь на изучении стеклообразного кремнезема и других аморфных твердых тел, Уоррен[32] в 1930-х годах указал на то, что кривая зависимости фактора рассеяния от отношения $(\sin\theta)/\lambda$ в твердых телах представляет собой сумму двух составляющих: монотонно убывающей (обусловленной наличием изолированных атомов) и осциллирующей компоненты, которая представляет собой в фурье-пространстве распределение вероятностей различных межатомных расстояний. Наличие в твердом теле даже в отсутствие дальнего порядка нескольких выделенных межатомных расстояний может привести к значительной осциллирующей компоненте при промежуточных значениях $(\sin\theta)/\lambda$. На рис. 1.45 представлена угловая зависимость фактора рассеяния кремнезема в сравнении с монотонно убывающей зависимостью фактора рассеяния от системы со случайным расположением атомов.

[32] *Warren B. E.*— J. Appl. Phys., 8, 645 (1938).

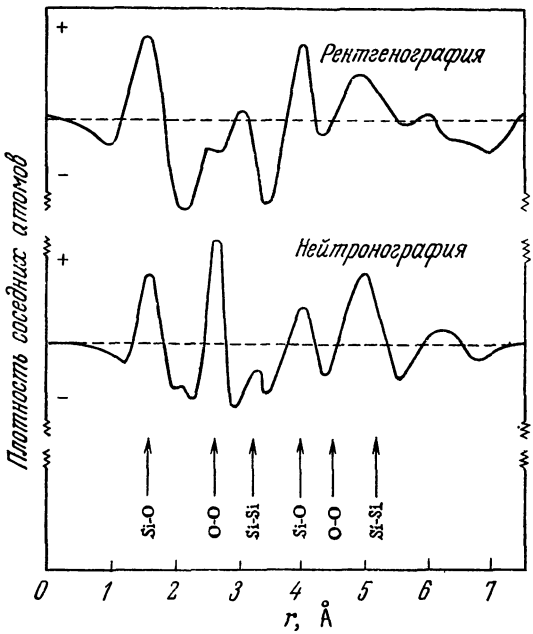


Рис. 1.46. Зависимость плотности соседних атомов от расстояния в стеклообразном кремнеземе (за вычетом общего увеличения плотности, пропорционального r^2). Кривые получены на основании результатов дифракции рентгеновских лучей и нейтронов, верхняя кривая построена по экспериментальным данным, представленным на рис. 1.45 [*Henninger E. et. al.*— J. Phys. Chem. Sol., 28, 423 (1967)].

Выполнив обратное преобразование Фурье кривой на рис. 1.45, можно получить зависимость плотности соседних атомов от расстояния до данного атома. На рис. 1.46 изображены такие кривые для кремнезема, построенные на основании как рентгенографических, так и нейтронографических данных. Здесь для сравнения указаны также межатомные расстояния в кристаллическом кварце. Из рисунка видно, что совпадают только расстояния между ближайшими соседями, в то время как расстояния между более удаленными атомами в кремнеземе никак не связаны с соответствующими расстояниями в кристаллическом кварце. Это подтверждает то, что стеклообразный кремнезем состоит из тетраэдрических образований SiO_4, случайно связанных между собой теми вершинами, в которых расположены атомы кислорода. Согласно Моззи и Уоррену[33], каждый атом кислорода связан с двумя атомами

[33] *Mozzi R. L., Warren B. E.*— J. Appl. Cryst., 2, 164 (1969).

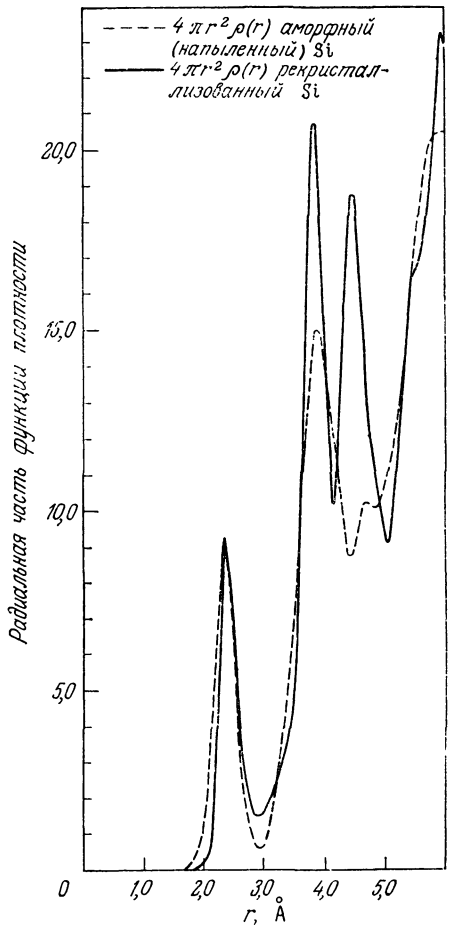


Рис. 1.47. Радиальное распределение плотности соседних атомов в кремнии, полученное на основании данных по дифракции электронов на тонких (напыленных) пленках аморфного кремния толщиной 100 Å и на тех же пленках, но после их рекристаллизации [*Moss S. C., Graczyk J. F.*— Proc. 10th Internat. Semiconductor Conference, A. E. C., 1970, p. 658.]

кремния, однако благодаря случайности связей тетраэдров SiO_4 угол между связями Si—O—Si может принимать любое значение в интервале 120°—180°.

Сравнение результатов исследования ряда других веществ, которые также принято считать аморфными, показывает, насколько далеко может распространяться ближний порядок в некристаллических твердых телах. Для смесей халькогенидных стекол, которые представляют интерес для создания электронных переключателей[31], этот вопрос еще совсем не изучен. К моменту написания данной книги было выявлено, что лучшие результаты получаются для аморфных структур различных элементов и простых бинарных соединений. Так, исследования по дифракции в аморфном селене указывают на то, что спиральные цепочки в структуре селена (рис. 1.6) состоят из групп, у которых всего три или четыре атома, причем направления связей между этими малыми группами случайны. Исследования проводились на аморфных структурах таких хорошо известных полупроводниковых элементов, как кремний и германий, у которых хорошо изученная кристаллическая структура имеет решетку алмаза, показанную на рис. 1.31, *а*. Зависимость плотности соседних атомов в аморфном и рекристаллизованном кремнии от расстояния представлена на рис. 1.47. Эта зависимость получена путем

анализа данных по дифракции *электронов*. (Кривая, приведенная на рис. 1.47, отличается от соответствующих кривых на рис. 1.46 тем, что при построении последних было вычтено общее возрастание плотности, пропорциональное $4\pi r^2$). Исследуя положения максимумов на рис. 1.47, можно показать, что как в аморфном, так и в кристаллическом кремнии каждый атом Si имеет четырех ближайших соседей, находящихся от него на расстоянии 2,35 Å, однако расположение более удаленных атомов в аморфном состоянии оказывается случайным. Поэтому на кривой для аморфного кремния почти не заметен максимум, который соответствовал бы третьим по удаленности соседям, расположенным на расстоянии 4,5 Å. В то же время влияние искажения угла между связями достаточно заметно и проявляется в увеличении ширины максимума для вторых по дальности соседей.

1.5. Обратное пространство

Обратная решетка

На данном этапе рассмотрения периодических структур целесообразно, наконец, ввести понятие обратного пространства и обратной решетки. Точно так же как любую величину, изменяющуюся во времени, можно представить в виде суммы фурье-компонент в частотной области, так и любую пространственную характеристику кристалла можно записать в виде суммы компонент в фурье-пространстве, или обратном пространстве. Для идеального монокристалла (представляющего собой набор повторяющихся в трехмерном пространстве идентичных блоков) обратная решетка в фурье-пространстве образуется бесконечным множеством точек, периодически расположенных в трехмерном пространстве, расстояния между которыми обратно пропорциональны расстояниям между *плоскостями* в прямой решетке. Поэтому очевидно, что расстояния в обратной решетке можно связать с условием дифракции.

Поскольку в реальном пространстве векторы имеют размерность длины, в обратном пространстве векторы имеют размерность (длина)$^{-1}$. Векторы обратного пространства можно сопоставить с волновыми векторами[34] возбуждений (таких, как

[34] Мгновенное значение амплитуды волны, распространяющейся со скоростью v и имеющей частоту ν, можно записать в виде $A=A_0\exp[2\pi i\nu(x/v-t)]$ или через длину волны $\lambda=v/\nu$ в виде $A_0\exp[2\pi i(x/\lambda-\nu t)]$. Нередко более удобно описывать волну с помощью угловой частоты $\omega=2\pi\nu$ и волнового вектора $k=2\pi/\lambda=\omega/v$. При этом амплитуда волны записывается как $A=A_0\exp[i(kx-\omega t)]$. В трехмерном пространстве, где волновой вектор действи-

фотоны, колебания решетки или движущиеся свободные электроны). Умножение каждой координаты на $\hbar$ преобразует обратное пространство в пространство импульсов. Понятие обратной решетки мы будем часто использовать в последующих главах (гл. 2—4). В этой же главе мы рассмотрим прежде всего вопрос о том, какие новые преимущества дает понятие обратной решетки с точки зрения условия дифракции Брэгга.

Пусть $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ — примитивные векторы трансляции трехмерной решетки в реальном пространстве. Тогда *основные векторы* обратной решетки можно *определить* следующим образом:

$$\mathbf{a}^* \equiv \frac{2\pi\, \mathbf{b}\times\mathbf{c}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}},$$
$$\mathbf{b}^* \equiv \frac{2\pi\mathbf{c}\times\mathbf{a}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}}, \qquad (1.38)$$
$$\mathbf{c}^* \equiv \frac{2\pi\mathbf{a}\times\mathbf{b}}{\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}}.$$

(Впоследствии мы покажем, что множитель 2π в этих выражениях можно опустить.) В знаменателях выражений (1.38) стоит одна и та же величина, а именно объем элементарной ячейки в реальной (прямой) решетке, который не зависит от перестановки векторов $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$.

Из (1.38) следует, что основной вектор обратной решетки (например, вектор $\mathbf{a}^*$) может и не быть параллелен соответствующему вектору реальной решетки (а именно, вектору $\mathbf{a}$), но *обязательно* должен быть перпендикулярен как $\mathbf{b}$, так и $\mathbf{c}$. Можно записать следующие соотношения для векторов прямой и обратной решеток:

$$\mathbf{a}^*\cdot\mathbf{b} = \mathbf{a}^*\cdot\mathbf{c} = \mathbf{b}^*\cdot\mathbf{a} = \mathbf{b}^*\cdot\mathbf{c} = \mathbf{c}^*\cdot\mathbf{a} = \mathbf{c}^*\cdot\mathbf{b} = 0, \qquad (1.39)$$
$$\mathbf{a}^*\cdot\mathbf{a} = \mathbf{b}^*\cdot\mathbf{b} = \mathbf{c}^*\cdot\mathbf{c} = 2\pi. \qquad (1.40)$$

Эти соотношения иллюстрируются на рис. 1.48 на примере двумерной косоугольной решетки.

Обратное преобразование к прямой решетке можно записать аналогично выражениям (1.38). Нетрудно проверить, что

$$\mathbf{a} = \frac{2\pi\, \mathbf{b}^*\times\mathbf{c}^*}{\mathbf{a}^*\cdot\mathbf{b}^*\times\mathbf{c}^*} \qquad (1.41)$$

и т. д.

тельно является вектором, амплитуду мы должны записать как $A = A_0\exp[i(\mathbf{k}\cdot\mathbf{r}-\omega t)]$. Согласно представлениям квантовой механики, импульс возбуждения с волновым вектором $\mathbf{k}$ равен $\mathbf{p} = \hbar\mathbf{k}$.

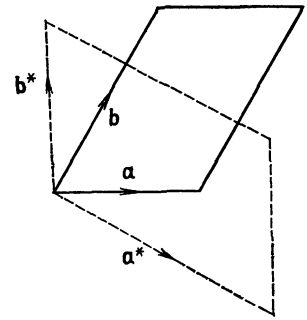


Рис. 1.48. Основные векторы обратной решетки для двумерной косоугольной «прямой» решетки. Заметим, что вектор **a*** перпендикулярен вектору **b**, а вектор **b*** перпендикулярен вектору **a**; кроме того, **a***·**a** равно **b***·**b**.

Множитель 2π введен в выражения (1.38), (1.40) и (1.41) для того, чтобы обратное пространство совпадало с пространством волновых векторов не только в смысле *размерности*, но также и *численно*. В ряде случаев при рассмотрении обратного пространства множитель 2π опускают, и тогда единичная трансляция в обратном пространстве оказывается эквивалентной изменению обратной длины волны на единицу. Опасность запутаться в этих двух системах обратных векторов не столь уж велика, как может сначала показаться, поскольку из контекста обычно ясно, о какой системе идет речь. Действительно, в одном случае некоторую точку в обратном пространстве обозначают как $(1/2\ 0\ 0)$, а в другом эта же точка будет обозначаться как $(\pi\ 0\ 0)$.

Векторы обратной решетки

В самом начале рассмотрения операций симметрии мы выяснили, что особая роль в реальной решетке принадлежит векторам трансляции, которые образуют следующее семейство векторов:

$$\mathbf{T} = n_1\mathbf{a} + n_2\mathbf{b} + n_3\mathbf{c}. \qquad (1.42)$$

Эти векторы трансляции связывают в кристаллической решетке пары точек, которые имеют одинаковые атомные окружения. В обратном пространстве также вводится понятие трансляций, которые описываются *векторами обратной решетки*, образующих следующее семейство:

$$\mathbf{G}_{hkl} = h\,\mathbf{a}^* + k\,\mathbf{b}^* + l\,\mathbf{c}^*, \qquad (1.43)$$

где h, k и l — целые числа.

Покажем, что вектор обратной решетки, задаваемой тройкой целых чисел h, k и l, перпендикулярен (в обратном пространстве) плоскости (hkl) в реальном пространстве. Для этого напомним, что плоскость (hkl) проходит через точки $(\mathbf{a}/h)$,

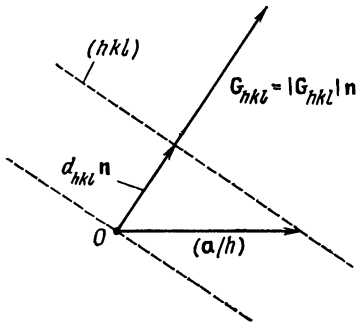


Рис. 1.49. Схема для определения расстояния между соседними плоскостями (hkl).

(b/k) и (c/l). При этом нетрудно показать (см. задачу 1.17), что любой вектор, лежащий в плоскости (hkl), записывается в виде

$$R = (A/h)\,\mathbf{a} + (B/k)\,\mathbf{b} + (C/l)\,\mathbf{c}, \quad (1.44)$$

где коэффициенты A, B и C связаны условием

$$A + B + C = 0. \quad (1.45)$$

При этом из выражений (1.39), (1.40) и (1.45) ясно, что

$$\mathbf{R} \cdot \mathbf{G}_{hkl} = 0. \quad (1.46)$$

Поскольку $\mathbf{R}$ — произвольный вектор, лежащий в плоскости (hkl) реального пространства, очевидно, что вектор $\mathbf{G}_{hkl}$ является перпендикулярным этой плоскости.

Мы уже знаем, что для кубической решетки можно получить простое выражение для расстояния между соседними плоскостями (hkl). Использование вектора обратной решетки $\mathbf{G}_{hkl}$ позволяет получить компактное выражение для любой решетки. Из рис. 1.39 видно, что как вектор $\mathbf{G}_{hkl}$, так и вектор, соединяющий начало координат с плоскостью (hkl) в реальном пространстве, можно записать как кратное единичного вектора $\mathbf{n}$. Уравнение плоскости (hkl), записанное в виде

$$d_{hkl} = \mathbf{r} \cdot \mathbf{n} = \frac{\mathbf{r} \cdot \mathbf{G}_{hkl}}{|\mathbf{G}_{hkl}|}, \quad (1.47)$$

будет выполняться для любого вектора $\mathbf{r}$, длина которого больше, чем d_{hkl}. Этому требованию, очевидно, удовлетворяет вектор $\mathbf{a}/h$. Таким образом,

$$d_{hkl} = \frac{\mathbf{a} \cdot \mathbf{G}_{hkl}}{h\,|\mathbf{G}_{hkl}|}. \quad (1.48)$$

Но из (1.40) и (1.43) следует, что $\mathbf{a} \cdot \mathbf{G}_{hkl} = 2\pi h$. Отсюда мы получаем, что расстояние между плоскостями (hkl) в реальном пространстве связано с длиной вектора обратной решетки $\mathbf{G}_{hkl}$ простым выражением

$$d_{hkl} = \frac{2\pi}{|\mathbf{G}_{hkl}|}. \quad (1.49)$$

Условие дифракции

Использование понятия обратной решетки позволяет по-новому сформулировать условие Брэгга для зеркального отражения. Вместо того чтобы рассматривать *длину волны* излуче-

ния, которое взаимодействует с последовательностью атомных плоскостей (см. рис. 1.40), мы будем теперь оперировать с начальным и конечным *волновыми векторами* ($\mathbf{k}$ и $\mathbf{k}'$) отраженной волны. При упругом рассеянии длина волнового вектора не меняется:

$$|\mathbf{k}| = |\mathbf{k}'| = (2\pi/\lambda). \quad (1.50)$$

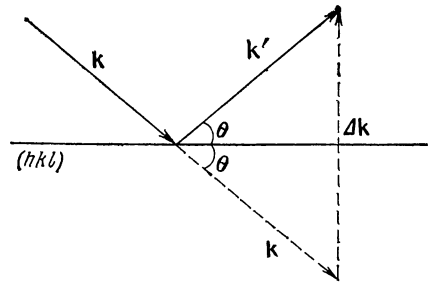


Рис. 1.50. Изменение волнового вектора при зеркальном отражении от плоскости (hkl).

Из векторного треугольника на рис. 1.50 видно, что вектор $\Delta\mathbf{k}$ перпендикулярен плоскостям $(h\,k\,l)$, т. е. имеет то же направление, что и вектор $\mathbf{G}_{hkl}$, или единичный вектор $\mathbf{n}$. Таким образом,

$$\Delta\mathbf{k} = (\mathbf{k}' - \mathbf{k}) = 2\sin\theta\,|\mathbf{k}|\,\mathbf{n} = \left(\frac{4\pi\sin\theta}{\lambda}\right)\mathbf{n} =$$
$$= \left(\frac{4\pi\sin\theta}{\lambda\,|\mathbf{G}_{hkl}|}\right)\mathbf{G}_{hkl} = \left(\frac{2d_{hkl}\sin\theta}{\lambda}\right)\mathbf{G}_{hkl}. \quad (1.51)$$

Если λ, θ и d_{hkl} таковы, что выполняется условие Брэгга, то

$$\Delta\mathbf{k} = \mathbf{G}_{hkl}. \quad (1.52)$$

Таким образом, множество точек, образованных вектором $\mathbf{G}_{hkl}$, соответствует распределению пятен Лауэ, полученных вследствие дифракции на кристалле. Расстояния между этими точками обратно пропорциональны расстояниям между плоскостями в реальной решетке.

С учетом выражения (1.52) соотношение между начальным и конечным волновыми векторами волны, испытавшей брэгговское отражение, можно записать в виде

$$\mathbf{k}' = \mathbf{G}_{hkl} + \mathbf{k}. \quad (1.53)$$

Возводя обе части этого соотношения в квадрат и вычитая из обеих частей величину $|\mathbf{k}|^2 = |\mathbf{k}'|^2$, мы получаем следующую запись условия Брэгга:

$$\mathbf{G}^2 + 2\mathbf{k}\cdot\mathbf{G} = 0. \quad (1.54)$$

Записанное таким образом условие дифракции является весьма полезным при рассмотрении зонной теории электронов в кристаллах, которую мы обсудим в гл. 3 настоящей книги.

Поскольку волновой вектор при отражении от последовательности плоскостей (hkl) изменяется на величину $\Delta\mathbf{k} = \mathbf{G}_{hkl}$, из соотношений (1.40) и (1.43) мы получаем следующие выражения:

$$\mathbf{a}\cdot\Delta\mathbf{k} = 2\pi h, \quad \mathbf{b}\cdot\Delta\mathbf{k} = 2\pi k, \quad \mathbf{c}\cdot\Delta\mathbf{k} = 2\pi l, \quad (1.55)$$

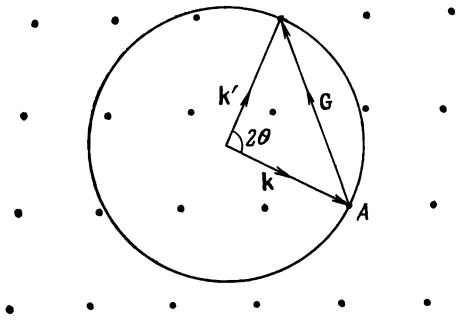


Рис. 1.51. Построение Эвальда, объясняющее существование дифрагированного излучения. Пусть волновому вектору падающей волны $\mathbf{k}$ соответствует вектор, конец которого находится в точке A обратной решетки. Другой конец этого вектора совпадает с центром сферы радиусом $|\mathbf{k}|$. При этом условие Брэгга $\Delta\mathbf{k}=\mathbf{k}'-\mathbf{k}=\mathbf{G}$ удовлетворяется всякий раз, когда сфера Эвальда проходит через какой-нибудь узел обратной решетки кристалла.

которые связывают между собой $\Delta\mathbf{k}$, тройку чисел hkl и векторы реальной (прямой) решетки. Выражения (1.55) называются *уравнениями Лауэ* и весьма полезны для описания симметрии и структуры кристаллов.

Исходя из формулировки условия Брэгга в виде (1.52), Эвальд в 1921 г. предложил интересную геометрическую интерпретацию условия дифракции (рис. 1.51). Дифракция будет обеспечиваться тем семейством плоскостей, которое (в реальном пространстве) перпендикулярно любому из векторов $\mathbf{G}$, соединяющих точку A с некоторой другой точкой на поверхности сферы (в обратном пространстве).

Зоны Бриллюэна

Если прямая решетка строго периодична, то обратная решетка, т. е. множество точек, удовлетворяющих условию (1.43), также периодична и бесконечна. Однако в тех задачах физики твердого тела, где удобно использовать представление обратной решетки, достаточно бывает ограничиться конечным объемом обратного пространства, или $\mathbf{k}$-пространства. Так обстоит дело в задачах, связанных с определением дисперсионных законов для электронов или элементарных возбуждений в кристалле.

Первую зону Бриллюэна определяют как область в обратном пространстве, окружающую один из узлов обратной решетки и ограниченную набором плоскостей, проходящих через середины векторов, соединяющих в обратной решетке данную точку с ее ближайшими соседями. На рис. 1.52 показана зона Бриллюэна для двумерной косоугольной решетки. Аналогично определяется зона Бриллюэна и в трехмерных решетках (задача 1.19). Любой точке $A'(\mathbf{k}')$ в обратном пространстве соответствует точка $A(\mathbf{k})$ в первой зоне Бриллюэна, так что выполняется соотношение

$$\mathbf{k}'=\mathbf{k}+\mathbf{G}. \qquad (1.56)$$

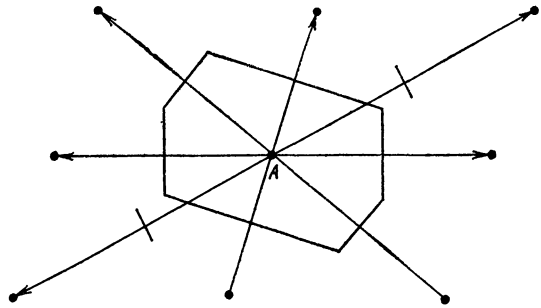


Рис. 1.52. Часть обратного пространства для двумерной косоугольной решетки. Показаны прямые линии, проходящие через середины векторов обратной решетки, проведенных из точки *A*. Шесть этих векторов наименьшей длины используются для определения первой зоны Бриллюэна относительно точки *A*.

В последующих главах мы убедимся в том, что, в твердых телах волну, характеризуемую волновым вектором $\mathbf{k}'$, невозможно отличить от волны, описываемой приведенным волновым вектором $\mathbf{k}$. Зонная теория твердых тел, которую мы подробно рассмотрим в гл. 3 и 4, является прямым следствием свойства периодической решетки создавать брэгговское отражение электронов, обладающих критической скоростью (электронов, волновые векторы которых совпадают с границами зоны Бриллюэна).

В ряде мест этой книги мы будем пользоваться понятием зоны Бриллюэна как для гипотетических одномерных (1-*D*) и двумерных (2-*D*) твердых тел, так и для «реальных» 3-*D*-решеток. Зоны Бриллюэна для 1-*D* и 3-*D* г. ц. к.-решеток применяются в разд. 2.2 и 2.3 при рассмотрении колебательного спектра соответствующих кристаллов. В разд. 2.4 зона Бриллюэна 3-*D*-решетки, имеющая форму многогранника, аппроксимируется сферой с «дебаевским радиусом», т. е. сферой, имеющей тот же объем. В разд. 2.5 размерность зоны Бриллюэна (предположительно 2-*D*) учитывается при определении температуры, ниже которой процессы переброса при рассеянии на фононах становятся маловероятными.

В гл. 3 мы снова обратимся к исследованию зон Бриллюэна, но уже в связи с изучением вопроса о заполнении возможных электронных состояний. В частности, на рис. 3.23—3.27, 3.30—3.33, 3.34—3.36 изображены зоны Бриллюэна соответственно для 1-*D*-, 2-*D*- и 3-*D*-кристаллов. Фактически с зонами Бриллюэна мы будем иметь дело на всем протяжении гл. 3 и 4. Для

многих случаев важно знать структуру зоны Бриллюэна для кристалла с г.ц.к.-симметрией (такова зона Бриллюэна в самой г.ц.к.-решетке, в структуре типа NaCl, в алмазной решетке и решетке цинковой обманки). Эта зона Бриллюэна используется на рис. 3.46, 3.47, 3.50, 3.51, а также на рис. 4.45 и 4.50.

1.6. Дефекты кристаллической решетки

В предшествующих разделах много внимания уделялось вопросам симметрии идеальной периодической решетки и бесконечному множеству точек в обратном пространстве, которое описывает эту решетку в фурье-представлении. Однако реальные кристаллы *несовершенны*, хотя бы потому, что они имеют конечные размеры. Макроскопические образцы большинства кристаллов состоят из множества случайно ориентированных кристаллитов, отделенных друг от друга межкристаллитными границами. На этих границах собирается множество различных микроскопических дефектов. Кроме того, каждый кристаллит обладает конечной концентрацией *точечных дефектов*, а иногда и конечной плотностью *линейных дефектов* или *дислокаций*.

Точечные дефекты

Точечные дефекты (несовершенства на атомном уровне) в твердом теле можно разделить на две основные категории: примеси (чужеродные атомы) и собственные точечные дефекты. Точечными дефектами также называют и более *сложные* образования, состоящие либо из очень близко расположенных примесных атомов нескольких сортов, либо из примесных атомов, связанных с собственными точечными дефектами. На первый взгляд это противоречит самому понятию «точечного» дефекта. Однако выбор терминологии основан на том, что максимальный размер любого точечного дефекта в каждом направлении не должен превышать нескольких межатомных расстояний. Напротив, *линейные* или *двумерные* дефекты могут простираться на миллионы межатомных расстояний в одном или двух направлениях.

Присутствие примесных атомов сильно влияет на электрические, оптические, магнитные, тепловые и механические свойства твердых тел. В частности, в гл. 4 мы покажем, сколь большое влияние оказывают примеси на свойства неметаллических проводников. Следует заметить, что расчет, основанный на за-

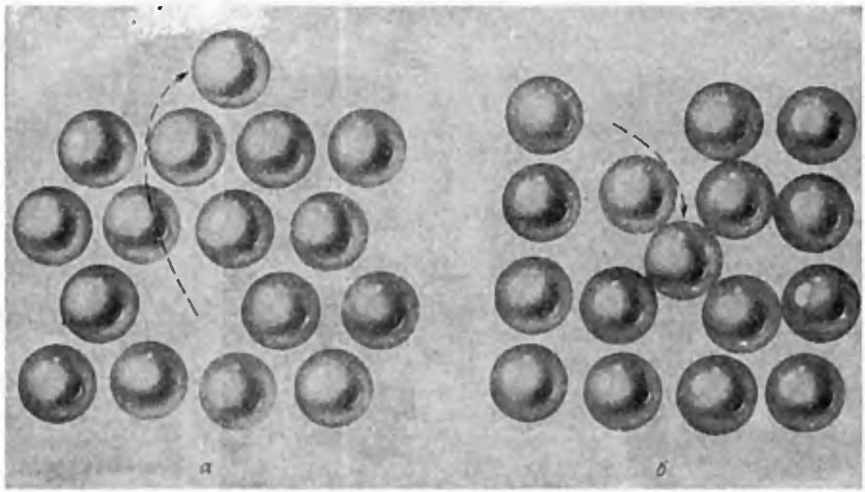


Рис. 1.53. Образование в кристаллической решетке вакансии (*а*) и междоузельного атома (*б*).

конах электростатики с привлечением термодинамического закона действующих масс, показывает, что равновесная концентрация *собственных* дефектов может зависеть от концентрации примесей. Исчерпывающий обзор взаимодействий собственных и примесных дефектов содержится в книге Крёгера (см. литературу в конце настоящей главы).

Простейшим случаем собственного дефекта является *вакансия*, т. е. отсутствие атома в узле кристаллической решетки и наличие дополнительного атома на поверхности кристалла (рис. 1.53, *а*), причем общее число атомов в кристалле не изменяется. Если предположить, что атомы в кристалле жестко закреплены, то для удаления атома из узла кристаллической решетки необходимо затратить энергию E_v; для помещения же этого атома на поверхности кристалла выделяется энергия E_s. Таким образом, энергия образования вакансии в модели жесткой решетки равна E_v—E_s, которая, как правило, составляет примерно 10 эВ на вакансию (см. задачу 1.20). Поскольку реальная решетка не является идеально жесткой, некоторая доля энергии E_v—E_s компенсируется за счет релаксации решетки: все соседние с пустым узлом атомы немного сдвигаются. При образовании вакансии в ионном соединении в объеме, окружающем вакантный узел, должно выполняться условие электронейтральности.

Дефекты в кристаллах галогенидов щелочных металлов, называемые «центрами окраски» (простейшими из них являются F-центры), обусловлены наличием отдельных вакансий или близко расположенных групп вакансий. Специфические электронные и оптические свойства этих дефектов определяются условиями электростатического равновесия. В книге Мотта и Герни (см. примечание 12) дан обзор работ, относящихся к начальному этапу исследований таких дефектов. Изучение этих свойств дефектов продолжается и в настоящее время.

В ионных твердых телах электростатическая нейтральность всего кристалла сохраняется автоматически, если числа вакансий для атомов различных сортов находятся в определенной пропорции, а именно в такой, что сохраняется *химическая стехиометрия*. В этом случае говорят о *разупорядочении типа Шоттки*. При таком разупорядочении в RbF число пустых мест, соответствующих ионам Rb^+, оказывается тем же, что и для ионов F^-, в то время как в Cu_2O для выполнения стехиометрического баланса необходимо, чтобы число вакансий для ионов меди в два раза превышало число вакансий для атомов кислорода. Обычно разупорядочение типа Шоттки реализуется только в бинарных соединениях, причем с соотношением 1 : 1 между компонентами, и под дефектом Шоттки мы понимаем в этом случае наличие пары вакансий противоположного знака. Два компонента дефекта Шоттки не обязательно находятся в непосредственной близости друг от друга.

Междоузельный атом (рис. 1.53, *б*) — это атом (или ион), который с поверхности попадает внутрь кристалла и размещается в некотором промежуточном положении в решетке. При этом вблизи междоузельного атома возникают сжатия и растяжения кристаллической решетки. Внедрение атомов в плотно упакованные структуры, такие, как медь (г. ц. к.) или цинк (г. п. у.), практически невозможно, зато в менее плотных решетках, например типа алмаза, существует много позиций, где мог бы разместиться атом внедрения (собственный или примесный), вызвав при этом лишь несущественное смещение ближайших соседей. Одновременное внедрение положительных и отрицательных ионов в стехиометрических пропорциях маловероятно, и поэтому для такого дефекта нет даже специального термина. Например, в решетке NaCl ионы Na^+ могут найти свободное место в междоузлиях решетки, тогда как ионы Cl^- для этого слишком велики.

Дефект типа Френкеля состоит из вакансии и междоузельного атома того же сорта (при этом обеспечивается электростатическая нейтральность и сохранение стехиометрии). Разупорядочение типа Френкеля может произойти либо при установлении теплового равновесия, либо при бомбардировке кри-

сталла ядерными частицами. Значительная часть исследований радиационных повреждений в твердых телах[35] связана главным образом с изучением возникновения и отжига дефектов типа Френкеля.

Для разупорядочения, известного под названием *антиструктурного*, характерно такое расположение атомов в соединении, при котором атомы одного сорта занимают места, предназначенные для атомов другого сорта. Антиструктурное разупорядочение в кристалле InSb не нарушает стехиометрии, если число атомов индия, занимающих позиции атомов сурьмы, равно числу атомов сурьмы, расположенных в позициях атомов индия. Антиструктурные дефекты возникают только тогда, когда размеры атомов двух сортов достаточно близки. В этом случае возникают причудливые образования — комплексы с неправильным расположением атомов. Эти комплексы занимают сравнительно малую область в кристалле и содержат от пятидесяти до ста атомов.

Плотность дефектов в состоянии теплового равновесия

Для образования точечного дефекта, такого, как вакансия, требуется затратить некоторое количество энергии (скажем, E эВ). Поэтому может показаться, что существование конечной концентрации вакансий в кристалле при тепловом равновесии энергетически невыгодно. Однако при наличии конечной концентрации вакансий энтропия системы увеличивается и *свободная энергия* может стать меньше, чем у идеального кристалла. Свободная энергия равна разности полной энергии кристалла и произведения энтропии на температуру. Предположим, что кристалл состоит из N атомов и n вакансий, случайно распределенных по узлам решетки. Тогда приращение энтропии записывается в виде

$$\Delta S = k_0 \ln\left[\frac{N!}{(N-n)!\ n!}\right]. \qquad (1.57)$$

[35] Радиационные повреждения в твердых телах широко изучаются начиная с 1940-х годов. Часть исследований проводится в военных целях, однако интерес к изучению радиационных повреждений связан также с применением радиации в мирных целях. Например, радиационные повреждения и деградация происходят в любых материалах, которые используются в реакторах, медицинском оборудовании, предназначенном для работы с применением облучения, и при работе в космосе. Обзор работ, относящихся к начальному периоду изучения этого вопроса, дан Д. Биллингтоном и Дж. Крауфордом в книге: *Billington D. S., Crawford J. H.* Radiation Damage in Solids, Princeton, 1961. Влиянию радиационных повреждений на свойства полупроводников посвящены работы Ф. Л. Вука [*Vook F. L.* Radiation Effects in Semiconductors, Plenum, 1969; Radiation Damage and Defects in Semiconductors, ed. J. Whitehouse, Institute of Physics, 1973].

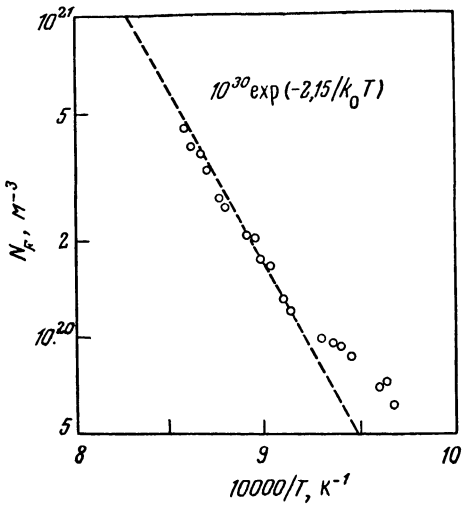


Рис. 1.54. Температурная зависимость плотности дефектов Френкеля в германии. Для графического представления процессов с энергией активации использован, как общепринято, полулогарифмический масштаб, причем по оси абсцисс отложены значения $1/T$. [*Mayburg S., Rotondi L.*— Phys. Rev., **91**, 1015 (1953)].

Используя приближенную формулу Стирлинга для факториала, получаем

$$\Delta S = k_0 [N \ln(N) - (N-n)\ln(N-n) - n\ln(n)]. \qquad (1.58)$$

При возникновении вакансий свободная энергия изменится на величину

$$\Delta F = nE - T\Delta S = nE - k_0 T [N \ln(N) - (N- \\ -n)\ln(N-n) - n\ln(n)]. \qquad (1.59)$$

Термодинамически наиболее вероятным значением n является такое, для которого производная $\partial\Delta F/\partial n$ равна нулю. Из уравнения (1.59) следует, что это условие выполняется, когда

$$n/(N-n) = \exp(-E/k_0T). \qquad (1.60)$$

Таким образом, для любых реальных значений плотности собственных дефектов их концентрация является экспоненциальной функцией обратной температуры.

Аналогичные выражения можно получить и для равновесных концентраций других точечных дефектов. В общем случае минимизация функционала свободной энергии $\Delta F = \Delta U - T\Delta S$ должна проводиться для *сочетаний* дефектов различных типов с разными энергиями активации.

На рис. 1.54 приведена кривая зависимости равновесной плотности дефектов типа Френкеля от температуры в германии, полученная Маубургом и Ротонди (1953). Попытки измерить плотность собственных дефектов в германии, а также в большинстве других кристаллов в течение ряда лет были не-

удачными из-за наличия в этих кристаллах различного вида примесей, в частности примесей меди. Для любого кристалла намного легче оказывается получить кривую плотности дефектов, которая экспоненциально зависит от $1/T$, чем удостовериться в том, что измеряемая величина действительно обусловлена именно исследуемыми дефектами.

Стехиометрия

Обычно считают, что соединения типа AB имеют одинаковые количества атомов A и B. Такой кристалл называется стехиометрическим. По-видимому, для многих кристаллов, в которых соотношение анионов и катионов отличается от номинального, нарушение стехиометрии не имеет большого значения. В нестехиометрических соединениях структура получается равновесной за счет существования либо вакансий, либо междоузельных атомов, либо тех и других одновременно.

В соединениях типа NiO или PbS отклонение от стехиометрии можно легко создать путем выдержки при высокой температуре либо в вакууме, либо при высоком давлении паров наиболее летучего компонента. При этом, как следует из закона действующих масс, изменение плотности дефектов пропорционально некоторой степенной функции парциального давления летучего компонента. (Более подробно см. об этом в книге Крёгера.) Если кристалл, который после такой процедуры стал нестехиометрическим, быстро охладить, то в нем можно зафиксировать высокую плотность дефектов. Эта плотность значительно выше равновесной плотности, соответствующей минимуму свободной энергии решетки при низких температурах, но колебания решетки при этих температурах слишком слабы, чтобы могла возникнуть обратная диффузия избыточных дефектов за конечный промежуток времени. Такие избыточные дефекты, полученные при быстром охлаждении, могут скапливаться на границах зерен или на дислокациях.

Дислокация

Точечные дефекты, например вакансии, вызывают нарушения решетки лишь в окрестности одного узла решетки (или в крайнем случае нескольких соседних узлов). Дислокации же приводят к линейным нарушениям структуры. При этом *не следует* отождествлять дислокации с линиями вакансий. В общем случае дислокации в кристалле могут иметь любую форму, однако понять их строение можно на примере двух простейших типов дислокации: *краевой* и *винтовой*.

В случае чисто *краевой дислокации* одна из атомных плоскостей внутри кристалла обрывается, напоминая лезвие ножа

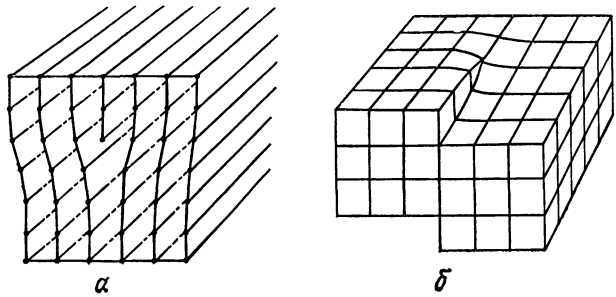


Рис. 1.55. Два вида простых дислокаций. *а* — краевая дислокация; *б* — винтовая дислокация. В общем случае линия дислокации искривлена; дислокация имеет смешанный характер с изменяющимся вдоль дислокации соотношением краевой и винтовой компонент.

при нарезании сыра. Этот тип дислокаций иллюстрирует рис. 1.55, *а*. Сдвиг в решетке (определяемый вектором Бюргерса **b**) перпендикулярен линии дислокации (направлению последнего ряда атомов в полуплоскости). Длина вектора Бюргерса равна целому числу единичных векторов решетки.

В случае *винтовой дислокации* (рис. 1.55, *б*) одна часть решетки смещена относительно другой в направлении, *параллельном* линии дислокации. Чтобы мысленно получить винтовую дислокацию, надо сначала сделать в совершенном кристалле надрез, а затем одну сторону надреза поднять относительно другой на величину одного или нескольких единичных векторов решетки и в таком положении состыковать обе стороны разреза. Для описания такого сдвига также пользуются вектором Бюргерса, который всегда равен целому числу единичных векторов решетки, а для чисто винтовой дислокации к тому же параллелен ее оси.

В общем случае дислокация в реальном кристалле может быть представлена как результат наложения краевой и винтовой компонент, причем соотношение между ними меняется вдоль линии дислокации. Заметим, однако, что вектор Бюргерса на всем протяжении дислокации не меняется. Дислокация не может обрываться внутри кристалла; она либо имеет вид замкнутой петли, либо оканчивается на свободных поверхностях кристалла или межкристаллитных границах. Для большинства материалов выходы дислокации на поверхность кристалла — единственные места, где можно «видеть» дислокацию.

Если кристалл пронизывает винтовая дислокация, то говорить об атомных плоскостях, перпендикулярных дислокации уже нельзя, поскольку все атомы образуют спиральную поверхность, которая простирается от одной внешней границы кри-

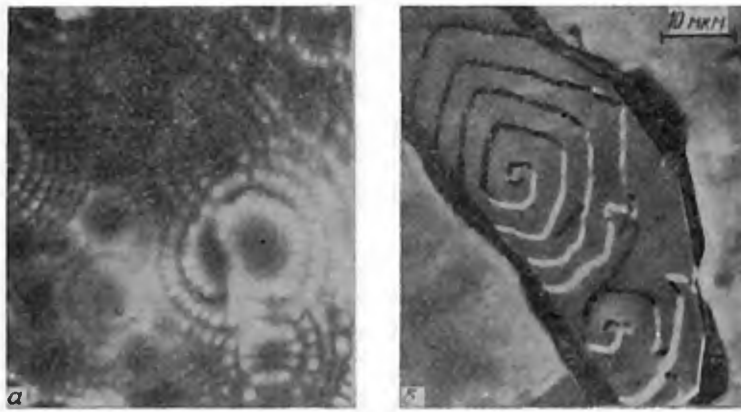


Рис. 1.56. Ступенчатые спиральные структуры на поверхности металла. *а* — атомная структура спирали на поверхности кристалла иридия, наблюдаемая с помощью ионно-эмиссионного микроскопа (с любезного разрешения О. Т. Инала); *б* — многовитковые спирали на поверхности электролитически осажденного серебра, наблюдаемые в просвечивающем электронном микроскопе методом реплик (с любезного разрешения Т. Сузуки) [*Murr L. E.* Interfacial Phenomena in Metals and Alloys, Addison-Wesley, 1975].

Рис. 1.57. Выявление дислокаций методом травления (100) поверхности монокристалла LiF. Кроме ямок травления, случайно расположенных на поверхности кристалла, в нижнем левом углу рисунка по диагонали проходит малоугловая граница, отмеченная рядом ямок травления, расположенных вплотную друг к другу. (С любезного разрешения Л. Е. Марра из Орегонского учебного центра.)

сталла до другой. При этом шаг спирали равен вектору Бюргерса. Эти спирали могут быть закручены либо вправо, либо влево. При медленном выращивании кристалла из газовой фазы рост за счет нарастания слоев на поверхностной ступеньке винтовой дислокации оказывается энергетически предпочтительным, т. е. кристаллизация может происходить путем непрерывного наращивания витков спирали. Для иллюстрации на рис. 1.56 приведены два примера; фотография слева получена с помощью ионно-эмиссионного микроскопа с разрешением отдельных атомов.

Во многих случаях следы выхода дислокации на поверхность кристалла можно обнаружить с помощью химического травления, которое сильнее всего действует на напряженные области, при этом появляются *ямки травления*, ограниченные кристаллографическими плоскостями. Несколько таких изолированных ямок травления видны на рис. 1.57. На этом рисунке показан также ряд ямок травления, расположенных вплотную друг к другу вдоль *малоугловой границы* между двумя частями почти идеального монокристалла. Приведенные фотографии подтверждают гипотезу Бюргерса (1939) о том, что границы с малыми углами разориентировки представляют собой регулярный ряд дислокации. Угол разориентировки обратно пропорционален расстоянию между дислокациями (см. задачу 1.22).

Плотность дислокаций N_D (размерность N_D равна обратному квадрату длины) характеризует степень совершенства кристалла. Для кристаллов с ковалентной связью, таких, как кремний или германий, можно вырастить (при строгом соблюдении необходимых условий) образцы с плотностью дислокаций $N_D < 10^6$ м$^{-2}$ *. Однако чаще всего твердые тела имеют плотность дислокаций порядка 10^8 м$^{-2}$, а у металлических кристаллов она не меньше, чем 10^{11} м$^{-2}$. В монокристаллах при тепловом равновесии существование дислокаций *не обязательно*, поскольку с дислокациями связана значительная энергия, а вклад их в энтропию очень мал. Однако в процессе кристаллизации в любом случае возникает некоторое количество дислокаций, которые в свою очередь могут быть источниками других дислокаций.

Механические напряжения вызывают движение дислокаций в кристалле, которое сопровождается генерацией точечных дефектов. Перемещение дислокации продолжается до тех пор, пока взаимодействие с примесными атомами или с другими ди-

* В кристаллах с решеткой алмаза и сфалерита особо важную роль играют так называемые 60°-ные дислокации, у которых вектор Бюргерса образует угол 60° с линией дислокации.— *Прим. ред.*

слокациями (деформационное упрочнение) не затормозит этого движения. Дислокационная петля, закрепленная в двух точках на поверхности кристалла, под действием напряжений может генерировать замкнутые дислокационные петли, проникающие в кристалл [36]. Возникновение дислокаций, их свойства и перемещения в кристалле заслуживают отдельного рассмотрения. Тем, кто интересуется этим вопросом, можно рекомендовать книги Фриделя и Амелинкса (см. литературу в конце настоящей главы).

Задачи

1.1. Известно, что в кристалле, в котором связи обусловлены силами Ван-дер-Ваальса, равновесное межатомное расстояние $r_0 = 1.50$ Å, а энергия на 10 % меньше, чем в случае, когда учитываются только силы притяжения. Чему равна характерная длина ρ, входящая в выражение (1.2)?

1.2. В молекуле KF равновесное межъядерное расстояние r_0 равно 2,67 Å, а энергия связи $(-E_i)$ относительно энергии бесконечно удаленных *ионов* на 0,50 эВ на молекулу меньше энергии кулоновского притяжения, что обусловлено отталкиванием при перекрытии электронных оболочек. Известно, что средство к электрону у фтора равно 4,07 эВ на электрон, а первый потенциал ионизации калия составляет 4,34 В. Покажите, что энергия, необходимая для разделения молекулы на два нейтральных атома, равна $-0{,}945\,E_i$.

1.3. На рис. 1.58 показана квадратная решетка гипотетического двумерного ионного кристалла. С помощью метода Эвальда — Эвьена (см. примечание 11) покажите, что при учете соседей данного атома вплоть до пятой координационной сферы значение постоянной Маделунга α равно 1,61. При расчетах следует учесть, что для каждого из квадратов, отмеченных штриховыми линиями, условие электронейтральности выполняется, если считать, что 50 % каждого иона, расположенного на одной из сторон квадрата, принадлежит этому квадрату, а 50 % находится вне его. Ион, расположенный в вершине квадрата, принадлежит ему на 25 %, в то время как остальные 75 % находятся вне данного квадрата. Вам нужно показать, что вклад, который дает площадь первого квадрата в постоянную Маделунга, равен +1,29, а площадь между первым и вторым квадратами дает дополнительный вклад, равный +0,32. Всегда ли вклад от площади, заключенной между двумя последующими квадратами, является положительным? Дайте объяснение ответа.

1.4. Расстояние r_0 между ближайшими соседями в KBr равно 3,30 Å. Бромистый калий имеет кристаллическую решетку каменной соли (для которой $\alpha = 1{,}748$). Сжимаемость этого вещества $\chi = 6{,}8 \cdot 10^{-11}$ м²/Н. Используя уравнения (1.10) — (1.15), покажите, что характерная длина потенциала от-

[36] Генерацию последовательности замкнутых дислокационных петель дислокацией, закрепленной в двух точках, можно сравнить с выдуванием мыльных пузырей из трубки. Модель такого источника была предложена Франком и Ридом [*Frank F. C., Read W. T.*— Phys. Rev., **79**, 722 (1950)], и ее принцип действия был непосредственно продемонстрирован как для ионных соединений, так и для кремния. Для наблюдения дислокаций в кремнии, предварительно декорированных медью, Дэш [*Dash W. C.*— J. Appl. Phys., **27**, 1193 (1956)] использовал инфракрасное излучение. Ему удалось наблюдать и сфотографировать источники Франка—Рида, которые создавали несколько замкнутых дислокационных петель.

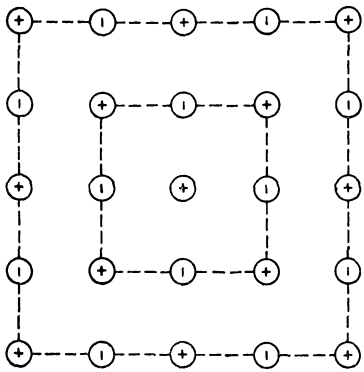


Рис. 1.58. Решетка гипотетического двумерного ионного кристалла, иллюстрирующая метод Эвьена для суммирования ряда при вычислении постоянной Маделунга в задаче 1.3.

талкивания ρ равна $3{,}37 \cdot 10^{-11}$ м, а энергия связи, отсчитанная от энергии удаленных ионов, $E_{\text{равн}} = 6{,}85$ эВ на молекулу.

1.5. Получите выражение для угла между соседними сторонами правильного многоугольника с n сторонами. Обсудите это выражение с точки зрения рис. 1.16 и данных по разрешенным порядкам вращательной симметрии из табл. 1 4.

1.6. Изобразите пять возможных решеток Браве в двумерном пространстве и покажите, где находятся центры инверсии. Укажите для этих решеток точечные группы с самой низкой симметрией, которые совместимы с каждой решеткой и при этом не исключают существование центра инверсии. Можно ли это сделать для каждой из пяти решеток?

1.7. Представьте себе двумерный кристалл, у которого базис состоит из пяти атомов, расположенных в вершинах пятиугольника. Может ли этот пятиугольник быть правильным? Какой симметрией обладает такой кристалл и какова его точечная группа?

1.8. Существование некоторых прямоугольных и квадратных плоских групп обеспечивается инвариантностью относительно операций скольжения (см. рис. 1.20). Объясните, почему скольжение не дает дополнительных плоских групп для других двумерных решеток Браве.

1.9. Нарисуйте примитивные ячейки для семи решеток, представленных на рис. 1.21 стандартными непримитивными ячейками.

1.10. Что произойдет, если углы между примитивными векторами в тригональной решетке будут стремиться к 90 или 120°?

1.11. На рис. 1.25 в качестве примера на использование четырехзначных индексов Миллера в тригональных или гексагональных кристаллах отмечена плоскость $(\bar{1}010)$. Используя этот рисунок, разберитесь в обозначениях и укажите, как расположены плоскости (0001), $(1\bar{2}10)$ и $(\bar{1}103)$ в г. п. у.-решетке (рис. 1.29) и тригональной решетке теллура (рис. 1.6).

1.12. Проверьте приведенные на рис. 1.30 обозначения коэффициентов заполнения различных кубических структур одинаковыми шарами. Покажите, как коэффициенты заполнения различных структур связаны с отношением диаметра шара к длине ребра соответствующего элементарного куба.

1.13. Рассмотрите структуру куприта на рис. 1.34. Изобразите несколько элементарных кубов этого вещества и покажите, что его структура образу-

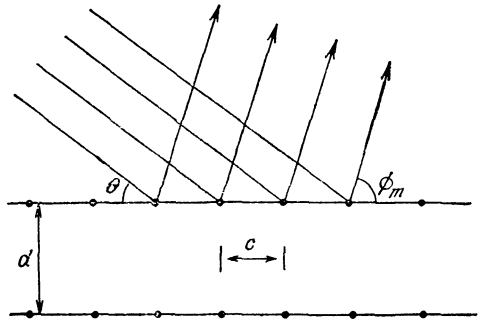


Рис. 1.59. Построение, используемое в задаче 1.14.

ется наложением двух подрешеток, которые вообще не соединены друг с другом связями между ближайшими соседями. Какую структуру будет иметь одна из подрешеток, если убрать атомы меди?

1.14. На рис. 1.59 показаны две параллельные атомные плоскости кристалла (с расстоянием d между ними). Каждая плоскость состоит из рядов атомов, перпендикулярных плоскости рисунка и отстоящих друг от друга на расстояние c. Рентгеновские лучи с длиной волны λ падают на плоскость под углом θ. Лучи, отраженные под углом ϕ_m, обусловлены дифракцией m-го порядка на первой плоскости, действующей как дифракционная решетка. Известно, что эти лучи усиливаются за счет интерференции с лучами, испытавшими дифракцию на нижней плоскости. Покажите, что

$$m\lambda = c\,[\cos\phi_m - \cos\theta],\quad n\lambda = d\,[\sin\phi_m + \sin\theta],$$

где m и n — целые числа. Покажите также, что существует кристаллическая плоскость, составляющая с данным набором плоскостей угол $(1/2)(\varphi_m - \theta)$, для которой дифрагированные лучи, изображенные на рис. 1.59, возникают благодаря зеркальному отражению, удовлетворяющему условию Брэгга.

1.15. Рассмотрите кристаллическую решетку, у которой векторы примитивной ячейки $\mathbf{a}$, $\mathbf{b}$ и $\mathbf{c}$ имеют разную длину, а углы отличны от 90°. Пусть две последовательные плоскости (hkl) расположены на расстоянии d друг от друга. Покажите, не прибегая к обозначениям обратной решетки, что

$$d = \frac{\mathbf{a}\cdot\mathbf{n}}{h} = \frac{\mathbf{b}\cdot\mathbf{n}}{k} = \frac{\mathbf{c}\cdot\mathbf{n}}{l},$$

где $\mathbf{n}$ — единичный вектор, перпендикулярный плоскости (hkl). Из данного выражения при $|\mathbf{a}|=|\mathbf{b}|=|\mathbf{c}|$ и $\mathbf{a}\cdot\mathbf{b}=\mathbf{a}\cdot\mathbf{c}=\mathbf{b}\cdot\mathbf{c}=0$ получите выражение (1.34) для кубической решетки.

1.16. Покажите, что геометрический структурный фактор г. ц. к.-решетки равен нулю для отражений от плоскостей (100) и (110) и отличен от нуля для отражений от плоскости (111).

1.17. Докажите, что вектор, принадлежащий плоскости (hkl), удовлетворяет условиям (1.44) и (1.45).

1.18. Покажите, что для простой кубической решетки обратной решеткой тоже будет простая кубическая решетка. Чему равен угол между векторами $\mathbf{G}_{100}$ и $\mathbf{G}_{111}$? Какой плоскости в прямой решетке перпендикулярен вектор $\mathbf{G}_{100}\times\mathbf{G}_{111}$?

1.19. Найдите удобную комбинацию примитивных векторов трансляций **a**, **b** и **c** для г. ц. к.-решетки. Используя формулы (1.38), определите фундаментальные векторы **a***, **b*** и **c*** обратной решетки. Покажите, что плоскости, проходящие через середины четырнадцати наименьших векторов обратной решетки, образуют зону Бриллюэна в форме усеченного октаэдра. Какой симметрией обладает обратная решетка? Попробуйте, не проводя дополнительных вычислений, описать зону Бриллюэна для о. ц к.-решетки

1.20. Вычислите энергию E_v—E_s образования вакансии для атома калия в кристалле KCl, пренебрегая той частью энергии, которая компенсируется за счет релаксации решетки. Для расчетов используйте данные задачи 1.4. Считайте, что ион калия перемещается на поверхность той грани кристалла, которая совпадает с плоскостью (100).

1.21. Предположим, что для образования вакансии в определенном кристалле необходима энергия 2 эВ. Покажите, что если при этом температура плавления вещества не превышает 1000 К, то отношение плотности вакансий к плотности атомов всегда меньше, чем 10^{-8} %.

1.22. Малоугловая граница между двумя частями кристалла образована дислокациями с вектором Бюргерса **b**, отстоящими друг от друга на расстоянии D. Покажите, что угол разориентировки равен $2\arcsin(b/2D)$.

Литература

Рост кристаллов

Aleksandrov L. Growth of Cristalline Semiconductor Materials on Crystal Surfaces, North-Holland, 1984.

Brice J. C. The Growth of Crystals from Liquids, North-Holland, 1973.

Gilman J. J. The Art and Science of Growing Crystals, Wiley, 1963

Crystal Growth: an Introduction, ed. P. Hartmann, North-Holland, 1973.

Henisch H. K. Crystal Growth in Gels, Pennsylvania State Univ. Press, 1970 [Имеется перевод: *Гениш Г.* Выращивание кристаллов в гелях — М.: Мир, 1973.]

Landise R. A. The Growth of Single Crystals, Prentice-Hall, 1970.

Crystal Growth, ed. B. R. Pamplin, Pergamon Press, 1975.

Pfann W. G. Zone Melting, Wiley, 1966. [Имеется перевод: *Пфанн В.* Зонная плавка.— М.: Мир, 1970.]

* Рост полупроводниковых кристаллов и пленок/Под ред. Л. Н. Александрова.— Новосибирск, 1984.

* Рост и дефекты металлических кристаллов/Под ред. Д. Е. Овсиенко.— Киев: Наукова думка, 1972.

* *Шубников А. В., Парвов В. Ф.* Зарождение и рост кристаллов.— М: Наука, 1969.

* Современная кристаллография. Т. 3. Образование кристаллов.— М.: Наука, 1980.

* *Лодиз Р. А., Паркер Р. Л.* Рост монокристаллов.— М.: Мир, 1974.

Кристаллическое состояние

Gay P. The Crystalline State, Oliver and Boyd, 1972.

Holden A. Nature of Solids, Columbia Univ. Press, 1965. [Имеется перевод: *Холден А. Н.* Что такое ФТТ? — М.: Мир, 1971.]

McKie D., McKie C. Crystalline Solids, Wiley, 1974.

Wells A. F. Structural Inorganic Chemistry, 5th ed., Oxford Univ. Press, 1983. [Имеется перевод 2-го издания: *Уэллс А. Ф.* Строение неорганических веществ.— М.: ИЛ, 1948.]

* *Леммелейн Г. Г.* Морфология и генезис кристаллов — М.: Наука, 1973.

Некристаллические твердые тела

Hlaváč J. The Technology of Glass and Ceramics, North-Holland, 1983.
Amorphous Solids and the Liquid State, eds. N. March, R. A. Street, M. Tosi, Plenum Press, 1985.
Zallen R. The Physics of Amorphous Solids, Wiley, 1983.
* *Мотт Н., Дэвис Э.* Электронные процессы в некристаллических веществах.— М.: Мир, 1974.
* Аморфные полупроводники/Под ред. М. Бродски.— М.: Мир, 1982.
* Физика гидрогенизированного аморфного кремния/Под ред. Дж. Джоунопулоса, Дж. Люковски. Вып. I.— М.: Мир, 1987; Вып. II.— М.: Мир, 1988.

Природа межатомной связи

Evans R. C. An Introduction to Crystall Chemistry, Cambridge Univ. Press, 1964.
Fletcher N. H. The Chemical Physics of Ice, Cambridge Univ. Press, 1970.
Structure and Bonding in Crystals, v. 1, 2, eds. M. O'Keefe, A. Navrotsky, Academic Press, 1981.
Pearson W. B. Crystal Chemistry and Physics of Metals and Alloys, Wiley, 1972.
Pauling L. The Nature of the Chemical Bond, 3rd ed., Cornell Univ. Press, 1960. [Имеется перевод 1-го издания: *Полинг Л.* Природа химической связи.— М.: ИЛ, 1947.]
* *Эванс Р. К.* Введение в кристаллохимию.— М.— Л.: Госхимиздат, 1948.
* *Мотт Н., Герни Р.* Электронные процессы в ионных кристаллах.— М.: ИЛ, 1950.

Операции симметрии и пространственные группы

Buerger M. J. Elementary Crystallography, MIT Press, 1978.
Knox R. S., Gold A. Symmetry in the Solid State, Benjamin, 1964. [Имеется перевод: *Нокс Р., Гольд А.* Симметрия в твердом теле.— М.: Наука, 1970.]
Koster G. F. Space Groups and Their Representations, Academic Press, 1964.
Lax M. Symmetry Principles in Solid State and Molecular Physics, Wiley, 1974.
Megaw H. D. Crystal Structures: a Working Approach, Saunders, 1973.
Phillips F. C. An Introduction to Crystallography, 4th ed., Halsted, 1979.
Salthouse J. A., Ware M. J. Point Group Character Tables, Cambridge Univ. Press, 1972.
Smith J. V. Geometric and Structural Crystallography, Wiley, 1982.
Wyckoff R. W. G. Crystal Structures, Krieger, 1982.
* *Федоров Е. С.* Симметрия и структура кристаллов.— М.: Гостехиздат, 1949.
* Современная кристаллография.— М.: Наука, 1979, т. 1.
* *Белов Н. В.* Структурная кристаллография.— М.: Физматгиз, 1951.
* *Най Дж.* Физические свойства кристаллов.— М.: ИЛ, 1960.

Дифракция в кристаллах и обратная решетка

Bacon G. E. Neutron Diffraction, 3rd ed., Oxford Univ. Press, 1975. [Имеется перевод 1-го издания: *Бэкон Дж.* Дифракция нейтронов.— М.: ИЛ, 1957.]
Brillouin L. Wave Propagation in Periodic Structures, Dover, 1972. [Имеется перевод: *Бриллюэн Л.* Распространение волн в периодических структурах.— М.: ИЛ, 1959.]
Buerger M. J. Crystal Structure Analysis, Krieger, 1979.
Lipson H. S. Crystals and X-Rays, Springer, 1970.
Pendry J. B. Low Energy Electron Diffraction, Academic Press, 1974.
Rymer T. B. Electron Diffraction, Methuen, 1970.
Warren B. E. X-Ray Diffraction, Addison-Wesley, 1969.

Woolfson M. M. An Introduction to X-Ray Crystallography, Cambridge Univ. Press, 1970.
* *Вайнштейн Б. К.* Симметрия кристаллов. Методы структурной кристаллографии.— М.: Наука, 1979.
* *Вайнштейн Б. К.* Структурная электронография.— М.: Изд-во АН СССР, 1956.
* *Липсон Г., Стипл Г.* Интерпретация порошковых рентгенограмм.— М.: Мир, 1972.
* *Бюргер М.* Структура кристаллов и векторное пространство — М.: ИЛ, 1961.
* *Пинскер З. Г.* Дифракция электронов — М.— Л.: Изд-во АН СССР, 1949.
* *Хейкер Д. М.* Рентгеновская дифрактометрия монокристаллов.— Л.: Машиностроение, 1973.
* *Александров Ю. А., Шарапов Э. И., Чер Л.* Дифракционные методы в нейтронной физике.— М.: Энергоиздат, 1981.
* *Кривоглаз М. А.* Теория рассеяния рентгеновских лучей и тепловых нейтронов реальными кристаллами.— М.: Наука, 1967.
* *Гуревич И. И., Тарасов Л. В.* Физика нейтронов низких энергий.— М.: Наука, 1965.
* *Александров Ю. А., Шарапов Э. И., Чер Л.* Дифракционные методы в нейтронной физике.— М.: Энергоатомиздат, 1981.

Дефекты в кристаллах

Amelinckz S. The Direct Observation of Dislocations, Academic Press, 1964. [Имеется перевод: *Амелинкс С.* Методы прямого наблюдения дислокаций.— М.: Мир, 1968.]
Bollmann W. Crystal Defects and Crystalline Interfaces, Springer, 1970.
Cottrell A. H. Theory of Crystal Dislocations, Gordon and Breach, 1964 [Имеется перевод: *Коттрелл А. Х.* Теория дислокаций.— М : Мир, 1969.]
The Chemistry of Extended Defects in Nonmetallic Solids, eds. L. Eyring, M. O'Keefe, North-Holland, 1970.
Flynn C. P. Point Defects and Diffusion, Oxford Univ. Press, 1972.
Friedel J. Dislocations, Addison-Wesley, 1964. [Имеется перевод: *Фридель Ж. Дислокации.*— М.: Мир, 1964.]
Girifalco L. A. Atomic Migration in Crystals, Blaisdell, 1964.
Hirth J. P., Lothe J. Theory of Dislocations, McGraw-Hill, 1968. [Имеется перевод: *Хирт Дж. П., Лоте И.* Теория дислокаций.— М.· Атомиздат, 1972.]
Hull D., Bacon D. J. Introduction to Dislocations, 3rd ed., Pergamon Press, 1984. [Имеется перевод 1-го издания: *Халл Д., Бэкон Д. Д.* Введение в дислокации.— М.: Атомиздат, 1968.]
Kröger F. A. Chemistry of Imperfect Crystals, v. 1—3, 2nd ed., North-Holland, 1973. [Имеется перевод 1-го издания: *Крёгер Ф.* Химия несовершенных кристаллов — М.: Мир, 1969]
Mrowec S. Defects and Diffusion in Solids, North-Holland, 1980.
Parsonage N. G, Stavely L. A. K. Disorder in Crystals, Oxford Univ. Press, 1979. [Имеется перевод: *Парсонидж Н., Стейвли Л.*— Беспорядок в кристаллах.— М.: Мир, 1982.]
Stoneham A. M. Theory of Defects in Solids, Oxford Univ. Press, 1975. [Имеется перевод: *Стоунхэм А. М.* Теория дефектов в твердых телах.— М.: Мир, 1978.]
* *Мильвидский М. Г., Освенский В. Б.* Структурные дефекты в монокристаллах полупроводников.— М.: Металлургия, 1984.
* *Мильвидский М. Г., Освенский В. Б.* Структурные дефекты в эпитаксиальных слоях полупроводников.— М.: Металлургия, 1985.
* *Лейбфрид Г., Бройер Н.* Точечные дефекты в металлах.— М.: Мир, 1981.
* *Вавилов В. С., Кив А. Е., Ниязова О. Р.* Механизмы образования и миграции дефектов в полупроводниках.— М.: Наука, 1981.

2

Динамика решетки

В настоящей главе мы займемся изучением спектра характеристических колебаний кристаллического твердого тела. При этом мы рассмотрим условие распространения волн в периодической решетке, энергию и теплоемкость решетки, корпускулярные аспекты квантованных колебаний решетки (фононов) и следствия ангармонизма взаимодействия между атомами. Обсуждение этих вопросов, составляющих значительную часть физики твердого тела, приведет нас к понятиям разрешенного и запрещенного интервала частот. С этими понятиями мы снова встретимся при рассмотрении электронных спектров твердых тел.

Непрерывные сложные колебания атомов определяют нулевую и тепловую энергию твердого тела. Этим колебаниям соответствует множество фурье-компонент с различными частотами. Если твердое тело подвергается внешним воздействиям, возникает дополнительное движение. Считают, что суммарный эффект такого движения подчиняется *принципу суперпозиции*, т. е. воздействие нескольких возмущений может быть найдено простым их суммированием. Это предположение выглядит правдоподобным в том случае, когда мы находимся в линейной области (или области упругой деформации), т. е. когда возвращающая сила, действующая на каждый атом, приблизительно пропорциональна его смещению (закон Гука). Как мы увидим при обсуждении вопросов, связанных с теплопроводностью, даже при очень небольших смещениях атомов возникают нелинейные эффекты, или «ангармонизм». Эффекты ангармонизма также существенны при взаимодействиях между фотонами и фононами, что мы рассмотрим в последней части разд. 2.3.

2.1. Упругие волны, смещения атомов и фононы

Скорость, с которой продольные волны адиабатически распространяются в жидкости с плотностью ρ, дается выражением

$$v_0 = \lambda\nu = (B_s/\rho)^{1/2}, \qquad (2.1)$$

где B_s — модуль объемной упругости, или коэффициент жесткости. В случае твердых тел это выражение необходимо модифицировать с учетом конечной жесткости. В общем случае для описания распространения звуковой волны в твердом теле следует применять тензор жесткости четвертого ранга. Даже в простейших кристаллических веществах скорость звука, вообще говоря, зависит от направления распространения. Кроме того, в твердых телах могут распространяться и поперечные волны; их скорость меньше скорости продольных волн. Мы будем рассматривать продольные волны в твердом теле, считая, что скорость их распространения правильно определяется выражением (2.1). Чем больше модуль объемной упругости и меньше плотность, тем быстрее будут распространяться звуковые волны.

В табл. 2.1 приведены скорости v_0 для различных твердых тел, вычисленные с помощью формулы (2.1) по известным значениям плотности кристалла и измеренным модулям объемной упругости. Эти величины v_0 можно сравнить со значениями, полученными путем непосредственного измерения скорости звука. Сравнение показывает, что для большинства представленных в табл. 2.1 кристаллов расчетные и экспериментальные значения скорости совпадают с достаточной степенью точности и что скорость звука в типичных металлических, ковалентных и ионных кристаллах равна примерно 5000 м/с. Некоторые значения v_0 будут использованы в дальнейших расчетах, результаты которых приводятся в табл. 2.2 и 2.3 (см. ниже).

Использование макроскопического подхода при описании распространения волн вполне правомерно при условии, что длина этих волн много больше межатомного расстояния. Для волн, амплитуда которых невелика (так что применим закон Гука), тензор жесткости может иметь не более чем 21 независимую компоненту даже в наиболее общем случае трехмерной решетки. Для кубического кристалла число независимых компонент уменьшается до трех. Разумеется, компоненты тензора жесткости теснейшим образом связаны с возвращающей силой, действующей в периодической решетке на каждый атом, смещенный из положения равновесия.

Следует заметить, что теоретические и экспериментальные работы, выполненные в рамках макроскопического приближе-

Таблица 2.1. **Модули объемной упругости и скорость звука для некоторых твердых тел***

Кристалл	Тип структуры	Расстояние между ближайшими соседями r_0, Å	Плотность ρ, кг/м³	Модуль упругости B_S, 10^{10} Н/м²	Расчетная скорость звука $v_0 = (B_S/\rho)^{1/2}$, м/с	Измеренная скорость звука, м/с
Натрий	О. ц. к.	3,71	970	0,52	2320	2250
Медь	Г. ц. к.	2,55	8 960	13,4	3880	3830
Цинк	Г. п. у.	2,66	7 130	8,3	3400	3700
Алюминий	Г. ц. к.	2,86	2 700	7,35	5200	5110
Свинец	Г. ц. к.	3,49	11 340	4,34	1960	1320
Никель	Г. ц. к.	2,49	8 900	19,0	4650	4970
Германий	Структура алмаза	2,44	5 360	7,9	3830	5400
Кремний	То же	2,35	2 330	10,1	6600	9150
SiO_2	Гексагональная	1,84	2 650	5,7	4650	5720
NaCl	Структура каменной соли	2,82	2 170	2,5	3400	4730
LiF	То же	2,01	2 600	6,7	5100	4950
CaF_2	Флюорит	2,36	3 180	8,9	5300	5870

* Данные по модулям упругости заимствованы из статьи Шнайднера [*Gschneidner K. A.* — In: Solid State Physics, v. 16, Academic Press, 1964]. Остальные данные взяты из справочника American Institute of Physics Handbook. 3rd ed., McGraw-Hill, 1971.

ния, играют важную роль для изучения процессов распространения волн в твердых телах. Диапазон частот, в котором звуковые волны распространяются в твердом теле без дисперсии, т. е. произведение длины волны λ на частоту ν является постоянной величиной, равной скорости распространения, очень широк. Этот диапазон привлек пристальное внимание экспериментаторов. Монографии Мэзона (1958) и Хантингтона (1964), приведенные в литературе к данной главе, включают в себя детальное исследование соотношений между тензором жесткости и модами распространяющихся звуковых волн.

Компоненты тензора жесткости можно измерить, пропуская через кристалл ультразвуковые импульсы. Этот метод был предложен Хирмоном[1]. В настоящее время диапазон частот ультразвука, который используется для исследования кристал-

[1] *Hearmon R. F. S.* — Rev. Mod. Phys., 18, 409 (1946).

лов, простирается до значений порядка нескольких ГГц[2]. Однако эти частоты все же далеки от той максимальной частоты, на которой еще возможно распространение волн в твердых телах. Поскольку звук в твердом теле распространяется со скоростью около 5000 м/с, частоте 1 ГГц соответствует длина волны 5 мкм, что более чем в 20 000 раз превышает межатомное расстояние.

Многие детали колебательного спектра решетки твердого тела удается выяснить лишь при учете периодичности кристалла. Макроскопическое рассмотрение оказывается несостоятельным, если длина волны только незначительно больше межатомного расстояния. Вместо этого рассмотрение взаимодействий на атомном уровне дает правильное описание распространения волн как при низких, так и при высоких частотах. Поэтому в настоящей главе мы будем применять главным образом микроскопический подход.

Волна колебаний кристаллической решетки представляет собой повторяющуюся и систематическую последовательность смещений атомов из положений равновесия (продольных, поперечных или некоторых их комбинаций), которая характеризуется следующими параметрами:

скоростью распространения v,

длиной волны λ или волновым вектором $|\mathbf{k}| = 2\pi/\lambda$,

частотой ν или угловой частотой $\omega = 2\pi\nu = vk$. (2.2)

Рассматривая возвращающие силы, действующие на смещенные атомы, можно написать уравнение движения для любого смещения и получить *дисперсионное уравнение,* связывающее частоту с длиной волны. Хотя в некоторых случаях мы будем рассматривать частоту и длину волны, но чаще всего нас будет интересовать соотношение между угловой частотой и волновым вектором.

С классической точки зрения волна, которая удовлетворяет данному дисперсионному уравнению, может иметь любую амплитуду. Однако здесь следует заметить, что, как и квантам электромагнитного излучения и частицам вещества, элементарным квантованным колебаниям решетки свойствен корпускулярно-волновой дуализм. Корпускулярный аспект этих колебаний приводит к понятию *фонона.* Иными словами, распространение в твердом теле волны смещения следует рассматривать как движение одного или многих фононов. При этом каждый

[2] См. статью: *Einspruch N. G.* Solid State Physics, v. 17, ed. F. Seitz, D. Turnbull, Acad. Press, 1965.

См также главу, написанную Якобсеном (*Jacobsen E. H.*), в книге: Phonons and Phonon Interactions, ed. T. A. Bak, Benjamin, 1964.

фонон переносит энергию $h\nu=\hbar\omega$ и импульс $\hbar\mathbf{k}$. Теплопроводность, рассеяние электронов и некоторые другие явления, происходящие в твердых телах, связаны с процессами рождения и уничтожения отдельного фонона. При этом важно учитывать как корпускулярные, так и волновые свойства фононов. Что касается других явлений, то наши возможности выявить дискретную, корпускулярную природу энергии возбуждения зависят от того, насколько велико число термически возбужденных фононов.

Удельная теплоемкость твердых тел была вычислена еще до развития квантовой теории на основании классических представлений о равнораспределении энергии по степеням свободы. Кристалл, состоящий из N атомов, рассматривался как совокупность $3N$ независимых гармонических осцилляторов, каждый с энергией k_0T. Тогда полная энергия колебаний решетки равна $U=3Nk_0T$. При постоянном объеме молярная теплоемкость такого ансамбля вычисляется следующим образом:

$$C_v=(\partial U/\partial T)_v=3N_Ak_0=24{,}94 \text{ Дж/(моль·К)}. \qquad (2.3)$$

Это хорошо известный закон Дюлонга и Пти (1869). Кривые, приведенные на рис. 2.1, характерны для многих твердых тел. Как видно из этого рисунка, классический закон хорошо описывает теплоемкость при комнатной и более высоких температурах и не в состоянии объяснить экспериментальные данные, полученные при низких температурах. Несоответствие между классической теорией и экспериментом при низких температурах является результатом ограничений, накладываемых квантовыми условиями. Это впервые было отмечено Эйнштейном в 1907 г. Позднее соответствующие расчеты были выполнены Дебаем (1912) и Борном и Карманом (1912). Мы обсудим эти расчеты в разд. 2.4.

В области частот, для которых энергия фононов $\hbar\omega$ становится сравнимой с тепловой энергией k_0T или превышает ее, квантовая статистика и классическое равнораспределение энергии дают существенно разные результаты. Учет *корпускулярных* свойств фононов становится важным, когда частота превышает зависящий от температуры порог $\nu_{\text{пор}}=(\omega_{\text{пор}}/2\pi)=(k_0T/h)$. При комнатной температуре этот порог соответствует $\nu_{\text{пор}}\approx 6\cdot 10^{12}$ Гц или $\omega_{\text{пор}}\approx 4\cdot 10^{13}$ рад/с, т. е. значительно превышает частоты, применяемые в акустических установках. Соответствующая пороговая длина волны равна

$$\lambda_{\text{пор}}=2\pi v/\omega_{\text{пор}}. \qquad (2.4)$$

Как мы уже заметили, волны смещения движутся со скоростью около 5000 м/с. Поэтому при комнатной температуре корпускулярные аспекты будут играть роль только для волн с $\lambda\leqslant$

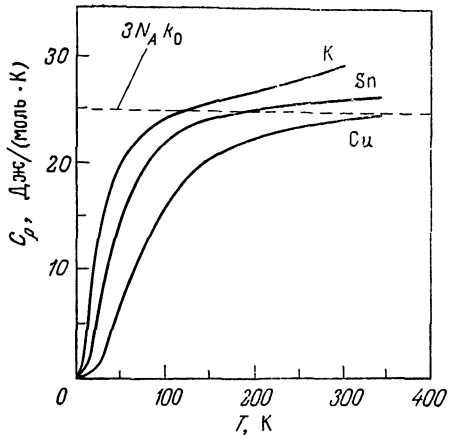


Рис. 2.1. Температурные зависимости теплоемкости при постоянном давлении для типичных металлов. Поскольку C_p и C_v для твердых тел почти одинаковы, можно сравнивать измеренное значение C_p с теплоемкостью, подчиняющейся классическому закону Дюлонга и Пти. Расхождение с классическим законом объясняется с помощью квантовой теории теплоемкости Дебая. Данные, приведенные здесь, взяты из справочника: American Institute of Physics Handbook, McGraw-Hill, 1971.

$\leqslant 10^{-9}$ м. Кроме того, следует заметить, что при таких длинах волн важную роль должна играть дискретная атомная структура твердого тела, поскольку межатомные расстояния в твердых телах лежат в пределах $(1-4)\cdot 10^{-10}$ м. Однако при очень низких температурах условия, связанные с квантовой природой, будут ограничивать амплитуду даже тех волн, длина волны которых много больше атомных размеров.

Таким образом, для объяснения многих вещей, в частности *амплитуды* волн смещения в твердых телах, нам придется учитывать корпускулярную природу фононов. Однако для рассмотрения энергетического *спектра* колебаний решетки достаточно волнового приближения, учитывающего периодичность расположения атомов в кристаллической решетке.

2.2. Колебательные моды одноатомной решетки

Линейная одноатомная цепочка

Плоскую волну, распространяющуюся в однородном твердом теле вдоль оси x, можно записать в виде

$$u = A\exp[i(kx-\omega t)], \tag{2.5}$$

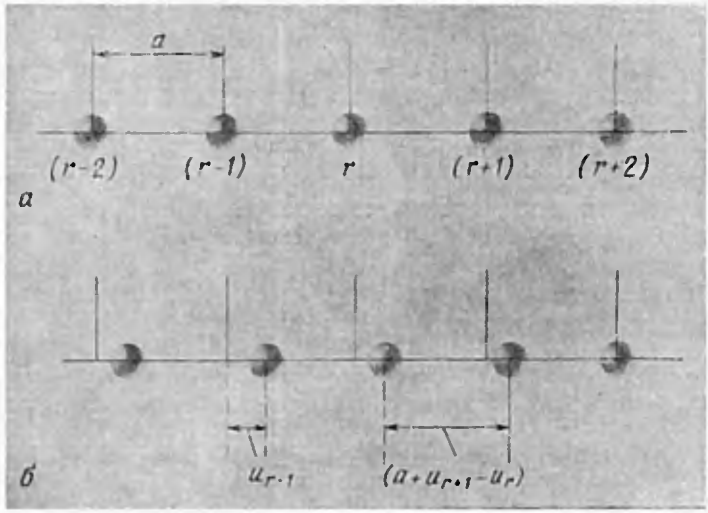


Рис. 2.2. Пять соседних атомов в линейной одноатомной решетке. *а* — атомы находятся в равновесных положениях; *б* — атомы смещены вследствие прохождения продольной волны.

где u — смещение, A — амплитуда, k — волновой вектор и ω — угловая частота. Рассмотрим линейную цепочку одинаковых атомов (рис. 2.2), имеющих массу m и расположенных на расстоянии a один от другого. Для этой последовательности атомов создаваемое волной (2.5) смещение u имеет смысл только в точках расположения атомов и не имеет смысла в промежуточных точках x. Таким образом, смещение r-го атома запишется в виде

$$u_r = A \exp[i(kra - \omega t)]. \tag{2.6}$$

Дифференцируя это выражение два раза, получаем ускорение r-го атома:

$$d^2u_r/dt^2 = -\omega^2 A \exp[i(kra - \omega t)] = -\omega^2 u_r. \tag{2.7}$$

Согласно второму закону Ньютона, возвращающую силу, действующую на r-й атом, можно записать в виде

$$F_r = m(d^2u_r/dt^2) = -m\omega^2 u_r. \tag{2.8}$$

Для установления связи между ω и k эту возвращающую силу необходимо выразить через силовую постоянную, характеризующую смещения атомов в кристаллической решетке. Для этого представим нашу линейную цепочку в виде шариков, связанных пружинками. В этой модели возникающая при

растяжении или сжатии возвращающая сила, действующая на атом, линейно зависит от расстояния до *ближайших соседей* (с коэффициентом пропорциональности μ) и не зависит от положения более удаленных атомов. На первый взгляд кажется, что эта модель не может соответствовать сложному квантовомеханическому характеру взаимодействия между атомами. Однако следует напомнить, что для большинства твердых тел в случае малых деформаций выполняется закон Гука, устанавливающий линейную связь между смещением и приложенной силой.

В приближении закона Гука сила, действующая на r-й атом на рис. 2.2, имеет вид

$$F_r = \mu(u_{r+1} - u_r) - \mu(u_r - u_{r-1}) = \mu(u_{r+1} + u_{r-1} - 2u_r). \quad (2.9)$$

Сравнивая этот результат с выражением (2.8), получаем

$$\omega^2 = (\mu/m)[2 - (u_{r+1}/u_r) - (u_{r-1}/u_r)]. \quad (2.10)$$

Подставляя сюда выражение (2.6) для смещения, последнюю формулу можно переписать в виде

$$\omega^2 = (\mu/m)[2 - \exp(ika) - \exp(-ika)] = 2(\mu/m)[1 - \cos(ka)] = 4(\mu/m)\sin^2(ka/2). \quad (2.11)$$

Таким образом, дисперсионное уравнение для продольных волн, которые могут распространяться в линейной одноатомной цепочке, при учете взаимодействия только с ближайшими соседями запишется как

$$\omega = \pm 2(\mu/m)^{1/2} \sin(ka/2) = \pm \omega_m \sin(ka/2). \quad (2.12)$$

Знаки плюс и минус в выражении (2.12) отвечают волнам, распространяющимся в противоположных направлениях. Движение в любой точке решетки является периодическим во времени.

На рис. 2.3 представлена дисперсионная кривая, построенная в соответствии с уравнением (2.12). Эта кривая качественно подобна той, которую можно получить с учетом зависимости возвращающей силы от расстояний не только до ближайших, но и до более удаленных соседей (см. задачу 2.1).

В качестве независимой переменной на рис. 2.3 вместо длины волны λ выбран волновой вектор $\mathbf{k}$, что более удобно (как мы увидим при разборе других задач физики твердого тела). Область *малых* $\mathbf{k}$ соответствует спектральному диапазону *длинных волн*, для которых вполне годится макроскопическое рассмотрение распространения звуковых волн. В случае когда $ka \ll 1$, мы можем положить $\sin(ka/2) \approx ka/2$. При

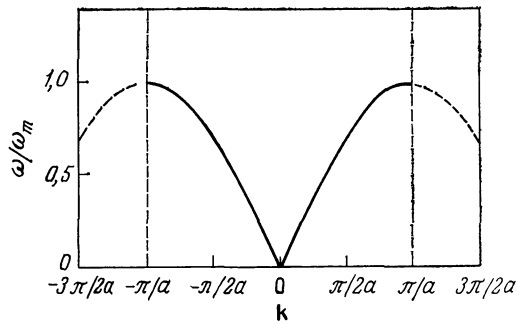


Рис. 2.3. Дисперсионная кривая для продольной волны, распространяющейся в линейной моноатомной решетке; получена с учетом взаимодействия только ближайших соседей [выражение (2.12)]. Область k-пространства, для которой $|\mathbf{k}| \leqslant \pi/a$, образует первую зону Бриллюэна. По оси ординат отложена безразмерная частота в единицах максимальной угловой частоты, для которой величина k вещественна, $\omega_m = 2(\mu/m)^{1/2} = 2(v_0/a)$. Поскольку звуковые волны в кристаллах распространяются, как правило, со скоростью $v_0 \approx \approx 5000$ м/с, эта максимальная угловая частота имеет порядок 10^{14} рад/с.

этом соотношение между угловой частотой и волновым вектором принимает вид

$$\omega \approx v_0 k, \qquad (2.13)$$

$$v_0 = a(\mu/m)^{1/2} \quad (ka \ll 1).$$

В области больших длин волн или низких частот дисперсия отсутствует, так что фазовая скорость ω/k и групповая скорость $d\omega/dk$ совпадают и оказываются равными скорости звука v_0. Скорость $v_0 = a(\mu/m)^{1/2}$, полученная для больших длин волн согласно этой модели, полностью согласуется с выражением (2.1) для непрерывной упругой среды, поскольку в одномерном случае m/a равна плотности, а $a\mu$ — модулю объемной упругости. Для трехмерного кубического кристалла $\rho = m/a^3$, а $B_s = \mu/a$.

Даже в рамках грубой одномерной модели, если известна скорость v_0 обычных звуковых волн в твердом теле, то можно рассчитать межатомный коэффициент жесткости по формуле

$$\mu \approx m(v_0/a)^2 \qquad (2.14)$$

и, следовательно, модуль объемной упругости $B_s = \mu/a$ или сжимаемость $K = 1/B_s$. При составлении табл. 2.1 мы решали и обратную задачу, а именно по измеренным значениям B_s были вычислены значения v_0.

Рассматриваемое на данном этапе низкочастотное приближение справедливо вплоть до частот 10^{12} Гц или около того,

что перекрывает так называемый «акустический» или «ультразвуковой» диапазон частот, в котором работают экспериментальные монохроматические установки. Однако из рис. 2.3 мы видим, что по мере перехода ко все более коротким волнам величина ω достигает предельного значения $\omega_m = (2\mu/m)^{1/2} = 2v_0/a$, когда $|\mathbf{k}| = \pi/a$, и что в области промежуточных частот наблюдается сильная дисперсия скорости звука. В общем виде зависимости фазовой и групповой скоростей от $\mathbf{k}$ имеют вид соответственно

$$v = \omega/k = v_0 \left[\frac{\sin(ka/2)}{ka/2}\right],$$

$$v_g = \partial\omega/\partial k = v_0 \cos(ka/2). \qquad (2.15)$$

При рассмотрении теории теплоемкости Дебая (разд. 2.4) мы увидим, что при выборе функциональной зависимости скорости от волнового вектора иногда допускается некоторый произвол.

Групповая скорость обращается в нуль, когда соседние атомы движутся в противофазе, что происходит при $\mathbf{k} = \pm\pi/a$ или при длине волны $\lambda = 2a$. Не имеет смысла описывать прохождение волны, обеспечивающей сдвиги фаз между колебаниями соседних атомов, превышающие 180°. Поэтому не имеют физического смысла значения $\mathbf{k}$, для которых $|\mathbf{k}| > \pi/a$ (участки кривой на рис. 2.3, изображенные штриховой линией). Нетрудно убедиться в том, что смещения атомов, удовлетворяющие выражению (2.6) для волновых векторов $\mathbf{k}$, абсолютные значения которых больше π/a, полностью эквивалентны смещениям с любым другим волновым вектором $\mathbf{k}'$, который удовлетворяет равенству $\mathbf{k}' = \mathbf{k} + \mathbf{G}$, где $\mathbf{G}$ — вектор обратной решетки[3]. Таким образом, первая зона Бриллюэна содержит полный спектр.

Из этих замечаний очевидно, что волна, соответствующая границе зоны Бриллюэна, т. е. условию $|\mathbf{k}| = \pi/a$, является *стоячей волной*, а не бегущей. Ее длина волны удовлетворяет условию брэгговского отражения на угол 180°. В дальнейшем мы увидим, что в случае двумерного или трехмерного кристалла возбуждения (такие, как фононы или электроны) также испытывают брэгговское отражение, если их волновой вектор упирается в границу зоны Бриллюэна. В случае более чем одного измерения волна с волновым вектором $\mathbf{k}$, соответ-

[3] Напомним (см. разд. 1.5), что вектор обратной решетки $\mathbf{G}$ соединяет эквивалентные точки в двух зонах Бриллюэна. Для рассматриваемого нами одномерного кристалла вектор обратной решетки $\mathbf{G}$ кратен $2\pi/a$, а первая зона Бриллюэна простирается от $\mathbf{k} = -\pi/a$ до $\mathbf{k} = \pi/a$. В реальных кристаллах максимальная величина $\mathbf{k}$ имеет порядок 10^{10} м$^{-1}$, а максимальная угловая частота $\omega_m \approx 10^{14}$ рад/с.

ствующим границе зоны Бриллюэна, в реальном пространстве должна иметь нулевую компоненту групповой скорости в направлении, перпендикулярном границе зоны Бриллюэна (представляющей собой линию или поверхность в пространстве волновых векторов $\mathbf{k}$).

Волна, у которой угловая частота больше, чем $\omega_m = 2v_0/a$, не может распространяться в воображаемом одномерном кристалле, поскольку в соответствии с выражением (2.12) она должна иметь комплексное значение волнового вектора. Наличие мнимой компоненты волнового вектора приводит к сильному затуханию волны. Таким образом, волны, для которых $\omega > \omega_m$, лежат в запрещенной области частотного спектра. Мы убедимся, что этот вывод вполне справедлив и для реальных трехмерных кристаллов. Однако, прежде чем перейти к обсуждению колебаний атомов в трехмерных кристаллах, обобщим некоторые из понятий, которыми мы до сих пор пользовались.

Колебания трехмерного одноатомного кристалла

При описании волнового движения в одноатомном одномерном кристалле мы сочли удобным рассмотреть только продольные смещения в линейной цепи. Конечно, можно рассматривать чисто продольные или чисто поперечные колебания и в трехмерных кристаллах, но только для относительно симметричных структур (например, кубических кристаллов) и для тех особых кристаллографических направлений, которые соответствуют плоскостям атомов, движущихся синхронно. В кубических кристаллах волна может быть полностью продольной или полностью поперечной, если она распространяется в направлениях [100], [110] или [111], и этим обусловлен выбор направлений в $\mathbf{k}$-пространстве, для которых приведены экспериментальные дисперсионные кривые на рис. 2.4, 2.12 и 2.13.

Однако в общем случае волна в трехмерном кристалле имеет как продольную, так и поперечную компоненты; следовательно, уравнения, описывающие ее распространение, должны содержать члены, в которые входят коэффициенты жесткости и скорости, связанные и с продольными, и с поперечными смещениями. Поэтому нетрудно представить себе, что дисперсионные кривые для колебаний в трехмерном кристалле оказываются значительно более сложными по сравнению с кривыми на рис. 2.3. Это подтверждается двумя экспериментально полученными спектрами, изображенными на рис. 2.4. В каждом из представленных на этом рисунке одноатомных кристаллов возбуждаются одна продольная и две поперечные

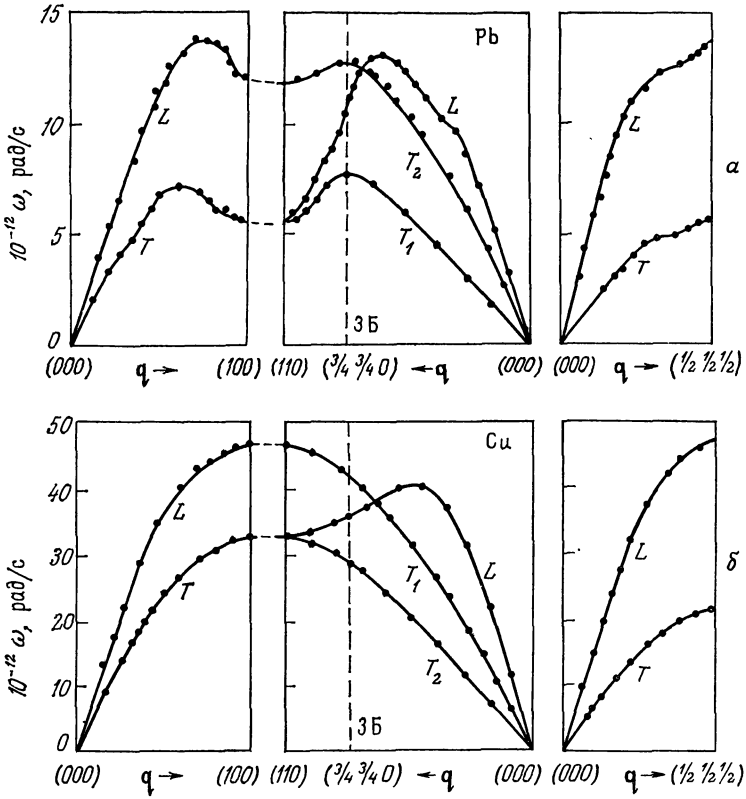


Рис. 2.4. Дисперсионные кривые фононов для свинца (*а*) и меди (*б*), кристаллизующихся в г. ц. к.-структуре. Угловая частота ω приведена как функция безразмерного вектора $\mathbf{q}=\mathbf{k}a/\pi$, измеряемого от центра зоны Бриллюэна в трех основных кристаллографических направлениях. Для направления [110] кривые продолжены за границу зоны Бриллюэна (ЗБ). Данные для свинца взяты из работы: *Brockhouse et al.*— Phys. Rev., **128**, 1099 (1962), а для меди из работы: *Svensson et al.*— Phys. Rev., **155**, 619 (1967); *Nilsson G., Rolandson S.*— Phys. Rev., **B7**, 2393 (1973). Все кривые получены с помощью неупругого рассеяния пучка монохроматических нейтронов.

волны, причем последние две совпадают в направлениях высокой симметрии [100] и [111].

Спектры типа показанных на рис. 2.4 были получены при исследовании *неупругого рассеяния* медленных нейтронов. В разд. 1.4 нас прежде всего интересовала та дополнительная информация о структуре кристалла, которую можно получить при исследовании *упругого* рассеяния нейтронов, подчиняющегося условию Брэгга. Тогда же мы отметили, что для изуче-

ния колебаний решетки можно использовать тепловые нейтроны, имеющие малые скорости. Теперь настал момент, когда нам необходимо вернуться к процессам неупругого рассеяния.

Мы уже отметили, что колебание решетки с угловой частотой ω эквивалентно движению фононов с энергией $\hbar\omega$ и квазиимпульсом $\hbar\mathbf{k}$. Использование понятия *квазиимпульс* не означает, что через кристалл передается «обычный» импульс с соответствующей скоростью; однако квазиимпульс — это величина, которая сохраняется в различных процессах взаимодействия, включающих в себя рождение и уничтожение фононов. Поэтому торможение нейтрона при возбуждении в решетке колебательной моды подчиняется как закону сохранения энергии, так и закону сохранения квазиимпульса, причем последний сводится к сохранению волнового вектора.

Скорости нейтронов $\mathbf{v}$ соответствует волновой вектор

$$\mathbf{K}_n = M_n\mathbf{v}/\hbar \tag{2.16}$$

и кинетическая энергия

$$E = \hbar^2 K_n^2/2M_n. \tag{2.17}$$

Предположим, что при поглощении или рождении фонона энергия и волновой вектор нейтрона изменились и стали равными E' и $\mathbf{K}'_n$ соответственно. Тогда для угловой частоты ω и волнового вектора $\mathbf{k}$ фонона, участвующего в процессе взаимодействия, будут выполняться следующие законы сохранения:

$$E - E' = \pm\hbar\omega, \tag{2.18}$$

$$\mathbf{K}_n - \mathbf{K}'_n = \mathbf{G} \pm \mathbf{k}. \tag{2.19}$$

Здесь вектор $\mathbf{G}$ равен либо нулю, либо вектору обратной решетки. Его можно включить в выражение (2.19), поскольку фонон с волновым вектором $\mathbf{k}$ эквивалентен фонону с волновым вектором $\mathbf{k}+\mathbf{G}$. Закон сохранения волнового вектора в простейшем случае иллюстрируется рис. 2.5.

В процессе неупругого рассеяния, описываемого выражениями (2.18) и (2.19), исследование можно проводить как для фотонов, так и для электронов. Однако из-за значительного различия скоростей света и звука лишь малая доля энергии фотонов может быть передана решетке. Более удобными для этих исследований оказались тепловые нейтроны, поскольку в процессах неупругого рассеяния благодаря их скоростям и энергии их волновой вектор претерпевает значительные изменения как по величине, так и по направлению.

Попытки изучения колебаний решетки с помощью неупругого рассеяния нейтронов возникли в начале 1950-х годов в связи с разработкой урановых реакторов — мощных источников медленных нейтронов. Первые успешные результаты были

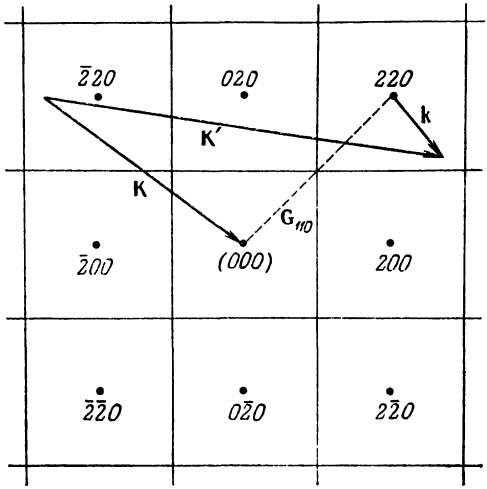


Рис. 2.5. Сечение **k**-пространства для кристалла с простой кубической решеткой. Показаны первая зона Бриллюэна и восемь окружающих ее зон. Координаты центра каждой зоны выражены в единицах безразмерной переменной $\mathbf{q}=\mathbf{k}a/\pi$. **K** и **K′** — волновые векторы нейтрона соответственно до и после неупругого рассеяния, в результате которого рождается фонон с волновым вектором **k**. Векторное соотношение (2.19) между **K**, **K′** и **k** выполняется при учете вектора обратной решетки.

опубликованы в 1955 г.[4] Этот метод оказался весьма полезным, и сейчас у нас имеются сведения о колебательных спектрах для многих твердых тел[5]. Изменение скорости нейтрона вследствие рассеяния измеряется либо методами брэгговской дифракции, либо времяпролетными методами.

Свинец и медь, спектры колебаний которых представлены на рис. 2.4, кристаллизуются в г. ц. к.-структуру. Различия ме-

[4] Первые успешные исследования спектра колебаний решетки с помощью неупругого рассеяния нейтронов были выполнены на алюминии Брокхаузом и Стюартом [*Brockhouse B. N., Stewart A. T.*— Phys, Rev., **100**, 756 (1955)] и на бериллии и ванадии Картером, Хьюджем и Палевски [*Carter R. S., Hughes D. J., Palevsky H.*— Phys. Rev., **99**, 611 (A) (1955)].

[5] Исчерпывающее обсуждение вопросов, связанных с неупругим рассеянием нейтронов и его применением для изучения дисперсионных кривых колебаний решетки, представлено в двух работах Брокхауза [*Brockhouse B. N.*— In: Phonons in Perfect Lattices and in Lattices with Point Imperfection, ed. R. W. H. Stevenson, Plenum Press, 1966; *Brockhouse B. N.*— In: Phonons and Phonon Interactions, ed. T. A. Bak, Benjamin, 1964]. Основное внимание в этих работах уделено результатам, полученным с использованием брэгговской дифракции для анализа импульсов падающих и рассеянных нейтронов. Для этого требуется трехосный спектрометр. [Неупругое рассеяние медленных нейтронов на колебаниях решетки обсуждается также в монографии: *Гуревич И. И., Тарасов Л. В.* Физика нейтронов низких энергий.— М.: Наука, 1965.— *Прим. ред.*]

жду фононными спектрами этих металлов обусловлены главным образом тем, что в них удаленные по отношению к данному атомы оказывают разное влияние. Для меди, спектры которой представлены в нижней части рисунка, предполагалось[6], что можно пренебречь влиянием атомов из координационных сфер дальше восьмой и что для большинства кристаллографических направлений силы взаимодействия с соседями из первой координационной сферы преобладают над силами взаимодействия с другими более удаленными атомами. В то же время количественное объяснение более сложного характера дисперсионных кривых для свинца (рис. 2.4, *а*) требует того, чтобы мы учитывали влияние соседей по крайней мере из десяти ближайших координационных сфер[7].

Кроме того, на дисперсионных кривых для свинца наблюдаются небольшие аномалии, обусловленные иной причиной. В металле сила взаимодействия электронов с фононами зависит от соотношения между волновым вектором фонона и удвоенным волновым вектором свободных электронов, имеющих наибольшую энергию. В случае когда волновой вектор фонона превышает удвоенный волновой вектор таких электронов, возникают *коновские аномалии*[8], которые наблюдаются для многих металлов с сильным электрон-фононным взаимодействием.

Качественно аналогичные приведенным на рис. 2.4 кривые получены и для других твердых тел с одноатомным базисом. Выбор наиболее важных направлений и положения максимумов на оси абсцисс при построении дисперсионных кривых определяются формой зоны Бриллюэна для рассматриваемой кристаллической структуры. (Зона Бриллюэна для металлов, спектры которых приведены на рис. 2.4, рассматривается в задаче 2.3).

Число мод; плотность состояний

В упругой среде могут возбуждаться любые колебания, однако число различающихся друг от друга мод колебаний решетки конечного размера из дискретных атомов строго ограничено.

Для того чтобы найти распределение этих мод по частотам или волновым векторам, рассмотрим линейную одноатомную цепочку из $N+1$ атомов, имеющую длину Na. Предположим, что концевые атомы, находящиеся на концах це-

---

[6] *Nilsson G., Rolandson S.*— Phys. Rev., **B7**, 2393 (1973).
[7] *Brockhouse B. N. et al.*— Phys. Rev., **128**, 1099 (1962).
[8] *Kohn W.*— Phys. Rev. Lett, **2**, 393 (1959).

почки, закреплены, т. е. $u_1=0$ и $u_{N+1}=0$. Тогда в этой цепочке могут возбудиться лишь такие продольные и поперечные колебания, для которых на длине Na уложится 1, 2, 3... или N полуволн. Волновые векторы этих разрешенных колебательных мод удовлетворяют соотношениям

$$k=\frac{\pi}{Na},\ \frac{2\pi}{Na},\ \frac{3\pi}{Na},\ \ldots \text{ или } \frac{\pi}{a}. \tag{2.20}$$

Заметим, что эти колебания разделены одинаковыми интервалами.

Разумеется, при больших N интервал очень мал, и при этом удобно ввести число *состояний* (число различимых колебательных мод); в области от k до $k+dk$ число состояний равно $(Na/\pi)\,dk$. При этом число состояний на *единицу длины* одномерного кристалла в интервале dk запишется в виде

$$\begin{aligned} g(k)\,dk&=(1/\pi)\,dk, \quad && k\leqslant \pi/a,\\ g(k)\,dk&=0, && k>\pi/a. \end{aligned} \tag{2.21}$$

Будем рассматривать $g(k)$ как *плотность состояний* (на единицу длины и на единичный интервал k) в пространстве волновых чисел. Для выбранного нами одномерного случая мы получили, что $g(k)$ не зависит от k в разрешенном спектральном интервале. В дальнейшем мы увидим, что для значительной части других случаев $g(k)$ имеет явную зависимость от k.

Может показаться, что приведенный выше расчет плотности состояний не совсем правильный из-за принятых нами допущений о закреплении атомов на концах линейной цепочки (поскольку такое условие приводит к тому, что существуют только стоячие, а не бегущие волны). Поэтому рассмотрим бесконечную линейную цепочку атомов, расположенных на расстояниях a друг от друга. Пусть в ней распространяются бегущие волны с тем лишь ограничением, что смещения в любой разрешенной моде повторяются через расстояние $L=Na$. Таким образом, мы имеем $u_r=u_{N+r}$ и т. д.[9] Моды, которые мо-

---

[9] Эти условия называются периодическими граничными условиями Борна и Кармана [*Born M., Kármán Th., von.*— Phys. Zeit., **13**, 297 (1912)]. Как мы покажем в последующих главах, искусственность периодических граничных условий не существенна при рассмотрении колебательных и электронных спектров кристаллов. Иногда вводят не периодические, а **циклические** граничные условия. Их можно представить себе **таким образом:** цепь из N атомов замыкается в кольцо, так что $u_r=u_{N+r}$, поскольку r-й атом *совпадает* с $N+r$-м. Большинству из нас циклические граничные условия в трехмерном случае представить себе труднее, чем периодические. Воображение — это все, что нужно, поскольку введенное число N выпадает из выражения для плотности состояний на *единицу длины* или *единичный объем* кристалла.

гут распространяться при этих условиях, должны иметь следующие волновые векторы:

$$k = \pm \frac{2\pi}{L}, \pm \frac{4\pi}{L}, \pm \frac{6\pi}{L}, \ldots \text{ или } \pm \frac{\pi}{a}, \qquad (2.22)$$

где знаки плюс и минус соответствуют волнам, распространяющимся в противоположные стороны.

Из сравнения выражений (2.20) и (2.22) с очевидностью следует, что в случае одномерного кристалла для стоячих или бегущих волн плотность состояний $g(k)$ на единичный интервал $|\mathbf{k}|$ равна $1/\pi$. Действительно, можно утверждать, что плотность состояний бесконечной цепочки не зависит от наложенных граничных условий. Но бесконечная линейная цепочка атомов — это не то, с чем мы имеем дело в реальности. Однако полученный нами факт, что граничные условия не играют существенной роли для линейной цепочки, позволяет утверждать, что в реальном трехмерном кристалле плотность состояний как функция волнового вектора, частоты или энергии не зависит от формы и природы поверхности кристалла, но лишь при условии, что размеры кристалла значительно превышают атомные размеры.

В задачах физики твердого тела нередко встречаются случаи, когда необходимо знать зависимость плотности состояний не от k, а от каких-то других переменных. (Это касается как электронов, так и фононов.). Например, во многих случаях необходимо иметь не только зависимость $g(k)$, но и $g(\omega)$. Для линейной одноатомной цепочки, которую мы только что рассмотрели, можно записать

$$g(\omega) = g(k)(dk/d\omega) = (1/\pi)(dk/d\omega). \qquad (2.23)$$

А поскольку

$$\omega = 2(\mu/m)^{1/2} \sin(ka/2) = \omega_m \sin(ka/2), \qquad (2.24)$$

мы имеем

$$d\omega/dk = (a\omega_m/2)\cos(ka/2) = (a/2)\left[\omega_m^2 - \omega^2\right]^{1/2}. \qquad (2.25)$$

Подставляя это выражение в (2.23), получаем

$$g(\omega) = (2/\pi a)\left[\omega_m^2 - \omega^2\right]^{-1/2} = (1/\pi v_0)\left[1 - (\omega/\omega_m)^2\right]^{-1/2}. \qquad (2.26)$$

Таким образом, $g(\omega)$ *имеет* явную зависимость от ω и действительно обращается в бесконечность на верхнем пределе. К этому выражению для $g(\omega)$ мы будем нередко обращаться, а при рассмотрении квантовой теории теплоемкости в разд. 2.4 мы будем иметь дело с более сложным выражением для $g(\omega)$.

Мы видели, что разрешенные моды для линейной цепочки с заданной периодичностью на длине L образуют последова-

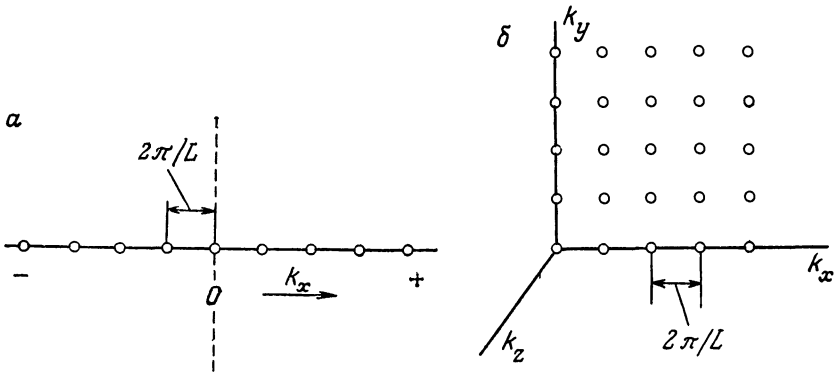


Рис. 2.6. Распределение разрешенных колебательных состояний в k-пространстве, когда периодичность задается длиной L. *а* — для одноатомной линейной решетки; *б* — для трехмерной решетки; показаны состояния лишь в положительном квадранте плоскости $k_x k_y$.

тельность точек в одномерном k-пространстве, расстояние между которыми равно $2\pi/L$. Это представление иллюстрируется рис. 2.6, *а*.

Такое же представление можно применить к разрешенным модам в k-пространстве для трехмерного кристалла. На рис. 2.6, *б* в целях удобства изображено всего лишь несколько разрешенных мод, все остальные смещения разрешенных мод должны быть периодическими с периодом L вдоль каждой оси декартовых координат. Из рисунка видно, что в k-пространстве каждое разрешенное состояние занимает объем $(2\pi/L)^3$. Таким образом, поскольку объем шарового слоя радиусом $k \equiv |\mathbf{k}|$ и толщиной dk с центром в начале координат равен $4\pi k^2 dk$, число разрешенных колебательных состояний на *единицу объема* в интервале dk должно быть равно

$$g(k)\,dk = \frac{4\pi k^2 dk}{L^3 (2\pi/L)^3} = \frac{k^2}{2\pi^2}\,dk. \qquad (2.27)$$

Изображенное на рис. 2.6, *б* распределение мод в k-пространстве будет простираться до границ зоны Бриллюэна во всех направлениях.

В случае трехмерного кристалла вывод зависимости $g(\omega)$ оказывается значительно более сложным, чем для одномерного кристалла, хотя совсем нетрудно показать, что в низкочастотной бездисперсионной области $g(\omega)$ может изменяться как ω^2 (см. задачу 2.4). Вычисление *полного* числа мод является сложной задачей для большинства реальных кристаллов, но, к счастью, мы можем воспользоваться тем, что любая совокупность N атомов в трехмерном пространстве в целом может колебаться $3N$ различными способами. Разумеется, это

число $3N$ равно числу классических степеней свободы у N атомов.

Две трети из этих $3N$ мод (т. е. $2N$) соответствуют поперечным волнам и одна треть (N) — продольным. На языке **k**-пространства это означает, что в объеме зоны Бриллюэна, соответствующей кристаллической структуре, могут разместиться все продольные моды по одной на каждый атом и, кроме того, все поперечные моды из расчета по две моды на каждый атом. Если мы отрежем часть кристалла, то объем зоны Бриллюэна не изменится, но точки в **k**-пространстве, соответствующие различным колебательным состояниям, раздвигаются.

Чтобы обосновать высказанное выше утверждение о том, что в трехмерном пространстве для N одинаковых атомов существует N продольных и $2N$ поперечных мод, рассмотрим простой кубический кристалл с боковыми ребрами длиной L, так что $N=(L/a)^3$. В этом случае зона Бриллюэна представляет собой куб с длиной ребра $2\pi/a$ и объемом $(2\pi/a)^3$. Поскольку, как это видно из рис. 2.6, *б*, каждая разрешенная продольная мода занимает объем $(2\pi/L)^3$, во всей зоне помещается, как и ожидалось, N состояний. Заинтересовавшийся читатель может, хотя и приложив значительно больше усилий, доказать, что в случае г. ц. к.-решетки (рассмотренной в задаче 2.3) объем зоны Бриллюэна таков, что в нем помещается N мод.

Дисперсионные кривые для продольных и поперечных фононов в кристаллографических направлениях высокой симметрии (примеры этих кривых приведены на рис. 2.4) могут дать информацию о наиболее важных атомных силовых постоянных. Эта информация может быть затем использована для численного расчета плотности колебательных состояний в зависимости от частоты. В наш компьютерный век такие расчеты плотности состояний можно выполнить со значительно более высокой точностью, чем это было возможно в первых исследованиях[10]. Результаты такого численного расчета для меди с использованием данных, представленных на рис. 2.4, иллюстрируются кривой на рис. 2.7; при построении этой кривой учитывались смещения, обусловленные влиянием всех ближай-

[10] В классических работах Блэкмана [*Blackman M.*— Proc. Roy. Soc., **A148**, 365 (1935); **A159**, 416 (1937)] показано, как $g(\omega)$ меняется при изменении ω для моделей со сравнительно простой кристаллической структурой, если учитывать взаимодействие с ближайшими соседями и атомами во второй координационной сфере. Расчет $g(\omega)$ оказался возможным благодаря использованию как компьютера, так и вычислительной программы, которая значительно сокращает машинное время [*Gilat G., Raubenheimer L. J.*— Phys. Rev., **144**, 390 (1966)]. В этой программе используется разбиение зоны Бриллюэна на кубические ячейки, а поверхности постоянной частоты внутри каждой из них аппроксимируются набором параллельных плоскостей.

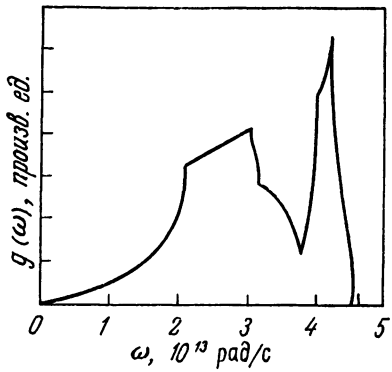


Рис. 2.7. Общая плотность колебательных состояний в меди как функция частоты ω. Кривая построена по результатам численного анализа разных ветвей экспериментальных дисперсионных кривых на рис. 2.4, б. [*Svensson E. C. et al.*— Phys. Rev., **155**, 619 (1967).]

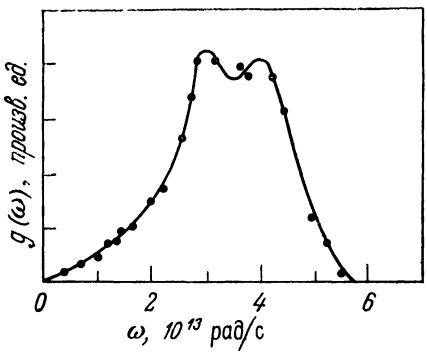


Рис. 2.8. Общая плотность колебательных состояний в ванадии как функция частоты ω. Кривая, построенная по результатам исследования *некогерентного* рассеяния нейтронов, заимствована из работы: *Stewart A. T., Brockhouse B. N.*— Rev. Mod. Phys., **30**, 250 (1958). Эта кривая существенно отличается от результатов работы Картера и др. [*Carter R. S. et al.*— Phys. Rev., **104**, 271 (1956)], в которой указывается на трудности такого рода экспериментов.

ших атомов до шестой координационной сферы [11]. Более 75 % колебательных состояний лежат в области частотного спектра, в которой $g(\omega)$ *не* пропорциональна ω^2.

Для кристаллов, в которых рассеяние медленных нейтронов происходит преимущественно *некогерентно* (из-за хаотичности распределения ядерных спинов), как показали Плачек и Ван Хов [12], функцию $g(\omega)$ можно получить непосредственно, измеряя некогерентную составляющую рассеяния нейтронов. Этот способ определения $g(\omega)$ в особенности интересен в тех случаях, когда некогерентная компонента рассеяния нейтронов в кристалле превосходит когерентную, что затрудняет получение дисперсионных кривых. Найденная таким образом функция $g(\omega)$ для ванадия приведена на рис. 2.8.

Острые пики и резкие перегибы имеются у кривых на рис. 2.7 и 2.8; особенно ярко они выражены на кривой на рис. 2.7. Значения ω, при которых кривые $g(\omega)$ имеют резкие

[11] *Svensson E. C., Brockhouse B. N., Rowe J. M.*— Phys. Rev., **155**, 619 (1967).

[12] *Placzek G., Van Hove L.*— Phys. Rev., **93**, 1207 (1954).

перегибы, называются *критическими точками*, или особенностями Ван Хова [13]. Этим значениям угловых частот соответствуют нулевые групповые скорости фононов в некоторых направлениях, поскольку на границе зоны Бриллюэна изменяется топология поверхности постоянной ω в $\mathbf{k}$-пространстве. Например, самый большой пик на кривой на рис. 2.7 соответствует максимальной разрешенной частоте продольных фононов для направления [110]; аналогичным образом наклон кривой $g(\omega)$ резко уменьшается при частоте, при которой кривая дисперсии поперечных фононов совпадает с границей зоны Бриллюэна для направления [111].

Анализ положения критических точек очень важен при интерпретации колебательных спектров. Как мы увидим позднее, сингулярности Ван Хова наблюдаются и в электронных спектрах кристаллов и тоже очень помогают при определении структуры энергетических зон.

2.3. Колебательный спектр решетки с базисом

До сих пор мы ограничивались обсуждением структур, в которых базис состоит из одного атома. В спектре более сложных структур, в которых на элементарный базис приходится два или больше атомов (независимо от того, являются ли они атомами одного или разных химических элементов), появляются дополнительные особенности. Наиболее примечательной особенностью является чередование разрешенных и запрещенных частотных интервалов [14]. Мы выясним, каким образом

---

[13] Эти *критические точки* впервые описаны в работе Ван Хова [*Van Hove L.*— Phys. Rev., **89**, 1189 (1953)] для колебательных спектров решетки. Термины «критические точки» или «сингулярности Ван Хова» употребляются также при обсуждении плотности *электронных состояний* как функции электронной энергии в кристаллах, поскольку их можно рассматривать аналогичным образом.

[14] Для читателей, знакомых с теорией электрических цепей, заметим, что чередование разрешенных и запрещенных областей ω аналогично чередованию полос пропускания и полос затухания в линиях передачи или комбинациях настраиваемых контуров. Аналогия между теорией цепей и рассмотрением частотных областей вещественного и комплексного $\mathbf{k}$ в кристаллах широко используется в классической монографии Шокли: *Shockley W.* Electrons and Holes in Semiconductors, Van Nostrand, 1950. [Имеется перевод: *Шокли В.* Теория электронных полупроводников.— М.: ИЛ, 1953.] Тесная аналогия между акустическими, электрическими и другими волнами, которые испытывают периодические воздействия, широко используется также в книге Бриллюэна: *Brillouin L.* Wave Propagation in Periodic Structures, Dover, 1953. [Имеется перевод: *Бриллюэн Л., Пароди М.* Распространение волн в периодических структурах.— М.: ИЛ., 1959.]

возникает чередование областей с вещественными и комплексными $\mathbf{k}$, на примере колебаний решетки. В дальнейшем это поможет нам понять аналогичное явление в электронном спектре кристаллов.

Удобно (хотя и не обязательно) начать рассмотрение этих более сложных явлений со случая, когда кристалл состоит из атомов двух сортов.

Нормальные колебания линейной двухатомной цепочки

Представим себе одномерный кристалл, изображенный на рис. 2.9. Он отличается от кристалла, изображенного на рис. 2.2, только тем, что в нем чередуются атомы с массами M и m. Предположим, что $m<M$. Если теперь в кристалле возбуждается продольное возмущение, распространяющееся вдоль цепочки, колебания двух сортов атомов обычно имеют различные амплитуды

$$u_{2r} = A \exp\{i[2kra - \omega t]\},$$
$$u_{2r+1} = B \exp\{i[(2r+1)ka - \omega t]\}. \qquad (2.28)$$

Предположим, что возвращающая сила обусловлена только ближайшими соседями и что смещения не выходят за пределы упругой области, в которой применим закон Гука. Тогда уравнения Ньютона для смещений u_{2r} и u_{2r+1} можно записать в таком же виде, как и уравнения (2.7) и (2.9) для линейной одноатомной цепочки

$$-m\omega^2 u_{2r} = m(d^2u_{2r}/dt^2) = \mu[u_{2r+1} + u_{2r-1} - 2u_{2r}],$$
$$-M\omega^2 u_{2r+1} = M(d^2u_{2r+1}/dt^2) = \mu[u_{2r+2} + u_{2r} - 2u_{2r+1}]. \qquad (2.29)$$

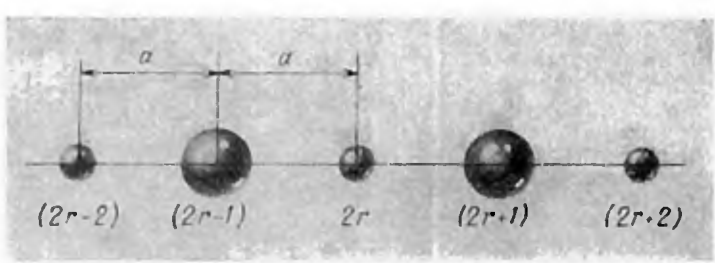


Рис. 2.9. Последовательность соседних атомов в линейной двухатомной цепочке. Масса m атомов меньших размеров меньше массы M бо́льших атомов. Следует заметить, что размер элементарной ячейки, которым определяется размер зоны Бриллюэна, в этом случае равен $2a$.

Подставляя (2.28) в (2.29), получаем два совместных уравнения:

$$-m\omega^2 A = \mu B\,[\exp(ika) + \exp(-ika)] - 2\mu A,$$
$$-M\omega^2 B = \mu A\,[\exp(ika) + \exp(-ika)] - 2\mu B, \qquad (2.30)$$

которые можно представить в следующем виде:

$$A(2\mu - m\omega^2) = 2\mu B\cos(ka),$$
$$B(2\mu - M\omega^2) = 2\mu A\cos(ka). \qquad (2.31)$$

Исключив A и B из этих уравнений, получим зависимость между k и ω, т. е. дисперсионное соотношение:

$$(2\mu - m\omega^2)(2\mu - M\omega^2) = 4\mu^2\cos^2(ka), \qquad (2.32)$$

откуда находим

$$\omega^2 = \mu\left(\frac{1}{m} + \frac{1}{M}\right) \pm \mu\left[\left(\frac{1}{m} + \frac{1}{M}\right)^2 - \frac{4\sin^2(ka)}{mM}\right]^{1/2}. \qquad (2.33)$$

Для линейной одноатомной цепочки, как было показано, существуют одно решение для $k>0$ (волна, бегущая вправо) и одно для $k<0$ (волна, бегущая влево). В рассматриваемом случае данному $\mathbf{k}$ даже в положительном квадранте соответствуют *два* решения для ω. Спектр частот $\omega(k)$ является двузначной функцией $\mathbf{k}$; он приведен на рис. 2.10.

Нижняя ветвь спектра на рис. 2.10 описывается формулой (2.33) с отрицательным знаком. Эта ветвь обычно называется *акустической ветвью*, она соответствует спектру, который уже был получен нами для одноатомной цепочки, за исключением двух следующих особенностей.

1. Для любой совокупности смещений в цепочке, состоящей из атомов двух видов, абсолютная величина волнового вектора не превышает $(\pi/2a)$, в то время как для одноатомной линейной цепочки граница зон Бриллюэна лежит в точках $\pm(\pi/a)$. Это связано с тем, что в двухатомной цепочке размеры зоны Бриллюэна определяются периодом $2a$, а не расстоянием между ближайшими соседями.

2. Из формулы (2.33) следует, что максимальная возможная угловая частота акустических колебаний равна

$$\omega_1 = (2\mu/M)^{1/2} \qquad (2.34)$$

и не зависит от массы более легких атомов в цепочке.

Последнюю особенность можно понять, если рассмотреть зависимость отношения амплитуд колебаний двух сортов атомов от частоты. Из уравнений (2.31) получаем

$$\frac{\text{Амплитуда тяжелых атомов } M}{\text{Амплитуда легких атомов } m} = \frac{B}{A} = \frac{2\mu - m\omega^2}{2\mu\cos(ka)} = \frac{2\mu\cos(ka)}{2\mu - M\omega^2}. \qquad (2.35)$$

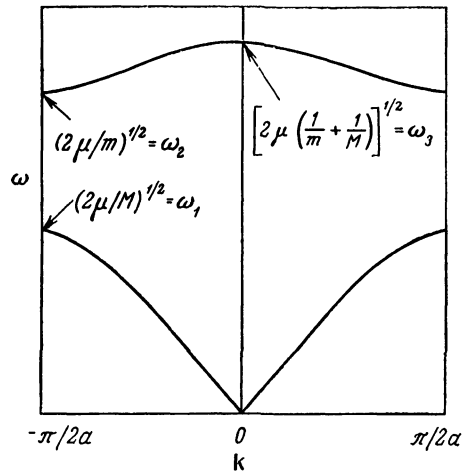


Рис. 2.10. Дисперсионные кривые для продольной волны, распространяющейся в линейной двухатомной решетке. Первая зона Бриллюэна охватывает область **k**-пространства, для которой $|\mathbf{k}|\leqslant\pi/2a$. Нижняя кривая (которую можно сравнить с кривой на рис. 2.3) является *акустической* ветвью, а верхняя — *оптической* ветвью колебательного спектра.

Таким образом, отношение амплитуд почти равно единице (все атомы движутся одинаково) в случае длинных волн и низких частот, т. е. для акустических волн, для которых выполняются следующие условия:

$$\begin{aligned}&\text{волновой вектор } |\mathbf{k}| \ll (\pi/2a),\\ &\text{скорость } v_0=\left[\frac{2\mu a^2}{M+m}\right]^{1/2},\\ &\text{угловая частота } \omega = kv_0 \ll (2\mu/M)^{1/2}.\end{aligned} \tag{2.36}$$

При увеличении **k** и ω в акустической ветви возрастает крайняя правая часть равенства (2.35), при этом возрастает также и отношение B/A. В предельном случае, когда

$$\begin{aligned}&\text{волновой вектор } \mathbf{k}=\pm(\pi/2a),\\ &\text{угловая частота } \omega=\omega_1=(2\mu/M)^{1/2},\\ &\text{фазовая скорость } \omega/k=(8\mu a^2/\pi^2 M)^{1/2},\\ &\text{групповая скорость } d\omega/dk=0,\end{aligned} \tag{2.37}$$

получаем, что $B/A\to\infty$ при $\omega\to\omega_1$. Это означает, что амплитуда A должна быть равна нулю независимо от величины B. Поскольку более легкие атомы не движутся в предельном режиме колебаний на частоте ω_1, не удивительно, что величина m не входит в выражение (2.34) для ω_1.

Заметим, что в некотором интервале частот, превышающих угловую частоту ω_1, не существует решений для вещественных $\mathbf{k}$. Аналогичную ситуацию мы имеем в случае одноатомной цепочки для *всех* частот выше ω_m. Область частот, соответствующих комплексным значениям $\mathbf{k}$, образует запрещенную зону, поскольку любая волна, возбуждаемая на частоте, лежащей в этом диапазоне, будет сильно затухать. Выражение (2.33) и рис. 2.10 указывают на другую интересную особенность: в двухатомной цепочке существует второй диапазон разрешенных частот — область оптических мод колебаний.

Оптическая ветвь

Оптическая ветвь колебаний решетки (т. е. совокупность оптических мод) называется так потому, что в кристаллах, в которых имеется по крайней мере некоторая доля ионности связи, колебания такого типа можно возбудить светом соответствующей частоты. Это объясняется тем, что отношение амплитуд B/A является *отрицательной* величиной во всем частотном диапазоне оптических колебаний; ближайшие соседи движутся не в фазе друг с другом. Такого рода движение возбуждается поперечной электрической компонентой соответствующей электромагнитной волны, когда ближайшие соседи имеют электрические заряды противоположного знака. К этому вопросу мы вернемся позднее в данном разделе.

Для линейной двухатомной цепочки оптическая ветвь описывается формулой (2.33) с *положительным* знаком. Оказывается, что для оптических мод с большими длинами волн (которые занимают область $\mathbf{k}$-пространства, близкую к $k=0$), выполняется любопытный набор соотношений:

$$\begin{aligned}
&\text{волновой вектор } \mathbf{k} \to 0,\\
&\text{угловая частота } \omega \to \omega_3 = \left[2\mu\left(\frac{1}{m}+\frac{1}{M}\right)\right]^{1/2},\\
&\text{фазовая скорость } \omega/\mathbf{k} \to \infty,\\
&\text{групповая скорость } d\omega/d\mathbf{k} \to 0.
\end{aligned} \tag{2.38}$$

Используя соотношение (2.35), получаем следующее отношение амплитуд колебаний двух типов атомов:

$$B/A = -m/M, \text{ если } \begin{cases} \mathbf{k}=0, \\ \omega=\omega_3. \end{cases} \tag{2.39}$$

Таким образом, длинноволновые колебания соответствуют движению атомов в противоположных направлениях, причем общий центр масс пары атомов не движется.

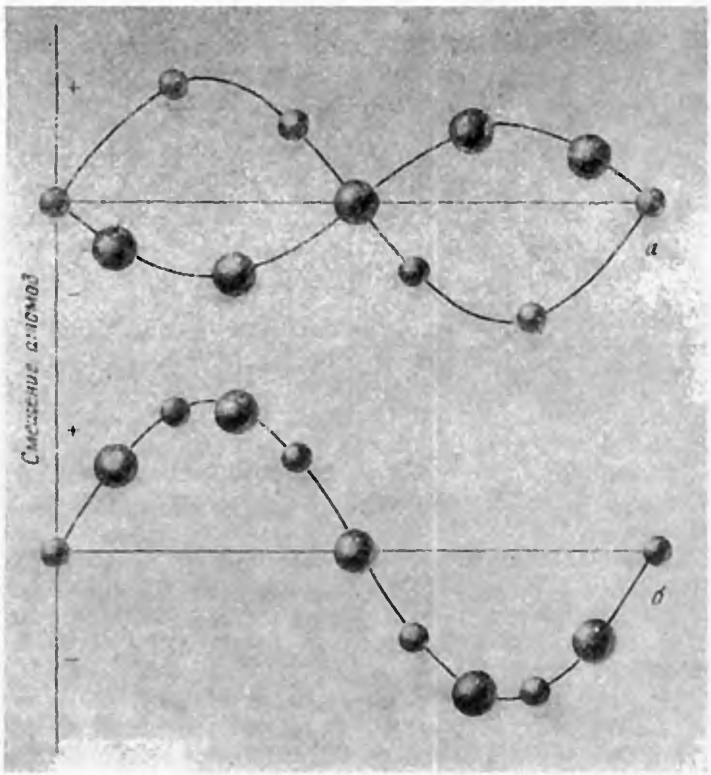


Рис. 2.11. Смещение атомов при распространении поперечной волны вдоль двухатомной линейной цепочки. Различие между акустической (*б*) и оптической (*а*) модами становится более очевидным при схематическом изображении не продольной, а поперечной волны.

До сих пор мы рассматривали только *продольные* колебания двухатомной цепочки. Однако различие между смещениями соседних атомов в акустических и оптических модах становится яснее, если изобразить *поперечные* смещения. Это сделано на рис. 2.11 для волны с некоторым промежуточным значением волнового вектора.

Энергия, которая затрачивается на создание атомных смещений (как продольных, так и поперечных) в случае длинноволновых оптических колебаний намного больше, чем в случае акустических колебаний той же длины волны. Как видно из рис. 2.11 (или аналогичного рисунка для продольной волны), при оптических колебаниях изменения расстояний до атомов второй координатной сферы имеют минимальную величину,

а изменения расстояний между ближайшими соседями максимальны. Поскольку константа жесткости для взаимодействия ближайших соседей обычно гораздо больше, чем для любых других более удаленных соседей, каждый гармонический осциллятор должен запасать значительное количество энергии.

Отношение амплитуд B/A отрицательно для всех оптических мод. Из выражения (2.35) можно видеть, что когда $\mathbf{k}$ возрастает и, следовательно, частота ω убывает вплоть до коротковолнового предела ω_2, отношение B/A стремится к нулю со стороны отрицательных значений. Таким образом, при оптических колебаниях с наиболее короткой из возможных длин волн ($\lambda=4a$) тяжелые атомы не движутся вовсе и в колебаниях участвуют только легкие атомы. Для предельной частоты

$$\omega_2=(2\mu/m)^{1/2} \tag{2.40}$$

групповая скорость $d\omega/dk$ опять оказывается равной нулю. Заметим, что волны, соответствующие частотам ω_1 и ω_2, являются стоячими волнами с нулевой групповой скоростью; они испытывают брэгговское отражение на 180°.

Отношение масс m/M определяет как ширину запрещенной области частот, так и ширину оптической ветви. Когда m и M различаются незначительно, запрещенная область узкая, а отношение предельных частот оптической ветви достигает 1,4 : 1. Однако, если M значительно больше m, запрещенная область оказывается довольно широкой и частоты всех оптических колебаний лежат в узком интервале вблизи $\omega\approx(2\mu/m)^{1/2}$. Со всем этим мы вновь встретимся, когда будем рассматривать теоретические расчеты энергии решетки и решеточной теплоемкости.

Трехмерный кристалл с многоатомным базисом

Деление колебательных мод на акустические и оптические ветви применимо и к колебаниям в трехмерных кристаллах с многоатомным базисом. В этом случае, как и в случае кристаллов с одноатомным базисом, число поперечных мод в два раза превышает число продольных мод. Таким образом, для pN атомов (где p — число атомов в базисе, а N — число базисных групп во всем кристалле) возникнет

N	продольных акустических мод,
$2N$	поперечных акустических мод,
$(p-1)N$	продольных оптических мод,
$2(p-1)N$	поперечных оптических мод.

Следовательно, в кристалле с большим количеством атомов в базисе (например, в монокристалле какого-нибудь органиче-

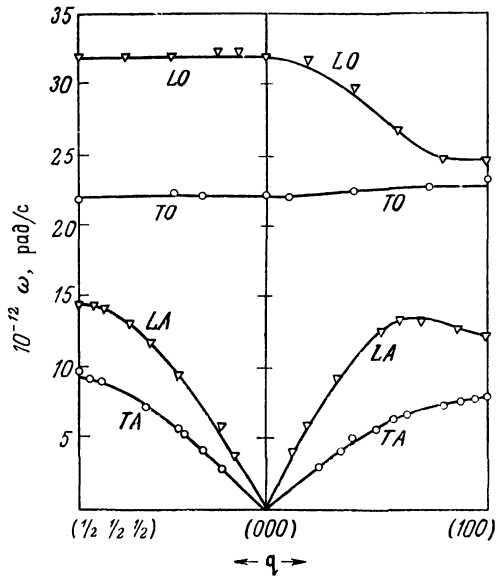


Рис. 2.12. Дисперсионные кривые в случае колебаний решетки иодистого натрия. Приведенный волновой вектор **q** отсчитывается от центра зоны Бриллюэна вдоль двух направлений. Буквами *L*, *T*, *O* и *A* обозначены соответственно продольная, поперечная, оптическая и акустическая ветви. Данные получены при температуре 100 К и взяты из работы: *Woods A. D. B. et al.*— Phys. Rev., **131**, 1025 (1963).

ского соединения с большой молекулярной массой) оптических мод гораздо больше, чем акустических. Спектр $g(\omega)$ оптических мод часто очень сложен. Это объясняется тем, что существует много различных способов возбуждения незвуковых колебаний *внутри* группы атомов, составляющих каждый базис.

Дисперсионные кривые для некоторых кристаллов, фононный спектр которых содержит акустические и оптические ветви (как и для более простых кристаллов), были найдены экспериментальными методами, основанными на явлении когерентного рассеяния нейтронов. На рис. 2.12 показаны типичные дисперсионные кривые для двух направлений в **k**-пространстве для ионного кристалла иодистого натрия. Все дисперсионные кривые для поперечных колебаний двукратно вырождены. Поскольку атом иода гораздо тяжелее атома натрия, запрещенная щель имеет большую ширину.

Фононные дисперсионные кривые не должны меняться с температурой при условии, что температура, при которой проводятся измерения, гораздо ниже температуры плавления.

Смысл этого утверждения становится ясным, если принять во внимание, что гармоническое приближение (закон Гука) справедливо лишь в тех случаях, когда амплитуды смещений атомов малы. Например, измерения на сравнительно тугоплавких твердых телах можно проводить при комнатной температуре. Однако, как отмечают авторы работы [15], цитированной в подписи к рис. 2.12, для того чтобы исключить температурные искажения экспериментальных данных, измерения на кристалле NaI следует проводить при температуре 100 К. *Кривые*, представленные на рис. 2.12, были рассчитаны теми же авторами [16] на основе «оболочечной модели», которая учитывает деформируемость изначально сферических ионов и неполный (менее 100 %) перенос заряда от катиона к аниону. Множество подобных «оболочечных моделей» использовалось для предсказания характера колебательных спектров других ионных кристаллов.

Надо отметить, что частоты продольных и поперечных оптических ветвей колебательного спектра на рис. 2.12 довольно сильно различаются при всех значениях $\mathbf{k}$ (или $\mathbf{q}$). Эта особенность присуща любому частично *ионному* кристаллу. Ниже будет показано, как отношение ω_L/ω_T в длинноволновой области связано с дисперсией диэлектрической проницаемости ионного или частично ионного кристалла в инфракрасной области спектра.

Спектр колебаний полностью *неионных* кристаллов, имеющих кристаллическую структуру с многоатомным базисом, также содержит акустические и оптические ветви. Такая ситуация реализуется, например, в кристалле алмаза, базис которого состоит из двух одинаковых атомов. Соответственно различные ветви дисперсионных кривых алмаза (рис. 2.13) отвечают движению ближайших соседей в фазе и в противофазе. Дополнительная симметрия, являющаяся следствием равенства масс двух атомов базиса, приводит, в частности, к исчезновению щели между верхней точкой продольной акустической (*LA*) и нижней точкой продольной оптической (*LO*) ветвей на границе зоны в направлении [100]. Более того, *LO*- и *TO*-моды в центре зоны вырождены, поскольку из-за отсутствия ионности нет дисперсии диэлектрических свойств. В других неионных кристаллах также имеются акустические и оптические моды, причем длинноволновые *LO*- и *TO*-моды имеют одну и ту же частоту.

---

[15] *Woods A. D. B., Brockhouse B. N., Cowley R. A., Cochran W.* — Phys. Rev., **131**, 1025 (1963).

[16] *Cowley R. A., Cochran W., Brockhouse B. N., Woods A. D. B* — Phys Rev., **131**, 1030 (1963).

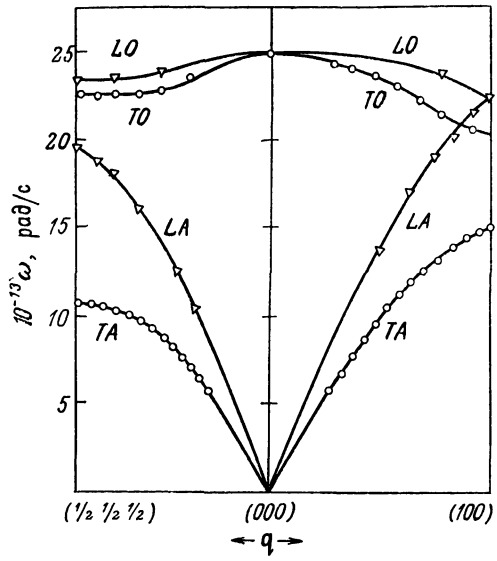


Рис. 2.13. Дисперсионные кривые для кристалла алмаза. Данные взяты из книги: *Warren J. L., Wenzel R. G., Yarnell J. L.* Inelastic Scattering of Neutrons, Vienna, International Atomic Energy Agency, 1965.

Ионная поляризуемость и остаточные лучи

Отношение импульса к энергии для акустических колебательных мод гораздо больше, чем для квантов электромагнитного излучения, поскольку свет распространяется гораздо быстрее, чем звук. Вследствие этого при процессах превращения фотона в акустический фонон не могут одновременно выполняться законы сохранения энергии и импульса, если в них не участвует третий объект, передающий практически весь импульс фонону. Однако такая трудность может и не возникнуть, если мы рассматриваем процессы, в которых происходит аннигиляция фотона и рождение оптического фонона. Так, оптический фонон с очень малым импульсом (соответствующим, возможно, длине волны, равной 10^4—10^5 межатомных расстояний), обладает конечной энергией вблизи максимума кривой рис. 2.10. Как видно из рис. 2.12 и 2.13, длинноволновые оптические фононы обычно обладают энергией 0,01—0,1 эВ. Таким образом, при наличии подходящего механизма взаимодействия генерация длинноволнового оптического фонона (в инфракрасной части электромагнитного спектра) оказывается возможной, при этом избыточный импульс не образуется. Однако при прохождении фотонов через ионную решетку поперечный

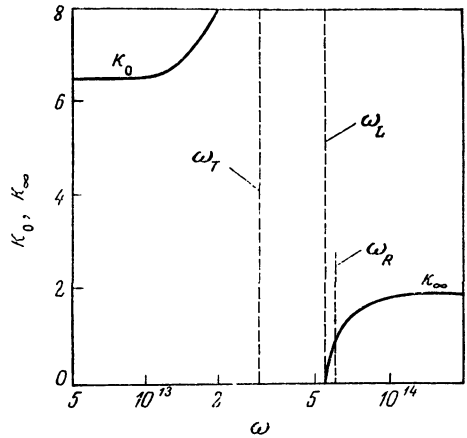


Рис. 2.14. Дисперсия диэлектрической проницаемости в ионном кристалле в области частот остаточных лучей, соответствующих длинноволновой части оптических колебаний. Использовались параметры кристалла фторида рубидия, несмотря на то что величина взаимодействия фотонов и фононов в этом твердом теле достаточно велика для того, чтобы сгладить разрыв при угловой частоте ω_T. Здесь ω_L — частота, для которой $\varkappa=0$ и которая соответствует длинноволновым продольным модам. Частота ω_R — это частота, при которой диэлектрическая проницаемость равна единице, что приводит к коэффициенту отражения, равному нулю.

осциллирующий электрический вектор электромагнитной волны поляризует ее, вынуждая соседние ионы противоположной полярности двигаться в противофазе. В *ионном* кристалле при поляризации неизбежно возникают нелинейные эффекты, приводящие к взаимодействиям, в результате которых фотон аннигилирует и рождается фонон. Название «оптические моды» объясняется существованием такого взаимодействия и сопутствующих ему эффектов. Такой термин применяется независимо от того, является ли связь в данном кристалле в некоторой степени ионной или нет.

Когда электромагнитная волна сравнительно низкой частоты проходит через ионный кристалл, вклад в диэлектрическую проницаемость $\varkappa_0$ вносит как реакция внешних электронов, так и смещение ионных остовов. При большей частоте волны на нее успевают реагировать только внешние электроны; при этом диэлектрическая проницаемость будет иметь гораздо меньшую величину $\varkappa_\infty$ (рис. 2.14). При дальнейшем увеличении частоты в ультрафиолетовой области оптического спектра возникнет дополнительная дисперсия, а при еще более высоких частотах даже электроны в кристалле не успевают реагировать на электромагнитные колебания.

В любом спектральном диапазоне, где есть дисперсия диэлектрической проницаемости, реакция кристалла описывается диэлектрической проницаемостью, имеющей вещественную и мнимую части, отвечающие отставанию по фазе и затуханию [17]. В области дисперсии, лежащей между $\varkappa_0$ и $\varkappa_\infty$, поглощение непосредственно связано с оптическими колебаниями ионов.

Предположим, что через кристалл проходит волна промежуточной частоты. Осциллирующий поперечный электрический вектор, связанный с этими волнами [18], вызывает поперечные колебания катионов и анионов в *противоположных* направлениях. Такой вид движения соответствует оптическим модам. Чем ближе угловая частота электромагнитной волны к ω_T, резонансной частоте поперечных оптических длинноволновых фононов (т. е. фононов с нулевым квазиимпульсом), тем больше амплитуда колебаний ионов. Решение простого дифференциального уравнения

$$\frac{d^2u}{dt^2} + \gamma\frac{du}{dt} + \omega_T^2 u = \frac{-eE_0}{m}\exp(i\omega t) \tag{2.41}$$

имеет вид комплексного смещения

$$u = \frac{-(eE_0/m)\exp(i\omega t)}{\omega_T^2 - \omega^2 + i\gamma\omega}. \tag{2.42}$$

Здесь γ представляет собой параметр, который определяет вклад ангармонических эффектов в процессы затухания (поглощения) электромагнитных возбуждений. Поглощение происходит довольно интенсивно, поскольку ионный кристалл ведет себя как совокупность осцилляторов, характеризующихся одной и той же частотой ω_T и обладающих спектром лоренцевой формы. На рис. 2.15 показана зависимость оптического пропускания тонкого кристалла щелочногалоидного соединения от энергии фотонов, лежащей в инфракрасной области спектра вблизи ω_T. Минимум пропускания соответствует коэффициенту поглощения, равному приблизительно 10^6 м$^{-1}$, или коэффициенту экстинкции (показателю поглощения), равному

---

[17] Мы вернемся к обсуждению вопросов, связанных с диэлектрической проницаемостью, в гл. 5. Дисперсия диэлектрической проницаемости и взаимодействие электромагнитного излучения с оптическими модами колебаний подробно обсуждаются в книге Борна и Хуан Куня (см литературу в конце главы).

[18] Осциллирующий вектор магнитного поля, связанный с этой волной, вызывает гораздо меньшие эффекты, поскольку магнитные взаимодействия определяются скоростями ионов, а не их положением. Таким образом, его влияние оказывается приблизительно во столько раз меньше, во сколько раз наибольшая скорость иона меньше скорости света. Оценить малость этой величины читатели могут самостоятельно (см. задачу 2.9).

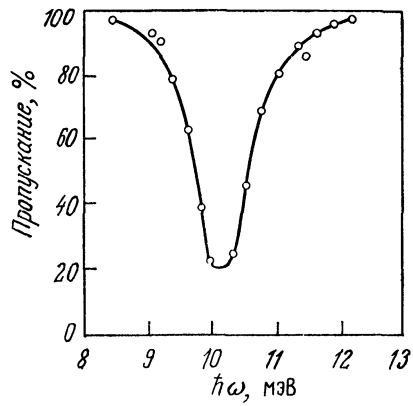


Рис. 2.15. Оптическое пропускание тонкой пленки (толщиной ~1 мкм) иодида рубидия при температуре 4,2 К в области полосы остаточных лучей ИК-спектра. Минимум пропускания имеет место при энергии фотона, равной энергии длинноволнового поперечного оптического фонона. [Из работы: *Jones G. O. et al.*— Proc. Roy. Soc., **A261**, 10 (1961).]

приблизительно 10. Последняя величина значительно больше показателя преломления при более низких энергиях.

Частоту ω_T принято называть частотой *остаточных лучей* (Reststrahlen). Происхождение этого термина объясняется тем, что ионные кристаллы, которые на частотах, близких к ω_T, сильно поглощают излучение, на этих частотах обладают высоким отражением. Обычный метод получения монохроматического излучения в далекой инфракрасной области спектра заключается в облучении немонохроматическим светом *нескольких пластинок*, отражающих на частоте, близкой частоте остаточных лучей. Пластинки изготавливают из кристаллов различных ионных соединений. Такие пластинки достаточно эффективно отражают лишь свет с частотой ω_T, соответствующей данному соединению, и поэтому только это излучение не поглощается полностью. Описанный метод получения спектральных пучков в далекой инфракрасной области спектра уже несколько устарел, поскольку решеточные монохроматоры перекрывают широкий интервал частот от видимой до микроволновой области. Тем не менее пластины для получения света на частоте остаточных лучей все еще находят применение.

Как будет показано в гл. 5, в которой проводится более подробное обсуждение диэлектрических свойств, учет электронных и ионных вкладов в спектральном интервале остаточных лучей приводит к следующему выражению для диэлектрической проницаемости:

$$\varkappa(\omega) = \varkappa_\infty + \frac{[\varkappa_0 - \varkappa_\infty]\,\omega_T^2}{(\omega_T^2 - \omega^2 + i\gamma\omega)}. \qquad (2.43)$$

Из этого выражения видно, что вещественная часть диэлектрической проницаемости становится отрицательной при частотах

ω, превышающих ω_T. В этой области частот показатель преломления мнимый, поэтому слой конечной толщины является идеально отражающим.

В ионных кристаллах угловая частота, соответствующая длинноволновым продольным оптическим фононам ω_L, больше ω_T. Причина такого расхождения связана с характером поляризации макроскопического объема. Когда возбуждается продольная оптическая мода, плоскости, в которых лежат ионы, движутся вперед и назад, и в моменты максимального отклонения на передней и задней поверхностях кристалла возникают слои положительного и отрицательного заряда. Таким образом, создается объемная поляризация кристалла, которой соответствует определенное количество запасенной энергии. При поперечных оптических колебаниях заряженные плоскости осциллируют в поперечных направлениях, макроскопическая поляризация не возникает, поэтому дополнительный вклад в энергию отсутствует.

Соотношение между ω_L и ω_T было исследовано Лидданом, Саксом и Теллером [19], которые показали, что дополнительную энергию, затрачиваемую на создание объемной поляризации продольными оптическими фононами, можно оценить, воспользовавшись соотношением

$$(\omega_L/\omega_T)^2 = \varkappa_0/\varkappa_\infty. \tag{2.44}$$

Подставив соотношение ЛСТ в выражение (2.43), находим, что вещественная часть диэлектрической проницаемости при $\omega = \omega_L$ проходит через нуль и становится положительной. Таким образом, ω_L является предельной верхней частотой для сильного отражения в полосе остаточных частот.

При угловой частоте, несколько превышающей ω_L, диэлектрическая проницаемость становится равной единице. Можно полагать, что отражательная способность на этой частоте близка к нулю, а затем опять медленно возрастает при увеличении частоты. Во многих ионных кристаллах наблюдается максимум, а затем минимум отражения. На рис. 2.16 представлена частотная зависимость отражательной способности сульфида цинка, диэлектрическая проницаемость которого

---

[19] Хорошее обоснование соотношения Лиддана—Сакса—Теллера (ЛСТ) и обсуждение взаимосвязи между свойствами кристалла в инфракрасной области спектра и характеристиками колебаний приводятся в книге Борна и Хуан Куня: *Born M., Huang K.* Dynamical Theory of Crystal Lattices, Oxford, 1954. [Имеется перевод: *Борн М., Хуан Кунь.* Динамическая теория кристаллических решеток.— М.: ИЛ, 1958 г.]. В этой же книге приводится таблица соответствующих параметров для многих ионных кристаллов. Первоначально соотношение ЛСТ опубликовано в работе: *Lyddane R. H., Sachs R. G., Teller E.*— Phys. Rev., 59, 673 (1941).

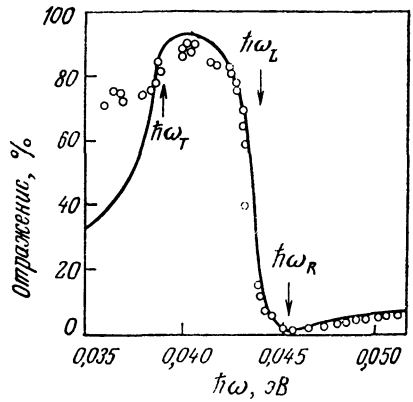


Рис. 2.16. Оптическое отражение сульфида цинка в области остаточных лучей при температуре 300 К [*Deutsch T.*— Intern. Conf. of Semiconductors, Exeter, Institute of Physics, 1962]. Значения энергии фотонов, отмеченные стрелками, согласуются со значениями $\omega_T = 5{,}9 \times 10^{13}$ рад/с, $\varkappa_0 = 6{,}2$ и $\varkappa_\infty = 4{,}7$.

при $\omega > \omega_T$ приблизительно на 25 % меньше, чем при $\omega < \omega_T$. В действительности нулевая отражательная способность наблюдается при некоторой частоте ω_R. На этой частоте, несколько превышающей ω_L, вещественная часть диэлектрической проницаемости равна единице. Из (2.43) и (2.44) можно получить соотношение

$$(\omega_R/\omega_T)^2 = \frac{\varkappa_0 - 1}{\varkappa_\infty - 1} \tag{2.45}$$

при условии, что γ в выражении (2.43) значительно меньше ω_T.

Комбинационные полосы фононов

Как уже отмечалось, оптические моды возникают даже в ковалентных кристаллах, но они не могут быть возбуждены оптическим путем на частоте остаточных лучей. Однако существуют оптические частоты, на которых при поглощении фотона могут возникать *многофононные* взаимодействия. Это — оптические частоты, соответствующие «комбинационным полосам» —полосам поглощения, которые значительно слабее полос, соответствующих остаточным лучам в ионных кристаллах. Фононные комбинационные полосы возникают, когда фотон обладает энергией, достаточной для создания двух или более фононов, полный квазиимпульс которых пренебрежимо мал, или когда фонон с меньшей энергией можно превратить в фонон с большой энергией, но с тем же значением волнового вектора. Если в результате аннигиляции фотона рождаются два фонона с одинаковой энергией, но противоположно направленными импульсами,— мы обычно употребляем термин «полоса двухфононной гармоники», а не «комбинационная полоса».

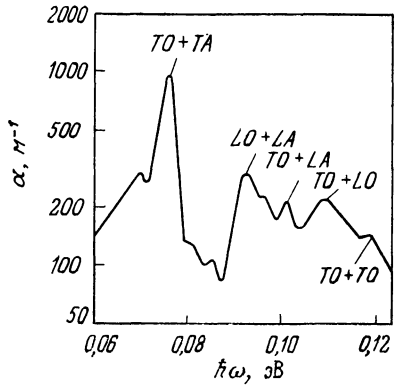


Рис. 2.17. Поглощение ИК-излучения в кремнии (α — коэффициент поглощения), демонстрирующее различные комбинационные полосы фононов и полосу TO-гармоники [*Jonnson F. A.*— Proc. Phys. Soc., 73, 265 (1959)].

Фононные комбинационные полосы могут появиться в спектре оптического поглощения, если разложение потенциальной энергии, соответствующей колебаниям решетки, содержит заметные ангармонические члены[20] или если малые электрические мультиплетные моменты высокого порядка имеют отличную от нуля величину[21]. После того как комбинационные полосы были обнаружены в германии и кремнии[22], количество работ, посвященных теоретическим и экспериментальным аспектам проблемы, растет очень быстро. На рис. 2.17 приведены данные по инфракрасному поглощению в кремнии; здесь отчетливо видно несколько комбинационных полос.

Ответ на вопрос о том, какие комбинационные полосы появятся, а какие — нет, непосредственно связан с симметрией данного кристалла. Можно установить правила отбора[21]. Таким образом, сравнение теории и экспериментальных данных по комбинационным полосам может дать дополнительные сведения о симметрии фононных спектров в критических точках.

Локальные фононные моды

Поглощение фотона, при котором возникает лишь один фонон, возможно и в ковалентных кристаллах при наличии примесей. Такие взаимодействия могут произойти, например, если энергия фотона совпадает с энергией *локальной фононной моды*, связанной с примесью. Как было показано[23], присут-

---

[20] *Born M., Blackman M.*— Z. Physik. **82**, 551 (1933).

[21] *Lax M., Burstein E.*— Phys. Rev., **97**, 39 (1955).

[22] См. работу: *Collins R. J., Fan H. Y.*— Phys. Rev., **93**, 674 (1954). После этой первой работы каждый год появлялись публикации других авторов, посвященные фононным комбинационным полосам в кристаллах. Эта тема освещалась и в трудах международных конференций по полупроводникам (Эксетер, 1962; Париж, 1964; Киото, 1966; Москва, 1968; Кембридж, 1970; Варшава, 1972 и т. д.).

[23] *Dawber P. G., Elliott R. J.*— Proc. Roy. Soc., **A273**, 222 (1963); Proc. Phys. Soc., **81**, 453 (1963).

ствие примесных атомов замещения, масса которых отличается от массы атомов решетки, приводит к тому, что 1) частоты нормальных мод слегка изменяются и 2) возникнут колебательные состояния внутри запрещенной полосы частот, лежащей ниже оптической ветви и выше акустической. Поскольку эти моды характеризуются комплексным значением $\mathbf{k}$, амплитуда колебаний экспоненциально затухает с увеличением расстояния до примесного атома.

Локализованные состояния появляются и в том случае, когда атомы примеси имеют такую же массу, как и атомы основной матрицы, но отличаются от них зарядом [23]. В общем случае примесные атомы отличаются от атомов решетки и массой и зарядом. Примеси нарушают исходную симметрию решетки, вследствие чего становится возможным взаимодействие вектора электрического поля, связанного с электромагнитной волной, с локализованными колебательными модами. Более того, вследствие нарушения симметрии становится возможным взаимодействие электромагнитной волны с нормальными (но слегка сдвинутыми) оптическими модами кристалла. Оба этих эффекта наблюдались в полупроводниковых кристаллах с высокой концентрацией ионизованных примесей.

Рассеяние фотонов фононами

До сих пор мы говорили об *аннигиляции* фотона при рождении фонона. Наряду с этим возможен целый ряд явлений, связанных с *рассеянием* фотонов при рождении (или уничтожении) фононов. Эти эффекты очень слабы, поскольку они в большой степени определяются ангармоническим взаимодействием электромагнитной волны с решеткой.

Эффект комбинационного (рамановского) рассеяния [24] первого порядка представляет собой процесс рассеяния фотона, сопровождающийся образованием или уничтожением одного оптического фонона с нулевым волновым вектором. При этом *стоксовский сдвиг* линии наблюдается в том случае, когда в процессе рассеяния создаются фононы (а фотон теряет часть своей энергии); *антистоксовский* сдвиг имеет место, когда столкновения приводят к уничтожению фононов. Комбинационное рассеяние обычно наблюдается под прямым углом к падающему свету, поскольку это наиболее выгодная ориентация для регистрации рассеяния этого типа, интенсивность которого очень мала.

[24] Этот эффект описан Раманом и Кришнаном (1928 г.), для жидкостей и паров, а для кристаллов — Ландсбергом и Мандельштамом (1928 г.). Обзор работ по комбинационному рассеянию в кристаллах приведен в работе: *Loudon R.*— Adv. in. Phys., **13**, 423 (1964).

Комбинационное рассеяние второго порядка связано с рождением (или уничтожением) пары фононов с равными, но противоположно направленными волновыми векторами. Этот эффект реализуется не на одной частоте, а в целой полосе частот, ширина которой определяется шириной оптической ветви. Анализ комбинационного рассеяния второго порядка тесно связан с анализом критических точек колебательного спектра.

Рассеяние фотона с образованием или уничтожением *акустических* фононов называется бриллюэновским рассеянием[25]. Из-за большой разницы в скоростях света и звука изменение энергии фотона при бриллюэновском рассеянии очень мало (см. задачу 2.10). Поэтому для изучения спектра фотонов, претерпевших бриллюэновское рассеяние, необходима аппаратура высокого разрешения и источники света с высокой степенью монохроматичности. Для исследований как бриллюэновского, так и комбинационного рассеяния в качестве источников света стали широко применяться лазеры.

Несмотря на малость ангармонических членов, которые ответственны за комбинационное и бриллюэновское рассеяние и низкую эффективность рассеяния, при использовании лазерных источников удается получать такие пучки фотонов, интенсивность которых достаточна для наблюдения многих интересных эффектов. Плотность ультразвуковых фононов можно сильно увеличить с помощью импульсного лазера, возбуждающего *вынужденное бриллюэновское рассеяние*, т. е. когерентное усиление ультразвуковых колебаний за счет увеличения рассеяния падающих фотонов[26]. Комбинационное рассеяние лазерного излучения оказалось также полезным для исследования колебательных состояний в кристаллах.

2.4. Статистика фононов и теплоемкость решетки

В последних двух разделах корпускулярный подход к описанию колебаний решетки использовался лишь иногда, например, при обсуждении экспериментов по рассеянию нейтронов или взаимодействия фотонов с фононами. Однако в этих разделах в первую очередь обсуждались дисперсионные соотношения, а при этом обычно удобнее пользоваться волновой тер-

[25] *Brillouin L.*— Ann. Physique, **17**, 88 (1922). См. также статью Лаудона (примечание 24).

[26] Вынужденное бриллюэновское рассеяние лазерного излучения наблюдалось в таких кристаллических твердых телах, как кварц и сапфир. Простая экспериментальная установка для наблюдения этого явления описана в работе: *Chiao R. Y., Townes C. H., Stoicheff B. P.*— Phys. Rev. Lett., **12**, 592 (1964).

мнологией. Теперь, когда мы располагаем сведениями о дисперсионных кривых и плотности мод как функции от $\mathbf{k}$ и ω, для получения зависимости колебательной энергии кристалла от температуры удобнее воспользоваться представлениями о квантованных частицах — фононах. Пусть U — полная колебательная энергия кристалла (в расчете на килограмм, кубический метр или моль). Тогда теплоемкость при постоянном объеме равна $C_v=(\partial U/\partial T)_v$. В реальных экспериментах гораздо удобнее измерять теплоемкость при постоянном давлении C_p, но, к счастью, разность $(C_p—C_v)$ для твердых тел очень мала, поскольку ничтожно малы затраты энергии на тепловое расширение.

Классическая модель для вычисления энергии решетки

Предположим, что атом кристаллической решетки массой m совершает гармонические колебания с амплитудой x_m и угловой частотой ω. Постоянная возвращающей силы равна μ. Если в любой момент времени отклонение атома от положения равновесия равно x, то его скорость $v=\dot{x}$, а ускорение $\ddot{x}=(-\mu x/m)=-\omega^2 x$. Тогда полная энергия, связанная с таким движением, равна:

$$E=(\text{кинетическая энергия})+(\text{потенциальная энергия})=$$
$$=(mv^2/2)+(\mu x^2/2)=(m/2)\,[v^2+\omega^2 x^2]. \tag{2.46}$$

После усреднения по больцмановскому распределению получаем классическое математическое среднее:

$$\langle E\rangle=\frac{\int_{v=0}^{v_m}\int_{x=0}^{x_m} E\exp(-E/k_0T)\,dv\cdot dx}{\int_{v=0}^{v_m}\int_{x=0}^{x_m}\exp(-E/k_0T)\,dv\cdot dx}. \tag{2.47}$$

Подставляя выражение (2.46) в (2.47), легко получить следующий результат (см. задачу 2.11):

$$\langle E\rangle=k_0T. \tag{2.48}$$

Если решетка состоит из N атомов, каждый из которых имеет 3 классические степени свободы, полная энергия решетки равна

$$U=3Nk_0T, \tag{2.49}$$

а теплоемкость решетки определяется ранее приведенным выражением (2.3).

Объяснение того, почему теплоемкость в действительности уменьшается при охлаждении, должно сводиться к объяснению причин, в силу которых средняя энергия, связанная с колебательной модой, зависит от температуры и ее частоты ω.

Модель Эйнштейна

Первый успех в этой области связан с именем Эйнштейна [27], который основывался на разработанной Планком (1900 г.) квантовой теории излучения черного тела. Согласно этой теории квантовый осциллятор, колеблющийся с частотой $\nu = (\omega/2\pi)$, может находиться только в состояниях с энергией

$$E_n = n\hbar\omega, \quad n = 0, 1, 2, 3, \ldots . \tag{2.50}$$

Вероятность пребывания в состоянии n равна

$$g_n = \exp(-E_n/k_0T). \tag{2.51}$$

Тогда средняя энергия этого осциллятора, находящегося в тепловом равновесии, определяется как

$$\langle E\rangle = \frac{\sum_{n=0}^{\infty} E_n \exp(-E_n/k_0T)}{\sum_{n=0}^{\infty} \exp(-E_n/k_0T)}. \tag{2.52}$$

Несложные выкладки (см. задачу 2.12) приводят к следующему результату для квантового гармонического осциллятора, колеблющегося с частотой ω:

$$\langle E\rangle = \hbar\omega/[\exp(\hbar\omega/k_0T) - 1]. \tag{2.53}$$

Другой способ определения средней энергии состоит в вычислении среднего числа фононов, соответствующего при температурах T колебаниям решетки с угловой частотой ω:

$$\langle n\rangle = [\exp(\hbar\omega/k_0T) - 1]^{-1}. \tag{2.54}$$

Величины $\langle n\rangle$ называют *числами заполнения* данной моды [28].

---

[27] *Einstein A.*— Ann. Phys., 22, 180 (1907).

[28] Согласно принципам современной квантовой механики, которые, как известно, были разработаны через двадцать лет после появления модели Эйнштейна, разрешенные значения энергии квантового осциллятора равны $E_n = (n+{}^1/_2)\hbar\omega$, так что

$$\langle E\rangle = \hbar\omega\left\{\frac{1}{2} + [\exp(\hbar\omega/k_0T) - 1]^{-1}\right\}.$$

Дополнительная энергия ${}^1/_2\hbar\omega$ для каждой моды представляет собой *энергию нулевых колебаний;* она названа так из-за того, что эта энергия существует при любой температуре, в том числе и при абсолютном нуле. Энергия нулевых колебаний, которая не зависит от температуры, автоматически входит в энергию решетки. Ее добавление к общей энергии не влияет на величину теплоемкости.

Значение энергии $\langle E\rangle$ стремится к классическому пределу k_0T при температурах гораздо выше $(\hbar\omega/k_0)$, но убывает существенно быстрее, чем k_0T, при низких температурах.

Для простоты Эйнштейн предположил, что кристалл с N атомами обладает $3N$ модами колебаний с одинаковой угловой частотой ω_E. Он использовал эту частоту в качестве подгоночного параметра для согласования своей модели теплоемкости с экспериментальными данными для твердых тел. Тогда, если каждой колебательной моде соответствует одна и та же энергия $\langle E\rangle$, определяемая выражением (2.53), общая колебательная энергия решетки равна

$$U=\frac{3N\hbar\omega_E}{\exp(\hbar\omega_E/k_0T)-1} \tag{2.55}$$

(без энергии нулевых колебаний, которая в модели Эйнштейна не учитывалась). Соответствующая теплоемкость при постоянном объеме равна

$$C_v=\left(\frac{\partial U}{\partial T}\right)_v=3Nk_0F_E(\omega_E,\ T), \tag{2.56}$$

где F_E — *функция Эйнштейна,* равная

$$F_E(\omega_E,\ T)=\frac{(\hbar\omega_E/k_0T)^2\exp(\omega_E/k_0T)}{[\exp(\hbar\omega_E/k_0T)-1]^{-2}}. \tag{2.57}$$

Функция Эйнштейна стремится к единице при высоких температурах, что приводит к классическому результату, т. е. к выражению (2.3). Однако при температурах значительно ниже *характеристической температуры* Эйнштейна $T_E=\hbar\omega_E/k_0$ эта функция убывает экспоненциально:

$$F_E(\omega_E,\ T)\approx(\hbar\omega_E/k_0T)^2\exp(-\hbar\omega_E/k_0T),$$
$$\text{при } T\ll T_E\ F_E(\omega_E,\ T)\approx(T_E/T)^2\exp(-T_E/T). \tag{2.58}$$

Эйнштейн провел сопоставление результатов, полученных с помощью его модели с экспериментальными данными по температурной зависимости теплоемкости алмаза (рис. 2.18). Несмотря на то что согласие между теорией и экспериментом нельзя считать идеальным, модель Эйнштейна явилась значительным шагом вперед по сравнению с классической моделью. Характеристическая температура T_E использовалась в качестве подгоночного параметра для согласования теоретических и экспериментальных *значений* C_v (но не ее производной); привязка осуществлялась при одном значении температуры в низкотемпературной области. Заметим, что параметр Эйнштейна T_E связан с универсальной угловой частотой ω_E, расположенной вблизи центра тяжести дисперсионных кривых для алмаза (см. рис. 2.13).

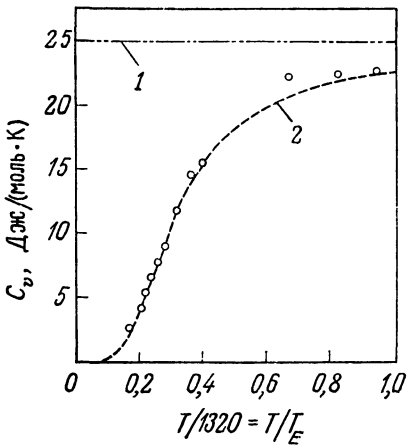


Рис. 2.18. Молярная теплоемкость алмаза. Для сравнения приведены прямая *1*, построенная в соответствии с классическим законом Дюлонга и Пти, а также кривая *2*, соответствующая модели Эйнштейна с характеристической температурой $T_E = 1320$ К, которой отвечает угловая частота $\omega_E = k_0 T_E/\hbar = 1{,}73 \cdot 10^{14}$ рад/с [*Einstein A.*— Апп. Physik, 22, 180 (1907)].

Модель Эйнштейна с физической точки зрения нереалистична, поскольку колебательные моды могут иметь одну и ту же частоту только в том случае, когда все атомы в кристалле совершают колебания независимым образом, что на самом деле не соответствует действительности. Другая, более правдоподобная модель, была предложена Дебаем[29]. Модель Дебая дает для многих кристаллов лучшее согласие с экспериментальными данными по теплоемкости в области низких температур. В этой температурной области C_v часто оказывается пропорциональной T^3. В модели Дебая как раз и получается такая зависимость теплоемкости от температуры, в то время как модель Эйнштейна, согласно выражению (2.58), предсказывает экспоненциальную зависимость.

Модель Дебая

Дебай, как и Эйнштейн, постулировал, что N атомов кристалла должны иметь $3N$ колебательных мод, причем каждая мода обладает энергией, описываемой выражением (2.52), и числом заполнения (2.54). Однако он заметил, что угловая частота ω моды должна зависеть от ее волнового вектора $\mathbf{k}$, причем должна существовать максимальная частота ω_m, такая, что полное число различимых мод равно

$$3N = \int_0^{\omega_m} g(\omega)\, d\omega. \qquad (2.59)$$

[29] *Debye P.*— Ann. Phys., **39**, 789 (1912).

Эта же частота является верхним пределом интеграла, описывающего полную колебательную энергию:

$$U = \int_0^{\omega_m} \frac{(\hbar\omega)\, g(\omega)\, d\omega}{[\exp(\hbar\omega/k_0T) - 1]}. \tag{2.60}$$

Определить *истинную* плотность состояний $g(\omega)$ реального кристалла довольно сложно. Попытки формального решения уравнения (2.60) впервые были предприняты Борном и Карманом[30], а позднее продолжены Блэкманом, опубликовавшим по этому вопросу серию работ, на которые мы уже ссылались ранее[10].

Дебай предположил, что можно получить полезные результаты, если выразить $g(\omega)$ через фазовую скорость $v = (\omega/k)$, положив ее равной соответствующим образом выбранной скорости звука v_0 для *всех* колебательных мод. Тогда необходимо выбрать верхний предел интегрирования ω_D в выражениях (2.59), (2.60) таким образом, чтобы правая часть выражения (2.59) равнялась $3N$. Это приводит к неправильному учету высокочастотных мод, но в силу квантовых ограничений при низких температурах таким модам, не подчиняющимся классическим законам, соответствует очень малое количество фононов. Таким образом, результат не должен существенным образом зависеть от выбора $g(\omega)$ вблизи верхнего края спектра.

Дебаевская модель плотности состояний может быть использована для описания реального или воображаемого кристалла любой размерности. В задаче 2.13 рассматривается модель Дебая для одномерного кристалла. В m-мерном кристалле температурная зависимость теплоемкости при низких температурах в этой модели подчиняется закону

$$C_v = AT^m. \tag{2.61}$$

Это согласуется с тем фактом, что во многих реальных (трехмерных) кристаллах при низких температурах она пропорциональна T^3.

Важными параметрами в модели Дебая является скорость звука v_0 и максимальная частота ω_D. У нас нет возможности выбрать эти параметры независимо при подгонке результатов эксперимента и дебаевской модели. В окончательные выражения обычно вместо ω_D входит *характеристическая температура Дебая* $\theta_D = (\hbar\omega_D/k_0)$ и через нее выражается C_v.

В трехмерной модели Дебая в выражение (2.60) подставляется соответствующая величина $g(\omega)$, а в качестве верхнего предела — величина, удовлетворяющая равенству (2.59). На-

[30] *Born M., von Kármán T.*— Z. Phys., 13, 297 (1912).

помним, что, согласно выражению (2.27), для продольных колебательных мод $g(k)=(k^2/2\pi^2)$, а для поперечных мод $g(k)$ имеет вдвое большую величину. Как показано в задаче 2.4, $g(k)$ можно преобразовать к виду

$$g(\omega)=\frac{\omega^2}{2\pi^2}\left(\frac{1}{v_L^3}+\frac{2}{v_T^3}\right), \quad \omega \ll (v_L/a). \tag{2.62}$$

Мы получили плотность состояний как функцию частоты для акустических колебаний в длинноволновом пределе (в предположении, что продольные и поперечные волны распространяются с различными скоростями). В дебаевской модели принимается, что

$$g(\omega)\approx(3\omega^2/2\pi^2v_0^3), \quad 0<\omega<\omega_D \tag{2.63}$$

для *всех* колебательных мод. Параметр v_0 в выражении (2.63) и предел интегрирования ω_D связаны определенным соотношением, поскольку в кристалле с N атомами в объеме V в единице объема должно содержаться $3N/V$ мод. Таким образом,

$$(3N/V)=\int_0^{\omega_D} g(\omega)\,d\omega=(\omega_D^3/2\pi^2v_0^3). \tag{2.64}$$

Искусственно введенный верхний предел можно выразить через дебаевскую характеристическую температуру

$$\theta_D=(\hbar\omega_D/k_0)=(\hbar v_0/k_0)(6\pi^2N/V)^{1/3}. \tag{2.65}$$

В дебаевской модели оказывается удобнее оперировать с параметром θ_D, имеющим размерность температуры, а не с предельной частотой $\omega_D=(k_0\theta_D/\hbar)$ или максимальным радиусом $k_D=(\omega_D/V_0)$ в обратном пространстве. Чтобы лучше понять модель Дебая, вспомним, что сфера радиусом k_D занимает такой же объем в k-пространстве, как и настоящая зона Бриллюэна. Таким образом, фононы, волновой вектор которых сравним с k_D (частота сравнима с ω_D, энергия — с $k_0\theta_D$), расположены вблизи границ зоны. При всех температурах, кроме высоких, число фононов с такой большой энергией и волновым вектором довольно мало.

На рис. 2.19 очень сложная функция $g(\omega)$ для меди сравнивается с более простой, квадратично возрастающей с ω вплоть до некоторого искусственно введенного предела. Заметим, что использование грубого приближения для высокочастотной части спектра не сильно отражается на результатах. Большой интерес представляет тот факт, что функция $g(\omega)$, полученная из экспериментов по рассеянию нейтронов, на начальном участке спектра возрастает быстрее, чем это следует из модели Дебая. Кривая, отвечающая на рис. 2.19 модели Дебая, построена для

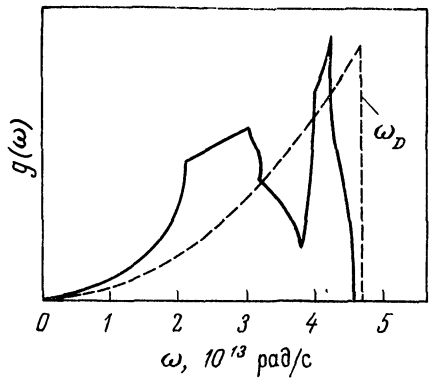


Рис. 2.19. Плотность состояний фононов в меди. Сплошная кривая построена по результатам экспериментального исследования рассеяния нейтронов согласно Свенссону и др. (1967), а также Нильсону и Роландсону (1973), причем эта же экспериментальная зависимость приведена на рис. 2.7. Штриховая кривая соответствует трехмерной модели Дебая и проведена таким образом, что площади под этими двумя кривыми одинаковы. При этом $\omega_D = 4{,}5 \cdot 10^{13}$ рад/с, а характеристическая температура Дебая $\theta_D = 344$ К.

некоторой температуры Дебая θ_D [и соответствующих ей значений ω_D, v_0 и $g(\omega)$], которая вычислялась по результатам исследования теплоемкости меди. Необходимо помнить, что θ_D играет роль *подгоночного параметра*, который обеспечивает наилучшее согласие между экспериментальными и теоретическими значениями теплоемкости.

В методе Дебая плотность состояний $g(\omega)$, определяемую формулой (2.63), следует подставить в выражение (2.60). Тогда для результирующей энергии колебаний решетки на единицу объема получаем

$$U = \left(\frac{3\hbar}{2\pi^2 v_0^3}\right) \int_0^{\omega_D} \frac{\omega^3 d\omega}{[\exp(\hbar\omega/k_0T) - 1]}. \tag{2.66}$$

Переходя к безразмерной переменной $x = (\hbar\omega/k_0T)$ можно записать

$$U = \left[\frac{9Nk_0T^4}{V\theta_D^3}\right] \int_0^{(\theta_D/T)} \frac{x^3 dx}{(e^x - 1)}. \tag{2.67}$$

Дифференцируя по температуре, получаем выражение для теплоемкости

$$C_v = \left(\frac{\partial U}{\partial T}\right)_v = \left[\frac{9Nk_0T^3}{V\theta_D^3}\right] \int_0^{(\theta_D/T)} \frac{x^4 e^x dx}{(e^x - 1)^2}. \tag{2.68}$$

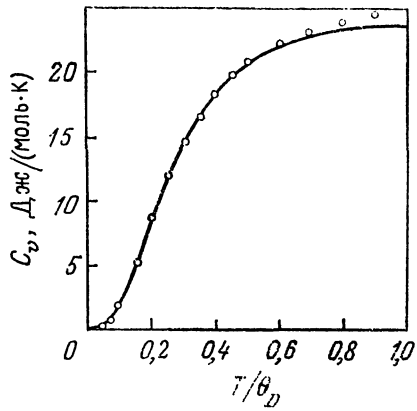


Рис. 2.20. Температурная зависимость молярной теплоемкости твердого тела, построенная по трехмерной модели Дебая. Экспериментальные точки получены для образца иттрия [*Jennings L. D., Miller R. E., Spedding F. H.*— J. Chem. Phys., **33**, 1849 (1960)]. Температура на оси абсцисс нормирована на температуру Дебая $\theta_D = 200$ К. (При самых низких температурах наилучшее согласие может быть достигнуто для несколько большего значения температуры Дебая.) То, что при высоких температурах экспериментальные точки лежат выше теоретической кривой, объясняется тем, что в действительности измеряется C_p, а не C_v.

Выражения для величин U и C_v записываются через интегралы, которые выражаются в аналитическом виде только в пределе высоких и низких температур. Однако численно C_v можно определить для любой температуры[31]. Сопоставление экспериментальных величин и теоретической кривой, полученной по модели Дебая для типичного случая, приведено на рис. 2.20.

Как и следовало ожидать, для температур гораздо выше θ_D интеграл в выражении (2.67) равен $^1/_3(\theta_D/T)^3$, так что для энергии и теплоемкости получаем классические формулы $U \approx (3Nk_0T/V)$ и $C_v \approx (3Nk_0/V)$. Более интересна область низких температур, в которой, согласно эксперименту, C_v убывает не по экспоненциальному закону, как предсказывает модель Эйнштейна, а имеет более слабую зависимость от температуры.

Если для температур, меньших $(\theta_D/10)$, в качестве пределов интегрирования в выражениях (2.67) и (2.68) взять нуль и бесконечность, то это не приведет к большой ошибке. Тогда

---

[31] Подробные таблицы функций Дебая и Эйнштейна приведены в монографии: *Gopal E. S. R.* Specific Heats at Low Temperatures, Plen. Press, 1966. В табулированном виде эти функции можно найти также в справочнике: The American Institute of Physics Handbook, McGraw-Hill, 1971 и Handbook of Mathematical functions, eds. M. Abramowitz, I. A. Stegun, Dover, 1965.

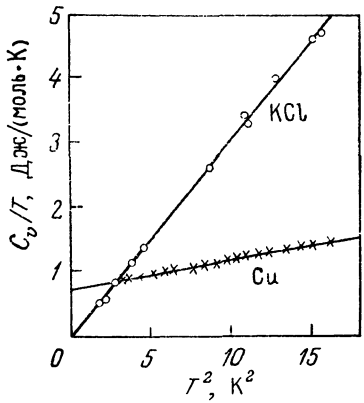


Рис. 2.21. Температурная зависимость теплоемкости KCl и Cu при очень низких температурах. Линейный ход в этих координатах указывает на то, что в C_v содержится член, пропорциональный T^3. В случае KCl этот член является единственным. В случае же меди имеется еще один член, который линейно зависит от температуры; этот член обусловлен электронным вкладом в теплоемкость. Данные для KCl взяты из работы: *Keesom P. H., Pearlman N.*— Phys. Rev., **91**, 1354 (1953), а для меди из работы: *Rosenberg H. M.* Low Temperature Solid State Physics, Oxford University Press, 1963.

интеграл в уравнении (2.67) будет равен $(\pi^4/15)$ и

$$U=\left[\frac{3\pi^4 N k_0 T^4}{5V\theta_D^3}\right] \text{ при } T<(\theta_D/10), \tag{2.69}$$

$$C_v=\left(\frac{\partial U}{\partial T}\right)_v=(12\pi^4 N k_0/5V)\,(T/\theta_D)^3= \\ =1944\,(T/\theta_D)^3 \text{ Дж/(моль·К) при } T<(\theta_D/10). \tag{2.70}$$

Согласно выражению (2.70) C_v характеризуется кубической зависимостью от температуры, что часто и наблюдается на практике для кристаллов при низких температурах. Модель Эйнштейна не может объяснить такую зависимость. При промежуточных температурах зависимость C_v от T с довольно хорошей точностью описывается или моделью Эйнштейна, или моделью Дебая.

Если при низких температурах C_v описывается выражением (2.70), измерения при одной температуре позволяют определить θ_D и предсказать величину C_v при других температурах. Иллюстрацией может служить рис. 2.21, где представлены данные для металла и для диэлектрика. Для кристалла KCl зависимость C_v/T от T^2 линейна и проходит через начало координат. Таким образом, закономерность $C_v \sim T^3$, предсказываемая моделью Дебая, выполняется и величину θ_D можно определить по наклону экспериментальной прямой.

Как видно из рис. 2.21, для меди тоже наблюдается линейная зависимость (C_v/T) от T^2, но соответствующая прямая отсекает на вертикальной оси некоторый отрезок. Это означает, что

$$C_v = AT + BT^3, \tag{2.71}$$

где второй член справа отвечает теплоемкости решетки (закон Дебая), а первый член соответствует теплоемкости газа свободных электронов. Мы подробно рассмотрим вопросы, связанные с энергией и другими свойствами электронного газа в металлах, в гл. 3. Сейчас же достаточно запомнить, что при сравнении экспериментальных данных по теплоемкости с более сложными моделями необходимо принимать во внимание теплоемкость газа свободных электронов, если они существуют в кристалле.

Уточнения модели Дебая

В модели Дебая используется существенное упрощение реальной функции $g(\omega)$. Поэтому можно ожидать, что величина θ_D, полученная путем подгонки при некоторой определенной температуре экспериментального значения C_v к теоретическому, не обязательно будет оптимальной для всех других температур. Однако для любой пары измеряемых величин (C_v и T) можно получить свой параметр θ_D. Совокупность таких измерений позволяет построить зависимость θ_D от T. Обычно всегда имеет место *некоторый* разброс θ_D при изменении температуры, хотя для большинства кристаллов этот разброс менее выражен, чем в случае, приведенном на рис. 2.22.

Основные черты кривой на рис. 2.22 (θ_D сначала убывает, а потом опять возрастает с увеличением T) учитываются во многих теоретических моделях, использующих представления, развитые Эйнштейном и Дебаем. Как уже говорилось, первые попытки расчета $g(\omega)$ для простых решеток, а также зависимостей $C_v(T)$ и $\theta_D(T)$ были предприняты Блэкманом [10]. Блэкман [32] показал также, что θ_D имеет минимум при некоторой промежуточной температуре, если в двухатомной решетке все акустические моды характеризуются дебаевским распределением, а оптические — эйнштейновским [33] (см. задачу 2.15).

---

[32] *Blackman M.* — Repts. Progr. Phys., 8, 11 (1941); Handbuch der Phys., 7, 325, Springer, 1956.

[33] Интересно, что для сравнения своей простой модели с экспериментом Эйнштейн выбрал алмаз. С позиций современных представлений можно считать, что в алмазе оптические моды находятся в узком частотном интервале (см. рис. 2.13), поэтому модель Эйнштейна должна хорошо описывать поведение C_v.

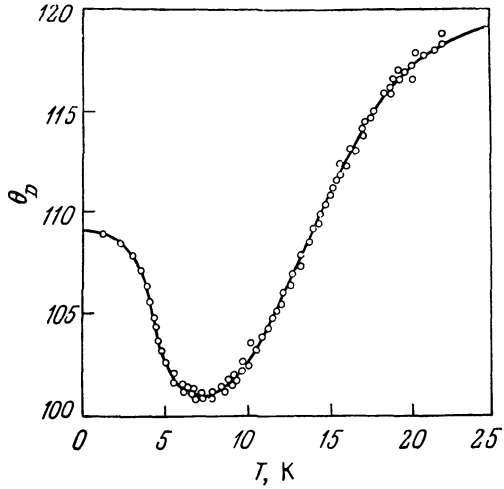


Рис. 2.22. Температурная зависимость характеристической температуры Дебая θ_D (по измерениям теплоемкости) для металлического индия. [Из работы: *Clement J. R., Quinnell E. H.*— Phys. Rev., 92, 258 (1953).] Теплоемкость кристаллической решетки, необходимая для определения температуры Дебая, получена вычитанием электронного вклада из измеренной полной теплоемкости.

Теперь благодаря использованию машинных расчетов колебательных мод, а также экспериментам по рассеянию нейтронов и других новых экспериментальных методик можно получать гораздо более подробные данные о виде зависимости $g(\omega)$. Однако это вряд ли сильно изменит наши представления о решеточной теплоемкости, поскольку C_v и C_p нечувствительны к мелким деталям функции $g(\omega)$. Метод использования изменения температуры для сканирования энергии решетки фактически соответствует спектроскопии с очень плохим разрешением. Тем не менее недавно полученные сложные зависимости $g(\omega)$ дают ценную информацию о важных группах фононов, в частности об их энергии, волновом векторе, направлении распространения, поляризации, способности взаимодействовать с другими фононами или иными возбуждениями.

Температура Дебая

Как это ни удивительно, но понятие температуры Дебая θ_D используется во многих задачах физики твердого тела, в том числе и не связанных с теплоемкостью. Из того, что до сих пор говорилось о модели Дебая, могло сложиться впечатление, что эта модель не учитывает периодичности кристаллической решетки и ограничений на физически реализуемые интер-

валы значений волнового вектора и частоты, связанных с существованием границ зоны Бриллюэна. Однако это не так, поскольку величина ω_D в модели Дебая, естественно, оказывается сравнимой с угловыми частотами фононов, волновые векторы которых близки к границам зоны. Эти фононы составляют *большинство* при температурах выше θ_D. Наоборот, при температурах значительно ниже дебаевской возбуждаются только фононы, волновые векторы $\mathbf{k}$ которых очень близки к центру зоны Бриллюэна и лежат достаточно далеко от ее границ. Поэтому такие явления, как теплопроводность (определяемая ангармоническим взаимодействием фононов друг с другом) и электропроводность (определяемая рассеянием электронов на фононах), существенно различны при температурах выше и ниже температуры Дебая.

Выше температуры Дебая у большей части фононов длина волны имеет порядок нескольких межатомных расстояний. Однако при температурах значительно ниже дебаевской наиболее вероятная длина волны фонона имеет порядок $a\theta_D/T$ (см. за-

Таблица 2.2. **Значения характеристической температуры Дебая для некоторых типичных твердых тел, определенные по результатам измерений упругих постоянных и теплоемкости***

Кристалл	Структура	Скорость распространения упругих волн $v_0 = (B_S/\rho)^{1/2}$, м/с	θ_D, К, определенная по упругим постоянным $(\hbar v_0/k_0) \times (6\pi^2 N/V)^{1/3}$	θ_D, К определенная по теплоемкости при низких температурах
Натрий	О. ц. к.	2320	164	157
Медь	Г. ц. к.	3880	365	342
Цинк	Г. п. у.	3400	307	316
Алюминий	Г. ц. к.	5200	438	423
Свинец	Г. ц. к.	1960	135	102
Никель	Г. ц. к.	4650	446	427
Германий	Структура алмаза	3830	377	378
Кремний	То же	6600	674	647
SiO_2	Гексагональная	4650	602	470
NaCl	Структура каменной соли	3400	289	321
LiF	То же	5100	610	732
CaF_2	Структура флюорита	5300	538	510

* Скорость распространения упругих волн вычислялась так же, как в табл 2.1 Температуры Дебая, полученные из данных по теплоемкости (после вычитания электронной компоненты) для кристаллов, состоящих из химических элементов, приведены в работе: *Gschneidner K. A.* In: Solid State Physics, v. 16, Academic Press, 1961, а для четырех соединений — в книге: American Institute of Physics Handbook, McGraw-Hill, 3rd ed, 1971.

дачу 2.16). Эта длина волны при достаточно низких температурах может иметь порядок нескольких сотен и даже тысяч межатомных расстояний.

Температура Дебая зависит от констант упругости, и, следовательно, от них зависит и температура, при которой будет выполняться сильное неравенство $T \ll \theta_D$. Кристаллы с сильным межатомным взаимодействием (алмаз, сапфир) характеризуются высоким значением θ_D. В таких кристаллах даже при небольшом охлаждении вымерзают все фононы, за исключением тех, длины волн которых очень велики по сравнению с размерами элементарной ячейки.

Температуры Дебая для наиболее характерных кристаллов приведены в последнем столбце табл. 2.2. Все эти величины были получены из данных по решеточной компоненте теплоемкости при низких температурах. Как можно видеть из таблицы, значения характеристической температуры, вычисленные по упругим постоянным, могут быть как больше, так и меньше «тепловой» температуры Дебая.

2.5. Теплопроводность

Корпускулярный или фононный подход к рассмотрению колебаний решетки особенно удобен при изучении процессов с преобразованием энергии. Эти процессы включают процессы *рождения* и *уничтожения* фононов. Теплопроводность удобнее всего описывать на языке *рассеяния* фононов на других фононах, статических несовершенствах решетки или на электронах.

Решеточная теплопроводность и длина свободного пробега фононов

Тепловая энергия может передаваться в кристалле

фононами,
фотонами,
свободными электронами (или свободными дырками),
электронно-дырочными парами,
экситонами (связанными электронно-дырочными парами).

Электронная компонента теплопроводности в металлах обычно является доминирующей. Однако в неметаллических кристаллах большая часть теплового потока переносится колебаниями решетки (фононами); лишь при самых высоких температурах преобладающим процессом может стать передача энергии фотонами. Рассмотрим поток фононов, который возникает при наличии в кристалле градиента температур.

Вспомним, что, согласно выражению (2.54), при температуре T количество возбуждаемых фононов с волновым вектором $\mathbf{k}$ и угловой частотой ω_k равно

$$\langle n_k \rangle = [\exp(\hbar\omega_k/k_0T) - 1]^{-1}. \qquad (2.72)$$

В условиях теплового равновесия, если нет температурных градиентов, $\langle n_k \rangle = \langle n_{-k} \rangle$, т. е. существует полное равенство скоростей фононных потоков в любых двух взаимно противоположных направлениях. Таким образом, общий тепловой поток равен нулю.

Когда имеется температурный градиент ∇T, теплопроводность $\varkappa_l$ можно выразить через ∇T и скорость потока энергии — через единичную площадку, перпендикулярную градиенту температуры:

$$\mathbf{Q} = -\varkappa_l \nabla T. \qquad (2.73)$$

Эту теплопроводность можно выразить через микроскопические характеристики фононов, воспользовавшись аналогией между проводимостью фононного газа и обычного молекулярного газа. Выражение для теплопроводности в рамках кинетической теории газов для простейшего случая, когда все частицы имеют одинаковые скорости (что справедливо для фононов при тем-

Таблица 2.3. **Теплопроводность решетки и средняя длина свободного пробега фонона в неметаллических кристаллах***

Кристалл	$T = 273$ K		$T = 77$ K		$T = 20$ K	
	Теплопроводность $\varkappa_l$, Вт/(м·К)	Средняя длина свободного пробега фонона $\Lambda = (3\varkappa_l/v_0C_v)$, м	Теплопроводность $\varkappa_l$, Вт/(м·К)	Средняя длина свободного пробега фонона $\Lambda = (3\varkappa_l/v_0C_v)$, м	Теплопроводность $\varkappa_l$, Вт/(м·К)	Средняя длина свободного пробега фонона $\Lambda = (3\varkappa_l/v_0C_v)$, м
Кремний	150	$4{,}3\cdot10^{-8}$	1500	$2{,}7\cdot10^{-6}$	4200	$4{,}1\cdot10^{-4}$
Германий	70	$3{,}3\cdot10^{-8}$	300	$3{,}3\cdot10^{-7}$	1300	$4{,}5\cdot10^{-5}$
Кристаллический кварц (SiO_2)	14	$9{,}7\cdot10^{-9}$	66	$1{,}5\cdot10^{-7}$	760	$7{,}5\cdot10^{-5}$
CaF_2	11	$7{,}2\cdot10^{-9}$	39	$1{,}0\cdot10^{-7}$	85	$1{,}0\cdot10^{-5}$
NaCl	6,4	$6{,}7\cdot10^{-9}$	27	$5{,}0\cdot10^{-6}$	45	$2{,}3\cdot10^{-6}$
^{7}LiF (изотопически чистый)	10	$3{,}3\cdot10^{-9}$	150	$4{,}0\cdot10^{-7}$	8000	$1{,}2\cdot10^{-3}$

* Данные взяты из статьи. *Berman R.* — Gryogenics, 5, 297 (1965) и работ, цитированных в ней. Средняя длина свободного пробега вычислялась для каждого значения теплопроводности с использованием значения v_0, взятого из табл 2.1, и теплоемкости C_v (на единицу объема), определенной через «тепловые» значения температуры Дебая, приведенные в табл. 2.2).

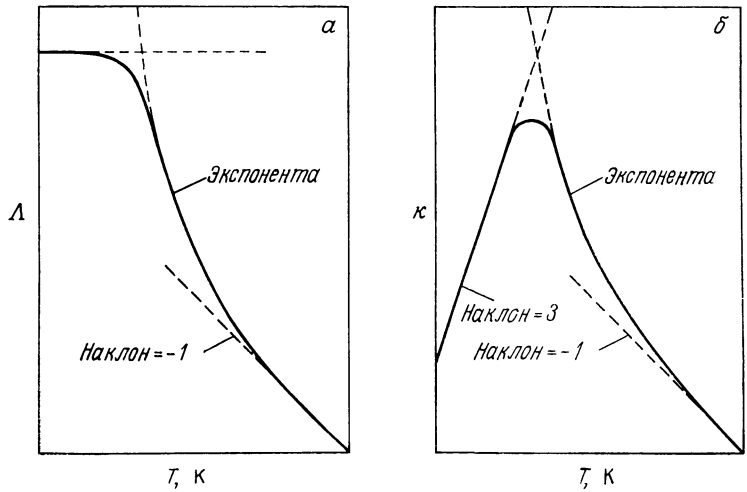


Рис. 2.23. Типичные кривые зависимости средней длины Λ свободного пробега фонона и фононной теплопроводности от температуры в двойном логарифмическом масштабе. Увеличение средней длины свободного пробега при понижении температуры, связанное с уменьшением вероятности процессов переброса (*U*-процессов), прекращается, когда существенным становится рассеяние фононов на дефектах кристаллической решетки и поверхности кристалла.

пературах много ниже температуры Дебая), записывается в виде

$$\varkappa_l = \frac{1}{3} C_v v_0 \Lambda. \qquad (2.74)$$

Здесь C_v — теплоемкость решетки для единицы *объема*, которая является мерой плотности фононов, v_0 — скорость фононов (скорость звука), а Λ — *средняя длина свободного пробега фононов*.

Длина пути, проходимого фононом с момента его рождения до момента уничтожения или иного превращения, часто сильно зависит от его энергии. Она может быть довольно большой для фононов малых энергий, но становится весьма малой для фононов, энергия которых превышает пороговую энергию процессов переброса $k_0\theta_u$ (о которой мы скажем ниже). Тем не менее для любого распределения фононов всегда можно *определить* среднюю длину свободного пробега Λ с помощью выражения (2.74).

Для температур вблизи точки плавления Λ может уменьшаться до 6—10 межатомных расстояний. При очень низких температурах Λ может достигать 1 мм. Общий вид темпера-

турной зависимости Λ и соответствующая зависимость $\varkappa_l$ показаны на рис. 2.23. Мы должны установить, как ведет себя $\Lambda(T)$, чтобы понять поведение теплопроводности.

В табл. 2.3 приведены значения теплопроводности и длины свободного пробега для шести неметаллических твердых тел при трех различных температурах. Значения длины свободного пробега фононов вычислены по формуле (2.74) с использованием данных по скорости звука (см. табл. 2.1) и теплоемкости. Последняя определялась по температурам Дебая, приведенным в табл. 2.2. Из табл. 2.3 видно, в каком широком интервале возможно изменение длины свободного пробега фононов.

Ангармонические эффекты

Если бы закон Гука в твердых телах выполнялся строго и энергию атома в положении r (равновесное положение r_0) можно было бы представить в виде

$$E_r = E_{r_0} + A'(r - r_0)^2, \tag{2.75}$$

то принцип суперпозиции был бы точным для колебаний решетки. В такой ситуации в совершенном кристалле бесконечных размеров взаимодействия (столкновения) фононов невозможны, а их длина свободного пробега Λ становится бесконечно большой. Однако в действительности энергия должна быть записана в виде

$$E_r = E_{r_0} + A(r - r_0)^2 + B(r - r_0)^3 + \ldots + \ldots . \tag{2.76}$$

Ангармонические члены более высоких порядков, чем второй, обусловливают взаимодействие фононов и, следовательно, ограничивают их длину свободного пробега. В отсутствие ангармонических эффектов установление теплового равновесия в системе неравновесных фононов было бы невозможно (в бесконечном идеальном кристалле)!

Теория влияния ангармонических эффектов на фонон-фононные взаимодействия достаточно сложна, и мы не будем обсуждать ее подробно. Более полные сведения по этому вопросу можно почерпнуть из библиографии, приведенной в конце этой главы (в особенности из монографии Займана и книги под редакцией Бэка, цитированной в примечании 5). Здесь лишь заметим, что при достаточно высоких температурах величина Λ должна меняться как T^{-1}, поскольку при температурах выше дебаевской число возбужденных фононов (способных взаимодействовать с данным фононом и уменьшить его длину свободного пробега) пропорционально T. В той же высокотемпературной области Λ и $\varkappa_l$ должны обладать одинаковой темпера-

турной зависимостью, поскольку теплоемкость C_v практически не зависит от температуры.

При рассмотрении более низких температур, для объяснения поведения Λ нужно учитывать два различных типа фонон-фононного взаимодействия, один из которых эффективно ограничивает длину свободного пробега, определяемую выражением (2.74), а другой не влияет на ее величину.

Нормальные процессы и процессы переброса

Допустим, что два фонона, сталкиваясь, рождают третий фонон. Вероятность такого столкновения определяется значением ангармонических членов в выражении для энергии. Свойства рожденного фонона определяются законами сохранения энергии и квазиимпульса.

Ранее при рассмотрении неупругого рассеяния медленных нейтронов на колебаниях решетки кристалла мы отмечали, что квазиимпульс фонона $\hbar\mathbf{k}$ не то же самое, что «обычный» импульс. Так, колебательная мода с волновым вектором $\mathbf{k}$, совпадающим с границей зоны Бриллюэна, представляет собой стоячую волну и обладает нулевым импульсом по отношению к кристаллу в целом. И все же при рождении или уничтожении фонона внешними возбуждениями (например, при рассеянии падающих нейтронов или β-частиц), обычный импульс передается кристаллу. Вопрос о том, сохраняется ли квазиимпульс $\hbar\mathbf{k}$ или нет, зависит от конкретных особенностей данного процесса.

Во вяком случае, квазиимпульс сохраняется при столкновении двух фононов с образованием одного нового, так что

$$\hbar\omega_1 + \hbar\omega_2 = \hbar\omega_3, \tag{2.77}$$

$$\hbar\mathbf{k}_1 + \hbar\mathbf{k}_2 = \hbar\mathbf{k}_3. \tag{2.78}$$

Это законы сохранения энергии и импульса. Такой процесс называется *нормальным процессом*, или N-процессом. Если построить векторную диаграмму, соответствующую выражению (2.78), то можно убедиться, что в случае N-процессов направление потока энергии не меняется. Поэтому такие процессы не дают вклада в тепловое сопротивление. Если бы N-процессы были единственным видом взаимодействия фононов друг с другом в идеальном кристалле, решеточная теплопроводность была бы бесконечно большой [34].

---

[34] Может показаться, что здесь имеется противоречие. Как можно получить бесконечную проводимость решетки, если существуют процессы рассеяния, ограничивающие длину свободного пробега? Ответ состоит в том, что *средняя длина свободного пробега* Λ, определенная выражением (2.74),

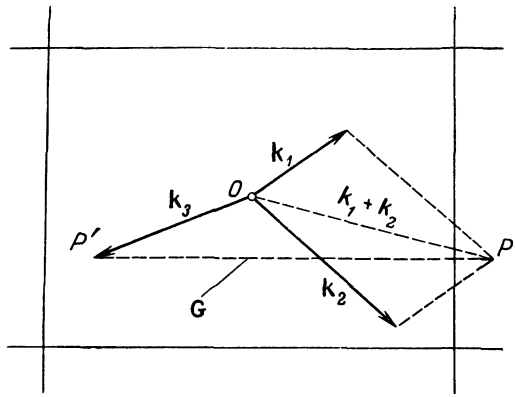


Рис. 2.24. Сечение **k**-пространства, иллюстрирующее соотношение между волновыми векторами фононов в случае *U*-процессов. Для осуществления *U*-процесса векторная сумма $\mathbf{k}_1$ и $\mathbf{k}_2$ должна выходить за границы первой зоны Бриллюэна. *P* и *P'* — эквивалентные точки в соседних зонах Бриллюэна.

Однако Пайерлс [35] показал, что на установление теплового равновесия в системе фононов могут влиять процессы, для которых выполняется закон сохранения энергии (2.77), а волновые векторы связаны соотношением

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{G}. \qquad (2.79)$$

Здесь **G** — вектор обратной решетки. Поскольку фонон с волновым вектором $(\mathbf{k}_3 + \mathbf{G})$ в периодической решетке неотличим от фонона с волновым вектором $\mathbf{k}_3$, выражение (2.79), очевидно, является хорошим законом сохранения. Пайерлс назвал такие процессы *процессами переброса* (Umklapp), или *U*-процессами. Интересной особенностью *U*-процессов является то, что они изменяют импульс и направление потока энергии. Это иллюстрируется на рис. 2.24; в изображенном на нем случае фононы с импульсами $\mathbf{k}_1$ и $\mathbf{k}_2$ переносят энергию вправо, а образовавшийся в процессе столкновения фонон с вектором $\mathbf{k}_3$ — влево. Таким образом, *U*-процессы отвечают за образование

бесконечна, поскольку она отражает только те процессы, которые могут привести в тепловое равновесие неравновесные фононы. Представление о том, что *N*-процессы ограничивают среднюю длину свободного пробега одного фонона, ошибочно. Независимо от того, равновесно или нет фононное распределение, *N*-процессы создают фононы с данной частотой ω_k с той же скоростью, с какой уничтожают их. Только в очень специальном случае (например, в твердом ^{4}He при температурах ниже 1 K, что будет далее обсуждаться в этом разделе более подробно) *N*-процессы могут влиять на *скорость* диффузии фононов.

[35] *Peierls R.* — Ann. Phys., 3, 1055, 1929; см. также: *Peierls R.* Quantum Theory of Solids, Oxf. Univ. Press, 1955. [Имеется перевод: *Пайерлс Р.* Квантовая теория твердых тел.— М.: ИЛ, 1956.]

теплового сопротивления потоку фононов и обеспечивают установление теплового равновесия в фононном распределении.

Пайерлс показал, что средняя длина свободного пробега, определяемая U-процессами, при высоких температурах (выше θ_D) меняется как T^{-1}, поскольку количество возбужденных мод решетки пропорционально T. При рассмотрении возможностей возникновения U-процессов при более низких температурах выясняются интересные обстоятельства. При низких температурах только области k-пространства, близкие к центру зоны Бриллюэна, остаются заселенными фононами. Вместе с тем U-процессы будут происходить только в тех случаях, когда волновой вектор $\mathbf{k}_1+\mathbf{k}_2$ лежит за пределами зоны Бриллюэна. Последнее условие требует, чтобы каждый из волновых векторов $\mathbf{k}_1$ и $\mathbf{k}_2$ был сравним с $\mathbf{G}/2$. Пайерлс предсказал, что при низких температурах вероятность U-процессов будет падать как $\exp(-\theta_u/T)$, где параметр θ_u по величине сравним с $\theta_D/2$. U-процессы сильно отличаются от N-процессов, вероятность которых при температурах значительно ниже температуры Дебая падает[36] по закону T^5.

Температурная зависимость длины свободного пробега фононов

$$\Lambda \sim \exp(\theta_u/T) \tag{2.80}$$

реально наблюдается во многих твердых телах, однако величина θ_u, требуемая для согласования теории с экспериментом, обычно меньше, чем $\theta_D/2$. Так, для сапфира (Al_2O_3) температура Дебая $\theta_D \approx 1000$ К, а *температура процессов переброса* $\theta_u \approx 250$ К. Средняя длина свободного пробега фононов в сапфире при 50 К приблизительно равна 100 мкм, что в 30 раз больше значения Λ, полученного экстраполяцией из области высоких температур зависимости $\Lambda \sim T^{-1}$.

Кроме того, возникают и более сложные процессы фонон-фононного взаимодействия, чем рассмотренные до сих пор. При более подробном изучении вопросов, связанных с теплопроводностью, нужно рассматривать и N-, и U-процессы с участием большего числа фононов. Такие процессы более высоких порядков оказывают существенное влияние на свойства кристалла только при достаточно высоких температурах.

---

[36] Длина волны большинства фононов, возбужденных при температурах значительно ниже θ_D, велика (см. задачу 2.16), и для них эффективно лишь рассеяние на малые углы. Эффективность рассеяния для N-процессов, усредненная по углам, пропорциональна квадрату температуры. Поскольку для того же температурного интервала общее число возбужденных фононов пропорционально T^3 можно ожидать, что $\Lambda_N \sim (T/\theta_D)^{-5}$. Похожие рассуждения, проведенные в гл. 3, показывают, что средняя длина свободного пробега *электронов* при низких температурах, определяемая электрон-фононными взаимодействиями, также пропорциональна T^{-5}.

Рассеяние фононов, обусловленное дефектами

Кроме рассеяния фононов на фононах, они могут рассеиваться также на

1) точечных дефектах (примесях и вакансиях);
2) линейных дефектах (дислокациях);
3) границах зерен в поликристаллах и внешних поверхностях монокристаллов;
4) разупорядочениях в сплавах (как ближнего, так и дальнего порядка);
5) случайном распределении различных изотопов химических элементов.

Рассеяние на всех этих дефектах может привести и к изменению энергии фононов и их квазиимпульса. Так, в кристаллах плохого качества величина Λ остается малой при всех температурах. В очень хороших кристаллах Λ растет при охлаждении (поскольку U-процессы становятся маловероятными) и стремится к насыщению при низких температурах, причем величина Λ при насыщении определяется распределением неизбежно присутствующих в кристалле дефектов.

На рис. 2.23 видно, что при низких температурах устанавливается некоторое постоянное значение Λ, зависящее от распределения дефектов. Поскольку в этом диапазоне C_v изменяется пропорционально T^3 (закон Дебая), очевидно, что теплопроводность $\varkappa_l = (C_v v_0 \Lambda/3)$ также должна быть пропорциональна T^3 для тех температур, для которых U-процессы не эффективны по сравнению с другими механизмами рассеяния.

Эти соображения иллюстрируются красивыми экспериментальными данными, приведенными на рис. 2.25. Здесь представлена теплопроводность монокристаллических образцов фторида лития с весьма совершенной структурой. Для того чтобы свести к минимуму изотопическое рассеяние, содержание изотопа ^{7}Li в них повышено до 99,9 %. Теплопроводность обоих образцов сначала возрастает при понижении температуры в соответствии с законом Пайерлса для U-процессов; затем следует участок кривой, определяемый рассеянием фононов на границах образцов. Чем больше размеры кристалла, тем больше должна быть средняя длина пробега фононов при низких температурах (при условии, что рассеяние на дефектах в объеме образца достаточно мало). Возможность ограничения переноса фононов за счет их рассеяния на границах кристалла была предсказана Казимиром [37]. Механизм этого явления напоминает «кнудсеновское течение» или «баллистический перенос» в кинетической теории газов.

---

[37] *Casimir H. B.*— Physica, 5, 495 (1938).

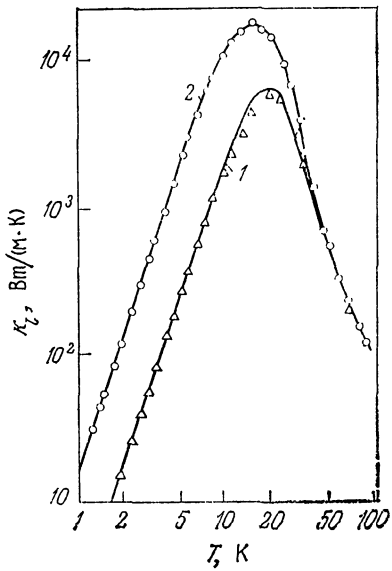


Рис. 2.25. Зависимость решеточной теплопроводности от температуры для монокристаллических стержней из фторида лития. Кривая *1* соответствует образцу с сечением 1,33××0,91 мм, кривая *2* — образцу с сечением 7,55·6,97 мм. При более высоких температурах теплопроводность определяется *U*-процессами, а при более низких — рассеянием на границах кристалла. Чтобы изотопический эффект не мешал наблюдению «размерного эффекта», литий был обогащен изотопом ^{7}Li до 99,9 % [*Berman R.*— Cryogenics, 5, 297 (1965)].

Кривые на рис. 2.25 являются хорошей иллюстрацией того, что при низких температурах действительно выполняется условие Казимира и средняя длина свободного пробега определяется диаметром образца (задача 2.17). Кроме того, уменьшение $\varkappa_l$ при более высоких температурах коррелирует с увеличением роли *U*-процессов. Однако для большинства образцов из многих других материалов максимум на кривых зависимости $\varkappa_l$ от температуры оказывается менее острым из-за присутствия иных конкурирующих механизмов рассеяния фононов. К их числу относится и механизм изотопического рассеяния фононов, которое по существу не зависит от температуры. Из рис. 2.26 видно, что теплопроводность кристалла LiF, который содержит смесь изотопов ^{6}Li и ^{7}Li, растет при охлаждении менее резко, чем в том случае, когда активны лишь процессы переброса. Кроме того, максимум теплопроводности этого образца имеет меньшую высоту и более сглажен. На рис. 2.26 приведены также данные для германия, у которого вклад изотопического рассеяния при комнатной температуре составляет обычно 10 % от всего теплового сопротивления, а при температурах жидкого водорода перекрывает вклад процессов переброса [38]. Высоту низкотемпературного максимума можно повысить, увеличив содержание изотопа ^{74}Ge, однако форма верхней кривой для германия на рис. 2.26 свидетельствует о том, что

[38] *Glassbrenner C. J., Slack G. A.*— Phys. Rev., **134**, A1058 (1964).

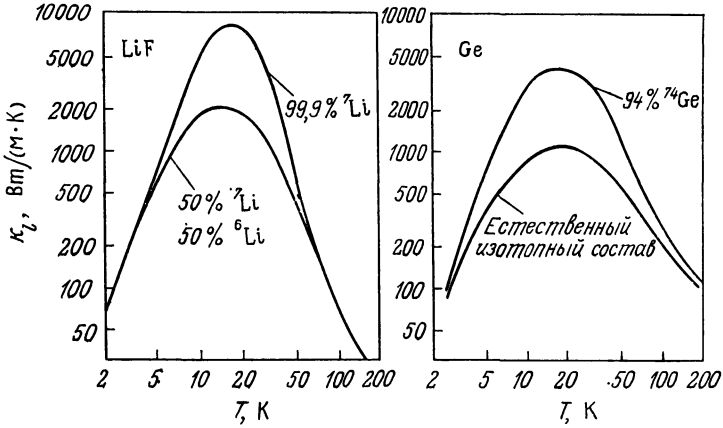


Рис. 2.26. Два примера, иллюстрирующие влияние изотопического рассеяния на величину решеточной теплопроводности. Слева представлена теплопроводность двух кристаллов LiF; один из них содержит смесь изотопов лития в отношении 50 : 50, а другой обогащен и содержит 99,9 % тяжелого изотопа лития [*Berman R.*— Cryogenics, 5, 297 (1965)]. Справа представлены кривые для «нормального» и для частично обогащенного германия [*Geballe T. H., Hull G. W.*— Phys, Rev., **110**, 773 (1958)].

теплопроводность ограничивается не только процессами переброса. Резкое возрастание Λ при низких температурах наблюдается лишь у тех материалов (например, у алмаза или алюминия), в которых отсутствует* изотопическое рассеяние, или у материалов, в которых возможно полное изотопическое обогащение.

В двухкомпонентных сплавах периодичность решетки для фононов нарушается из-за случайного распределения атомов двух сортов по эквивалентным узлам в кристаллической решетке. Поэтому теплопроводность решетки сплава значительно меньше теплопроводности любого из исходных материалов. Рис. 2.27 иллюстрирует это на примере твердых растворов квазибинарной системы GaAs—GaP. Для твердого раствора эквимолярного состава средняя длина свободного пробега фононов при комнатной температуре составляет лишь восемь межатомных расстояний, в то время как для кристалла GaP она порядка 30 межатомных расстояний.

---

* Отсутствие изотопического рассеяния в кристаллах алмаза и алюминия связано с тем, что алюминий состоит лишь из одного стабильного изотопа ^{27}Al, а углерод содержит 98,9 % изотопа ^{12}C. В связи с обсуждением экспериментальных данных по теплопроводности кристаллов LiF заметим, что фтор также состоит лишь из одного стабильного изотопа — ^{19}F.— *Прим. ред.*

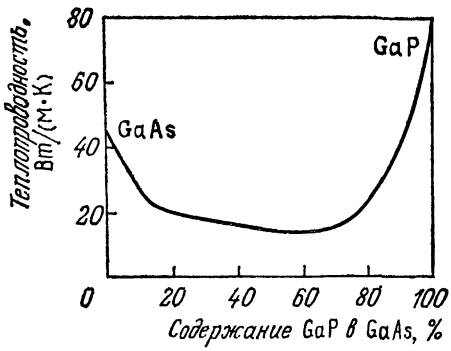


Рис. 2.27. Влияние на решеточную теплопроводность рассеяния фононов, обусловленного случайным распределением компонентов. Приведена зависимость теплопроводности от состава твердых растворов $GaAs_xP_{1-x}$. Экспериментальные результаты получены при $T=300$ К [*Maycock P. D.*— Sol. State Electronics, **10**, 161 (1967)].

Влияние N-процессов на теплопроводность

Как уже отмечалось, Пайерлс[35] пришел к заключению, что N-процессы сами не могут термализовать неравномерное распределение фононов. Тогда вполне логично при обсуждении вопросов, связанных с теплопроводностью, не рассматривать N-процессы и принимать во внимание лишь U-процессы (с участием 3-х или более фононов) и различные типы рассеяния на дефектах. Для большинства материалов это действительно так во всем температурном интервале, и все же *может* возникнуть ряд условий, при которых *наличие* N-процессов меняет ситуацию таким образом, что одно событие становится вероятнее другого. Такая возможность обнаруживается при описании явлений переноса фононов с помощью кинетического уравнения Больцмана[39]. Мы вернемся к уравнению Больцмана в гл. 3 при рассмотрении кинетических свойств электронов.

Подробное исследование метода кинетического уравнения Больцмана показало[40], что N-процессы действительно не играют роли, если среднее расстояние, на которое распространя-

---

[39] Кинетическое уравнение Больцмана используется при рассмотрении явлений переноса в любых статистических системах. Первоначально оно было выведено для описания кинетических свойств молекулярных газов. (Серьезный обзор этих вопросов дан в книге: *Harris S.* An Introduction to the Theory of the Boltzmann Equation. Holt, Rinehart, Winston, 1971.) Уравнение Больмана использовалось для описания кинетических свойств электронного газа еще со времен появления модели Лорентца в 1905 г. (см. разд. 3.2). Для описания потока фононов оно стало использоваться гораздо позже. Фактически впервые такое описание применили Займан [*Ziman J. M.* Canad J. Phys., **34**, 1256 (1956)] и Каллауэй [*Callaway J.*— Phys. Rev. **113**, 1046 (1959)]. Эти вопросы обсуждаются в гл. 8 книги Займана: *Ziman J. M.* Electrons and Phonons, Oxford, 1960. [Имеется перевод: *Займан Дж.* Электроны и фононы.— М.: ИЛ, 1962.]

[40] *Гуржи Р. Н.*— ЖЭТФ, 1964, т. 46. с. 719, см. также: *Guyer R. A., Krumhansl J. A.*— Phys. Rev., **148**, 778, 1966.

ется фонон до U-столкновения, мало по сравнению с размерами образца или кристаллита. При достаточно высоких температурах, когда реализуется такая ситуация, мы можем сказать, что перенос фононов напоминает течение довольно плотной жидкости по трубе, причем N-процессы играют роль упругих столкновений между молекулами, а U-процессы — неупругих столкновений, приводящих к вязкому торможению.

Как мы уже видели, в противоположном предельном случае очень низких температур поток фононов напоминает течение сильно разреженной жидкости. Если в среде нет значительного рассеяния на статических дефектах, то поток фононов аналогичен кнудсеновскому (баллистическому) течению, причем большинство столкновений происходит со стенками. При самых низких температурах диаметр образца D мал по сравнению со средним пробегом фонона между столкновениями с участием U-процессов (Λ_U) или N-процессов (Λ_N), поскольку число фононов с любой энергией вообще очень мало.

Однако при сверхнизких температурах величина Λ_U должна быть значительно больше Λ_N, поскольку в U-процессах должны участвовать фононы *больших энергий*, а N-процессы представляют собой ангармоническое взаимодействие фононов любой энергии. Как было отмечено раньше, при низких температурах вероятность U-процесса меняется как $\exp(-\theta_u/T)$, т. е. более резко, чем вероятность N-процессов, которая пропорциональна T^5. Таким образом, должен существовать ограниченный температурный диапазон, для которого

$$\Lambda_N \ll D \ll \Lambda_U. \tag{2.81}$$

При наличии заметного рассеяния на дефектах в этой температурной области ничего примечательного не происходит. Но если рассеяние на дефектах свести к минимуму, единственными эффектами, *приводящими* к *сопротивлению*, являются столкновения со стенками, хотя гораздо чаще происходят столкновения, не приводящие к сопротивлению (N-процессы). При таких условиях поток фононов аналогичен пуазейлевскому течению жидкости, а перемещение фонона представляет собой случайное блуждание, как и броуновское движение. Исследование этой проблемы как задачи о случайном блуждании (см. задачу 2.18) показывает, что в среднем фонон между последовательными столкновениями со стенками преодолевает путь длиной порядка (D^2/Λ_N). Таким образом, найденная из *экспериментов по измерению теплопереноса* средняя длина пробега оказывается больше диаметра образца и возрастает до тех пор, пока N-процессы не становятся более частыми! Верхний предел средней длины свободного пробега (и теплопроводности) достигается при температурах, достаточно высоких для того,

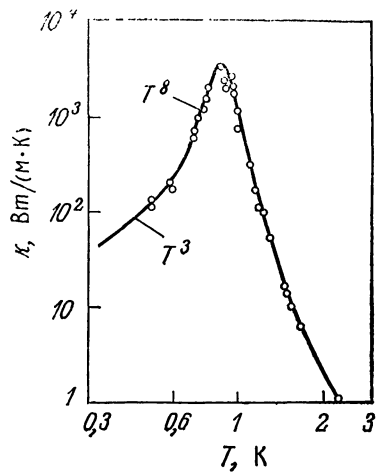


Рис. 2.28. Температурная зависимость теплопроводности кристаллического ^{4}He, имеющего г. п. у.-решетку [*Межов-Деглин* Л. П. — ЖЭТФ, 1965, т. 49, с. 66] (кривая построена в двойном логарифмическом масштабе, причем масштаб по вертикальной оси уменьшен в три раза). Для поддержания ^{4}He в твердом состоянии создавалось давление 85 атм. Максимальному значению теплопроводности соответствует средняя длина свободного пробега фононов, равная 7 мм, диаметр кристалла 2,5 мм.

чтобы U-процессы могли конкурировать с рассеянием на границах. При дальнейшем нагревании проводимость резко падает, а затем выходит на обычную зависимость для U-процессов.

Неравенства (2.81) должны выполняться для всех материалов в ограниченной температурной области. Однако пуазейлевское течение будет наблюдаться только в тех случаях, когда рассеяние на дефектах в том же температурном интервале достаточно мало. Последнее условие не выполняется в LiF, данные для которого представлены на рис. 2.25. Вследствие этого теплопроводность $\varkappa_l$ плавно возрастает с температурой по кубическому закону до тех пор, пока U-процессы не начинают превалировать примерно при 20 К. Пуазейлевское течение может быть в LiF (или другом диэлектрике, который можно изготовить в виде монокристалла), имеющем исключительно высокую изотопическую чистоту, бездефектную кристаллическую структуру и очень низкое содержание примесей.

Легче всего этот набор условий выполняется в твердом гелии, поскольку термодинамические свойства гелия позволяют довольно легко удалить примеси, а существование сверхтекучей фазы ^{4}He дает возможность привести достаточно полное разделение изотопов. На рис. 2.28 приведена экспериментальная зависимость $\varkappa_l(T)$ для кристалла ^{4}He с г. п. у.-структурой, кристаллизованного и сохраняемого в кристаллическом состоянии при давлении 85 атм. Низкотемпературное поведение Λ описывается моделью Казимира для рассеяния на границах, а именно средняя длина свободного пробега имеет порядок диаметра образца, а теплопроводность (как и C_v) меняется

пропорционально T^3. Это область кнудсеновского течения. Однако при нагревании выше 0,6 К теплопроводность $\varkappa_l$ начинает возрастать быстрее, достигая максимума примерно при 1,0 К. В интервале от 0,6 до 1,0 К поведение $\varkappa_l$ соответствует пуазейлевскому течению, причем переносу фононов *способствуют* N-процессы. Нужно заметить, что в этом малом температурном интервале $\varkappa_l \sim T^8$, как и должно быть, поскольку $C_v \sim (T/\theta_D)^3$, а $\Lambda_N \sim (T/\theta_D)^{-5}$. При 1 К полученная «средняя длина свободного пробега» в три раза больше диаметра кристалла, но при дальнейшем нагреве средняя длина свободного пробега (и теплопроводность) падает, поскольку определяющими становятся U-процессы.

Мы кратко обсудили интересную, хотя и не совсем ясную роль N-процессов в теплопроводности. Можно полагать, что в низкотемпературной области возможно различное их проявление. Например, возможность пуазейлевского течения связана с возможностью передачи без существенного затухания теплового импульса — второго звука. Явление второго звука хорошо известно в сверхтекучем жидком гелии; оно наблюдалось также в твердом ^{4}He[41]. Вероятно, второй звук возможен и в других твердых телах, для которых отношения вероятностей различных процессов рассеяния благоприятствуют его наблюдению.

Электронная теплопроводность в металлах и полупроводниках

Тепловая энергия в металлической среде переносится как фононами (которые рассеиваются на других фононах, электронах и дефектах), так и электронами (которые рассеиваются на фононах и дефектах). В чистом металле отношение частоты электрон-фононных столкновений к частоте U-процессов таково, что электронный вклад в общую теплопроводность составляет обычно 90 %. Вероятно, только в металлах с высоким содержанием примесей или в разупорядоченных сплавах средняя длина свободного пробега электрона настолько мала, что при обычных температурах может преобладать фононная теплопроводность.

Мы не будем здесь рассматривать более подробно теорию электронной теплопроводности. Обсуждение переноса электронов в металлах и полупроводниках (гл. 3 и 4) приводит непосредственно к вопросам переноса электронами как энергии, так и заряда. Подробное обсуждение экспериментальных дан-

[41] *Ackerman C. C., Bertman B., Fairbank H S, Guyer R. A.*— Phys. Rev. Lett., 16, 789 (1966).

ных по электронной теплопроводимости в металлах приводится в монографии Розенберга [42].

В полупроводниковых материалах плотность электронного газа, как правило, мала, поэтому электронная теплопроводность также невелика. Однако в некоторых полупроводниках при очень высоких температурах (когда велика плотность термически возбужденных электронов, а величина Λ_U очень мала) ее удается измерить. В собственном полупроводнике значительное количество энергии может переноситься за счет диффузии электронно-дырочных пар. Каждая такая пара, диффундируя как нейтральное образование, переносит энергию, сравнимую с шириной запрещенной зоны полупроводника. Для более глубокого ознакомления с вопросами переноса энергии в полупроводниках можно рекомендовать книгу Драбла и Голдсмида [43].

Задачи

2.1. Рассмотрим распространение продольных волн в одноатомной цепочке. Пусть m — масса каждого атома, a — равновесное межатомное расстояние и μ_j — коэффициент жесткости по отношению к j-му соседнему атому. Вспомним, что соотношение (2.11) было выведено при учете возвращающих сил, действующих только со стороны ближайших соседей, т. е. $\mu_j=0$ при всех $j\neq 1$. Покажите, что в этом случае, когда нельзя пренебречь влиянием атомов, не являющихся ближайшими соседями, соотношение (2.11) заменится дисперсионным соотношением

$$\omega^2 = (2/m)\sum_{j=1}^{\infty}\mu_j[1-\cos(jka)].$$

Постройте кривую $\omega(k)$ при $\mu_2=0{,}25\,\mu_1$ и $\mu_j=0$ для $j>2$. Что изменится, если мы теперь введем третий коэффициент жесткости $\mu_3=-0{,}2\,\mu_1$?

2.2. Рассмотрим трехмерный одноатомный кристалл, имеющий простую кубическую решетку с постоянной решетки a. Покажите, что зависимость плотности состояний от частоты в низкочастотной области колебательного спектра имеет вид

$$g(\omega) = (\omega^2/2\pi^2)\sum_i (m/a^2\mu_i)^{3/2},$$

где суммирование ведется по продольным и поперечным ветвям. Что произойдет, если $\omega_i=\omega_{m_i}\sin(ka/2)$ для каждой ветви во всем спектральном диапазоне? Получите выражение для $g(\omega)$ и постройте график этой зависимости для случая $v_L=2v_T$.

---

[42] *Rosenberg H. M.* Low Temp. Sol. St. Phys., Oxf. Univ. Press, 1963, ch. 5.

[43] *Drabble J. R., Goldsmid H. J.* Thermal Conduction in Semiconductors, Pergamon Press, 1962. [Имеется перевод: *Драбл Дж., Голдсмид Г.* Теплопроводность полупроводников — М.: ИЛ, 1963.]

2.3. Заметим, что оба металла, фононные спектры которых приведены на рис. 2.4, имеют г. ц. к.-структуру. Изобразите в **k**-пространстве расположение точек, соответствующее решетке, обратной по отношению к г. ц. к.-решетке, определите размер и форму первой зоны Бриллюэна. Убедитесь в том, что конечные точки векторов $\mathbf{q}=[100]$, $\mathbf{q}=[^3/_4, ^3/_4, 0]$ и $\mathbf{q}=[^1/_2, ^1/_2, ^1/_2]$ лежат на поверхности, ограничивающей зону Бриллюэна. Укажите несколько других точек **k**-пространства, которые лежат на поверхности, ограничивающей первую зону Бриллюэна и соответствуют направлениям высокой симметрии.

2.4. Рассмотрим изотропный трехмерный кристалл, в котором дисперсионные соотношения для колебаний решетки одинаковы для всех направлений и имеют вид $\omega=(2v_0/a)\sin(ka/2)=\omega_m\sin(ka/2)$. Известно также, что скорости распространения продольных и поперечных волн почти одинаковы. Покажите, что зависимость полной плотности состояний от частоты имеет вид

$$g(\omega)=\frac{6\,[\arcsin(\omega/\omega_m)]^2}{\pi^2a^2v_0\,[1-(\omega/\omega_m)^2]^{1/2}},$$

так что $g(\omega)\sim\omega^2$ в области частот, где отсутствует дисперсия. Постройте кривую $g(\omega)$ во всей области ω, пользуясь выведенной формулой, и объясните, почему выражение не справедливо в верхней части диапазона.

2.5. Нарисуйте плоскость (110) решетки, обратной г. ц. к., и отметьте границы зоны. Покажите, что расстояния от центра зоны Бриллюэна до ее границ в направлениях [111], [100] и [110] относятся как 0,87 : 100 : 1,06, что объясняет выбор масштаба по оси абсцисс на рис. 2.4. Изобразите линии постоянной частоты ω для продольных волн в меди, основываясь на кривых рис. 2.4. Сопоставьте построенные вами кривые с рис. 2.7.

2.6. Рассмотрим линейную двухатомную цепочку, состоящую из атомов с разными массами m_1 и m_2. Размер элементарной ячейки равен $2a$, а расстояние между ближайшими соседями вдвое меньше. Опишите спектр акустических и оптических фононов, если коэффициенты жесткости, определяющие изменение расстояний между ближайшими соседями, равны $\mu_{12}=\mu_{21}$, а дополнительные коэффициенты жесткости, которые определяют изменение расстояний между ближайшими атомами одинаковой массы, равны μ_{11} и μ_{22}. Изобразите графически полученные результаты для $m_1=2m_2$, $\mu_{12}=2\mu_{11}=4\mu_{22}$.

2.7. Получите выражение для групповой скорости продольных волн в линейной двухатомной цепочке и покажите, что она действительно равна нулю для значений угловых частот ω_1, ω_2 и ω_3, указанных на рис. 2.10. Получите зависимость от частоты отношения амплитуд смещений и групповой скорости для обеих ветвей колебательного спектра. Изобразите ваши результаты графически для двухатомной цепочки, у которой $M=3m$ и взаимодействием со всеми соседями, кроме ближайших, можно пренебречь.

2.8. Объясните, почему возникает брэгговское отражение, когда волновой вектор совпадает с границей зоны Бриллюэна. Почему в кристалле не может существовать волна с комплексным или чисто мнимым волновым вектором?

2.9. Электромагнитная волна проходит через кристалл с показателем преломления, равным 5,0 для частот, которые лежат значительно ниже частоты остаточных лучей $\omega_T=10^{14}$ рад/с. Кристалл состоит из однократно заряженных положительных и отрицательных ионов с массой ~ 50 а. е. м. Предположим, что интенсивность волны равна 10^6 Вт/м^2 при угловой частоте $\omega=10^{13}$ рад/с. Рассчитайте среднеквадратичное значение амплитуды электрического и магнитного вектора этой электромагнитной волны. Запишите урав-

нение движения для ионов, на которые действует переменное электрическое поле волны, и определите среднеквадратичную амплитуду их движения. Используя выражение для силы Лоренца, определите влияние осциллирующего магнитного поля электромагнитной волны на движение ионов.

2.10. Фотоны видимой области спектра с энергией 2 эВ проходят сквозь кристалл с показателем преломления, равным 2,5, а скорость звука в кристалле составляет 5000 м/с. Определите энергию фононов, образующихся при бриллюэновском рассеянии первого порядка, как функцию угла между падающим на кристалл светом и направлением импульса образовавшихся фононов. Постройте график зависимости частоты звука и изменения длины волны света от угла.

2.11. Покажите, что энергия $\langle E\rangle$ в выражении (2.47) есть сумма двух слагаемых, каждое из которых является отношением интегралов и равно $k_0T/2$, причем первое слагаемое соответствует средней кинетической энергии классического гармонического осциллятора, а второе — средней потенциальной энергии.

2.12. Получите выражение (2.53) из выражений (2.50) и (2.52). *Указание:* представьте знаменатель выражения (2.52) в виде бесконечной геометрической прогрессии и убедитесь, что выражение (2.52) можно получить, беря производную от натурального логарифма этой прогрессии. Найдите сумму прогрессии и получите искомый результат.

2.13. Рассмотрите в рамках модели Дебая энергию и теплоемкость одноатомной линейной цепочки с межатомным расстоянием a, скорость звука в которой равна v_0 [см. выражение (2.13)]. Покажите, что для дебаевской частоты $\omega_D=(\pi v_0/a)$ и дебаевской температуры $\theta_D=(\hbar\omega_D/k_0)$ сохраняется правильное полное число состояний. Получите интегральные соотношения для энергии решетки и теплоемкости и покажите, что при низких температурах теплоемкость на единицу длины $C_v=(\pi^2k_0T/3a\theta_D)$, а при высоких она равна классическому значению k_0/a.

2.14. В рамках модели Эйнштейна, согласно которой все моды вырождены и имеют частоту ω_E, получите выражение для температуры, при которой тепловая энергия равна энергии нулевых колебаний. Затем запишите условие равенства тепловой энергии и энергии нулевых колебаний в модели Дебая для кристалла, скорость звука в котором равна v_0.

2.15. Пусть в воображаемом твердом теле имеется N квантовых осцилляторов с частотой ω_1, $2N$ осцилляторов с частотой $2\omega_1$ и $3N$ осцилляторов с частотой $3\omega_1$. Покажите, как энергия колебаний и теплоемкость будут меняться с температурой (энергией нулевых колебаний пренебрегаем). Сравните эту зависимость с протабулированной функцией Дебая, чтобы установить, как в нашей модели температура Дебая θ_D зависит от температуры.

2.16. Рассмотрим трехмерный кристалл, плотность состояний в котором соответствует модели Дебая. Покажите, что при температурах ниже θ_D наиболее вероятная энергия фонона равна $\hbar\omega_p$, причем

$$\exp(\hbar\omega_p/k_0T)\,[1-(\hbar\omega_p/2k_0T)]=1.$$

Отсюда следует, что наиболее вероятная длина волны фонона $\lambda_p=(a\theta_D/T)$.

2.17. Покажите, что для температур значительно ниже температуры Дебая среднюю длину свободного пробега фонона можно записать следующим образом:

$$\Lambda=A\varkappa_l\theta_D^3/v_0T^3=B\varkappa_lv_0^2/T^3.$$

Рассчитайте Λ для кристалла LiF, который содержит только изотоп ^{7}Li с молекулярной массой 27. Фторид лития имеет структуру NaCl, а расстояние между ближайшими соседями равно 2,014 Å. Покажите, что если для самых низких температур на рис. 2.25 величину Λ считать равной диаметру образца, то восходящим участкам кривых $\Lambda(T)$ соответствуют значения $\theta_D \approx 720$ К или $v_0 \approx 5000$ м/с.

2.18. Рассмотрим броуновское движение фононов, которые испытывают лишь столкновения друг с другом типа N-процессов и неупругие столкновения со стенками кристалла. Покажите, что если выполняются неравенства (2.81), то расстояние, которое проходит фонон между двумя последовательными столкновениями со стенками кристалла диаметром D, имеет порядок D^2/Λ_N.

Литература

Упругие свойства твердых тел

Ultrasonics: Methods of Experimental Physics, ed. P. D. Edmonds, v. 19, Academic Press, 1981.

Ghatak A. K., Kothari L. S. An Introduction to Lattice Dynamics, Addison-Wesley, 1972.

Huntingdon H. B. Elastic Constants of Criystals, Academic Press, 1964.

Physical Acoustics, eds. W. P. Mason, R. N. Thurston, v. 1, 1964 — v. 17, 1984.— Academic Press. [Имеется перевод: Физическая акустика/Под ред. У. Мэзона.— М.: Мир, 1966, т. 1; 1974, т. 7.]

Musgrave M. J. P. Crystal Acoustics, Holden-Day, 1970.

Narasimhamurty T. S. Photoelastic and Electro-Optic Properties of Crystals, Academic Press, 1981. [Имеется перевод: *Нарасимхамурти Т. С.* Фотоупругие и электрооптические свойства кристаллов.— М.: Мир, 1984.]

* *Френкель Я. И.* Введение в теорию металлов.— Л.: Наука, 1972.

Динамика решетки

Bilz H., Kress W. Phonon Dispersion Relations in Insulators, Springer, 1979.

Born M., Huang K. Dynamical Theory of Crystal Lattices, Oxford Univ. Press, 1954. [Имеется перевод: *Борн М., Хуан Кунь.* Динамическая теория кристаллических решеток.— М.: ИЛ, 1958.]

Cochran W. Dynamics of Atoms in Crystals, Arnold, 1973.

Donovan B., Angress J. F. Lattice Vibrations, Chapman and Hall, 1971.

Dynamical Properties of Solids, eds. G. K. Horton, A. A. Maradudin, v. 1, 1974 — v. 5, 1984, North-Holland.

Maradudin A. A., Montroll E. W., Weiss G. H., Ipatova I. P. Theory of Lattice Dunamics in the Harmonic Approximation, Academic Press 2nd ed., 1971. [Имеется перевод 1-го издания: *Марадудин А. А., Монтролл Е. В., Вейсс Дж.* Динамическая теория кристаллической решетки в гармоническом приближении — М.: Мир, 1965.]

Venkataraman G., Feldkamp L. A., Sahni V. C. Dynamics of Perfect Crystals, MIT Press, 1975.

Willis B. T. M., Pryor A. W. Thermal Vibrations in Crystallography, Cambridge Univ. Press, 1975.

* *Киттель У.* Введение в физику твердого тела.— М.: Наука, 1978.

* *Лейбфрид Г.* Микроскопическая теория механических и тепловых свойств кристаллов.— М — Л.: Физматгиз, 1963.

Фононы и взаимодействия фононов

Bartolo B. Di., Powell R. C. Phonons and Resonances in Solids, Wiley, 1976.
Lovesey S. W. Theory of Neutron Scattering from Condensed Matter, Oxford Univ. Press, 1984. 2 v.
Phonon Scattering in Condensed Matter, ed. H. J. Maris, Plenum Press, 1980.
Reissland J. A. Physics of Phonons, Wiley, 1973. [Имеется перевод: *Рейсленд Дж.* Физика фононов.— М.: Мир, 1975.]
Ziman J. M. Electrons and Phonons, Oxford Univ. Press, 1960. [Имеется перевод: *Займан Дж.* Электроны и фононы.— М.: ИЛ, 1962.]
* *Заембо Л. К., Красильников В. А.* Введение в нелинейную акустику.— М.: Наука, 1966.
* *Гуревич И. И., Тарасов В. В.* Физика нейтронов высоких энергий.— М.: Наука, 1965.
* Динамические свойства твердых тел и жидкостей. Исследование методом рассеяния нейтронов/Под ред. С. Лавш, Т. Шприпгера.— М.: Мир, 1980.

Тепловые свойства

Berman R. Thermal Conduction in Solids, Oxford Univ. Press. 1976. [Имеется перевод: *Берман Р.* Теплопроводность твердых тел.— М.: Мир, 1979.]
Blackman M.— In: Handbuch der Physik, 7, 325, Springer, 1955. (О теплоемкости).
Mechanical and Thermal Behavior of Metallic Materials, eds. G. Caglioti, A. F. Milone, North-Holland, 1982.
Drabble J. R., Goldsmid H. J. Thermal Conduction in Semiconductors, Pergamon Press, 1962. [Имеется перевод: *Драбл Дж., Голдсмит Г.* Теплопроводность полупроводников.— М.: ИЛ. 1963.]
Gopal E. S. R. Specific Heats at Low Temperatures, Plenum Press, 1966.
Ho C. Y., Powell R. W., Liley P. E. Thermal Conductivity of Elements, NBS, 1975.
Keesom P. H., Pearlman N.— In: Handbuch der Physik, 14, 282, Springer, 1956. (О теплоемкости при низких температурах).
Klements P. G.— In: Handbuch der Physik., 14, 198, Springer, 1956. (О теплопроводности).
Parrott I. E., Stuckes A. D. Thermal Conductivity of Solids, Pion, 1975.
Yates B. Thermal Expansion, Plenum Press, 1972.
* *Карслоу Г., Егер Д.* Теплопроводность твердых тел.— М.: Наука, 1964.
* *Новикова С. И.* Тепловое расширение твердых тел.— М.: Наука, 1974.

3

Электроны в металлах

Хотя металлическая проводимость привлекала интерес исследователей в течение долгого времени, ясное понимание динамики электронов в металлах как это ни странно, было достигнуто лишь недавно. В данной главе мы будем соблюдать историческую последовательность, описывая в первую очередь простые модели свободных электронов. Эти модели предполагают, что внешние электроны в металлах могут освобождаться от своих атомов и по характеру своего беспорядочного теплового движения и вкладу в проводимость образуют «электронный газ». Такие модели объяснили некоторые важные свойства металлов, однако вызвали ряд новых вопросов, остававшихся без ответа, пока не стали учитывать, что движение электронного газа происходит в пространстве, занятом периодически распределенными положительно заряженными атомными остатками (ионами).

В двух предыдущих главах мы часто подчеркивали периодический характер кристаллической решетки. В данной главе мы увидим, что периодичность расположения ионов создает такое распределение электростатического поля, которое радикальным образом влияет на связь между энергией подвижного электрона и его импульсом. Учет этого поля приводит к теории, называемой «зонной теорией твердых тел». Она не только дает более реалистическую картину металлической проводимости, но и объясняет, почему многие твердые тела являются диэлектриками или полупроводниками.

Любопытно, что появление этой теории привело в 40—50-х годах к необычайно быстрому прогрессу в понимании свойств полупроводников и диэлектриков. Можно утверждать, что к середине 50-х годов теория некоторых простых полупроводников была гораздо более полной, чем теория металла. Однако с этого момента маятник начал обратное движение, и с развитием тонких экспериментальных методов вновь возродился интерес к металлам. Одним из проявлений этого интереса стали исследования зависимости величины импульса от его направ-

ления для электрона с постоянной энергией, т. е. изучение топологии «поверхности Ферми» в k-пространстве, позволившее внести много нового в понимание процессов проводимости в металлах.

В металле свободный электрон взаимодействует с периодически расположенными ионами, с дефектами кристалла и примесями, с тепловыми смещениями системы ионов (т. е. фононами) и с другими электронами. В данной главе мы уделим много внимания тому, как идеальная периодичность ионов позволяет электрону сохранять прямолинейное движение и как дефекты и фононы вызывают изменение этого движения. Будет дано краткое описание коллективных эффектов в металлах (включая «плазменные колебания» и сверхпроводимость), однако в основном стандартная теория металлов основана на предположении, что взаимной корреляцией в движении отдельных электронов можно пренебречь. Таким образом, стандартная теория металлов — это теория одноэлектронная, в которой отдельные электроны не взаимодействуют друг с другом и подчиняются статистике Ферми — Дирака. Простота этой статистики утрачивается, как только приходится иметь дело с ансамблями взаимодействующих частиц.

Обсуждение «теорий свободных электронов» (как классических, так и квантовых) занимает в данной главе достаточно много места. Наша основная цель при изучении моделей свободных электронов не является чисто исторической: эти модели заслуживают внимания потому, что они содержат все основные представления об электропроводности, принятые с незначительными изменениями в зонной теории твердых тел. Конечный результат зонной теории обычно обосновывает справедливость приближения «почти свободных электронов», но с перенормировкой их масс и заменой некоторых скаляров тензорами.

3.1. Некоторые особенности металлического состояния

Типичные свойства металлов

Прежде чем обсуждать различные теоретические модели металлического состояния, полезно вспомнить наиболее важные свойства металла. Если теория металлов удовлетворительна, эти свойства должны быть естественным ее следствием. Имея это в виду, следует признать, что в настоящее время еще не существует удовлетворительной во всех отношениях теории металлов. Последовательно развиваемые теоретические модели

позволяют только *приблизиться* к количественному и качественному описанию известных свойств металлов.

Внимания заслуживают следующие свойства металлов.

1. В изотермических условиях в металле хорошо выполняется закон Ома. Этот закон часто записывают в виде

$$\mathbf{J} = \sigma \mathbf{E}, \tag{3.1}$$

где скалярная электропроводность σ [$Ом^{-1} \cdot м^{-1}$] связывает плотность тока $\mathbf{J}$ [$А/м^2$] с напряженностью электрического поля $\mathbf{E}$ [В/м].

2. Металл — очень хороший проводник электричества. Если электропроводность изоляторов может быть настолько мала, что составляет 10^{-16} $Ом^{-1} \cdot м^{-1}$ (см. рис. 3.21), а электропроводность полупроводника обычно лежит в пределах от 10^{-4} до 10^5 $Ом^{-1} \cdot м^{-1}$, то большинство металлов при комнатной температуре обладает электропроводностью 10^6—10^8 $Ом^{-1} \cdot м^{-1}$. В табл. 3.1 приведены значения удельной электропроводности ряда металлов, начиная от одновалентных и кончая многовалентными и переходными металлами.

3. Металл обладает также высокой электронной теплопроводностью $\varkappa_e$. Видеман и Франц в 1853 г. заметили, что хороший проводник тепла является одновременно и хорошим проводником электричества. Совпадение отношения $(\varkappa_e/\sigma)$ для разных металлов при данной температуре с тех пор называется *законом Видемана — Франца*. При обычных температурах $\varkappa_e$, как правило, от температуры почти не зависит, в то время как

Таблица 3.1. **Электропроводность и число Лоренца для некоторых металлических элементов***

Металл	$T = 100$ K		$T = 273$ K	
	Электропроводность σ, $Ом^{-1} \cdot м^{-1}$	Число Лоренца $L = \varkappa_e/\sigma T$, $(В/К)^2$	Электропроводность σ, $Ом^{-1} \cdot м^{-1}$	Число Лоренца $L = \varkappa_e/\sigma T$, $(В/К)^2$
Медь	$2{,}9 \cdot 10^8$	$1{,}9 \cdot 10^{-8}$	$6{,}5 \cdot 10^7$	$2{,}3 \cdot 10^{-8}$
Золото	$1{,}6 \cdot 10^8$	$2{,}0 \cdot 10^{-8}$	$5{,}0 \cdot 10^7$	$2{,}4 \cdot 10^{-8}$
Цинк	$6{,}2 \cdot 10^7$	$1{,}8 \cdot 10^{-8}$	$1{,}8 \cdot 10^7$	$2{,}3 \cdot 10^{-8}$
Кадмий	$4{,}3 \cdot 10^7$	$2{,}1 \cdot 10^{-8}$	$1{,}5 \cdot 10^7$	$2{,}4 \cdot 10^{-8}$
Алюминий	$2{,}1 \cdot 10^8$	$1{,}5 \cdot 10^{-8}$	$4{,}0 \cdot 10^7$	$2{,}2 \cdot 10^{-8}$
Свинец	$1{,}5 \cdot 10^7$	$2{,}0 \cdot 10^{-8}$	$5{,}2 \cdot 10^6$	$2{,}5 \cdot 10^{-8}$
Вольфрам	$9{,}8 \cdot 10^7$	$2{,}8 \cdot 10^{-8}$	$2{,}1 \cdot 10^7$	$3{,}0 \cdot 10^{-8}$
Железо	$8{,}0 \cdot 10^7$	$3{,}1 \cdot 10^{-8}$	$1{,}1 \cdot 10^7$	$2{,}8 \cdot 10^{-8}$

* Значения σ взяты из работы: *Meaden G. T.* Electrical Resistance of Metals, Plenum Press, 1965. Данные по теплопроводности, позволившие вычислить число Лоренца — из работы: American Institute of Physics Handbook, McGraw-Hill, 3zd ed, 1971.

σ обнаруживает температурную зависимость вида T^{-1}, показанную для меди на рис. 3.1. В 1881 г. Лоренц заметил, что отношение $(\varkappa_e/\sigma T)$ не зависит от температуры и имеет одинаковую величину для многих металлов, поэтому величину $L \equiv (\varkappa_e/\sigma T)$ называют *числом Лоренца*. Значения L для некоторых металлов приведены в табл. 3.1.

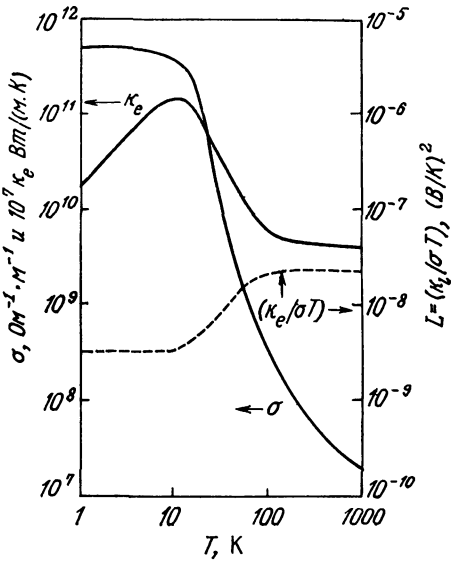


Рис. 3.1. Температурная зависимость электропроводности σ и электронной теплопроводности $\varkappa_e$ для меди высокой чистоты.

4. Если металл охлажден ниже некоторой характерной температуры (которая связана с характеристической температурой θ_D для теплоемкости), то наблюдается рост $\varkappa_e$ и еще бóльшее возрастание σ. Для меди (рис. 3.1) и некоторых других одновалентных металлов на самом крутом участке кривой электропроводность пропорциональна T^{-5}.

5. При достаточно низкой температуре электропроводность σ достигает насыщения; ее значение при этом определяется примесями и дефектами решетки. *Удельное электрическое сопротивление* (величина, обратная электропроводности, измеряемая в Ом·м) часто подчиняется *правилу Маттисена*, согласно которому вклад, вносимый примесями и дефектами, одинаков для всех температур:

$$1/\sigma(T) = 1/\sigma_{\text{деф}} + 1/\sigma_{\text{чист}}(T). \qquad (3.2)$$

Ясно, что вклад $\sigma_{\text{деф}}$ можно наблюдать только при очень низких температурах, когда $\sigma_{\text{чист}} \to \infty$.

6. Магнитные эффекты в ферромагнитных металлах и сплавах также дают вклад в электрическое удельное сопротивление.

7. Примерно половина металлических элементов при достаточно низких температурах становится сверхпроводниками.

8. Газ свободных электронов обладает очень малой удельной теплоемкостью, которая пропорциональна абсолютной тем-

пературе, а также очень малой магнитной восприимчивостью, которая не зависит от температуры.

9. При наличии комбинации электрического, магнитного и температурного градиентов возникают многочисленные термогальвано-магнитные эффекты (например, температурный градиент может вызвать электрический ток или разность потенциалов). При этом связь между данным воздействием и возникающим явлением должна описываться некоторым тензором. Многие из вторичных эффектов в металлах крайне малы, и их трудно измерить с достаточной точностью.

10. В очень чистых монокристаллах в сильных магнитных полях проявляются эффекты, зависящие от ориентации; их величина обнаруживает осциллирующую зависимость от напряженности магнитного поля.

Свободные электроны и положительные атомные остатки (ионы)

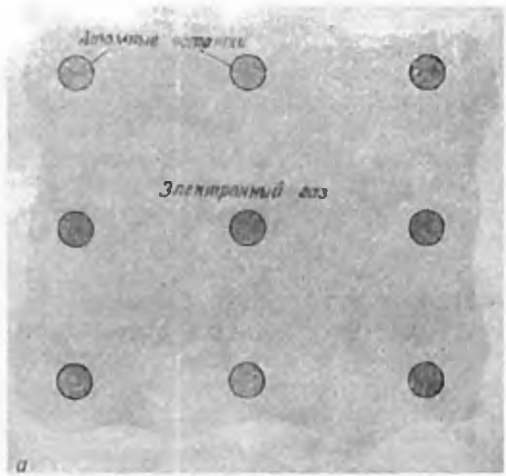
В разд. 1.1 были рассмотрены различные виды межатомной связи в твердых телах. Было указано, что в ионном и ковалентном кристалле валентные электроны связаны с определенными атомами, а в металле межатомная связь является результатом коллективизации электронов. В типичной металлической структуре каждый атом имеет много ближайших соседей и много связей. Каждая из связей очень слабая, но металл обладает заметной прочностью, потому что число связей велико.

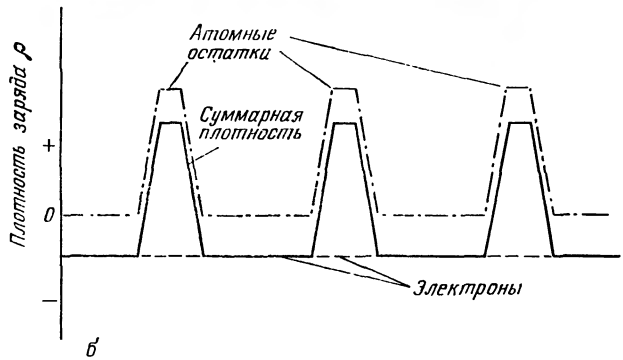
Согласно теории металлов, основные черты металлического кристалла представляются в следующем виде:

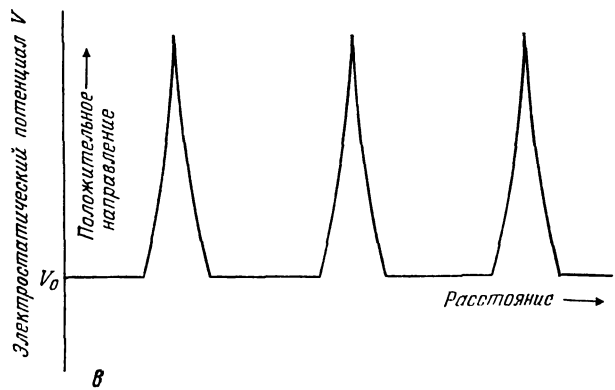
1. Периодическое расположение атомных остатков, или ионов. Каждый ион состоит из атомного ядра и электронов, образующих заполненные оболочки. Согласно принципу Паули электрон, находящийся в заполненной оболочке данного атома, должен быть связан исключительно с этим атомом.

2. Квазиоднородная плотность отрицательного заряда обусловлена всеми внешними электронами. Они образуют «элект-

Рис. 3.2. *а* — двумерный срез металлического кристалла, показывающий периодическое расположение компактных атомных остовов. «Газ» свободных электронов считается равномерно распределенным по всему пространству. *б* — зависимость плотности заряда от расстояния вдоль направления, пересекающего ряд положительных атомных остатков. Вдоль такого направления ρ периодически меняет знак с положительного на отрицательный. Если бы мы исследовали ρ вдоль параллельного направления, не пересекающего атомные остатки, то обнаружили бы, что ρ — однородно отрицательная величина вдоль всей прямой. *в* — периодическое изменение электростатического потенциала, обусловленное распределением плотности заряда, изображенным в части *б* рисунка.


Электронный газ
а


Плотность заряда ρ
Атомные остатки
Суммарная плотность
+
0
−
Электроны
б


Электростатический потенциал V
Положительное направление
V_0
Расстояние
в

ронный газ», частицы которого движутся с тепловыми скоростями. Коллективное действие этих электронов обеспечивает связь в твердом теле. Плотность свободных электронов может составлять от одного электрона на атом ($2 \cdot 10^{28}$ м$^{-3}$) и выше в зависимости от валентности атомов, образующих металл.

Схематически такая совокупность атомных остатков и электронов изображена на рис. 3.2, *а*, а рис. 3.2, *б* и *в* показывают, к каким выводам приводит такая схема. *Средняя* плотность заряда для комбинации, изображенной на рис. 3.2, *а*, должна быть равна нулю в силу электрической нейтральности твердого тела. Однако, рассматривая *локальное* распределение, можно предположить, что плотность заряда ρ должна иметь периодические максимумы, накладывающиеся на однородное распределение отрицательного заряда электронного газа. Такие осцилляции ρ показаны на рис. 3.2, *б*. Поскольку данный атомный остаток, состоящий из заполненных оболочек, обычно значительно меньше межатомного расстояния, ионы, как правило, занимают только малую часть всего объема [1] и максимумы плотности ρ велики, но сосредоточены в малом объеме.

Из уравнения Пуассона

$$\nabla^2 V = -\rho/\varepsilon_0 \tag{3.3}$$

следует, что в металлическом кристалле должно существовать трехмерное периодическое распределение электростатического потенциала, как показано на рис. 3.2, *в*. Форма и высота каждого максимума потенциала зависят от особенностей распределения заряда в атомном остатке.

В модели *свободных электронов* считается, что влиянием периодичности электростатического потенциала можно *пренебречь*. Расчеты делаются в предположении, что электроны движутся совершенно *свободно* по всему объему металла, наталкиваясь на потенциальный барьер у поверхности твердого тела. Этот внешний барьер (рис. 3.3) представляет собой работу выхода ϕ. Он преодолевается электроном в таких процессах, как термо- и фотоэлектронная эмиссия.

Плотность свободных электронов может быть порядка $2 \cdot 10^{28}$ м$^{-3}$ и выше в зависимости от места атомов металла в периодической таблице. Статистическое рассмотрение этих электронов при полном пренебрежении межатомным потенциалом наиболее успешно применимо (в квантовой форме) к одновалентным щелочным металлам, хотя в первом приближении они дают правильное представление и о поведении многовалентных металлов, таких, как алюминий и свинец.

[1] См. задачу 3.1 об определении объема, занимаемого свободными электронами в простейших металлах — щелочных металлах группы 1 А.

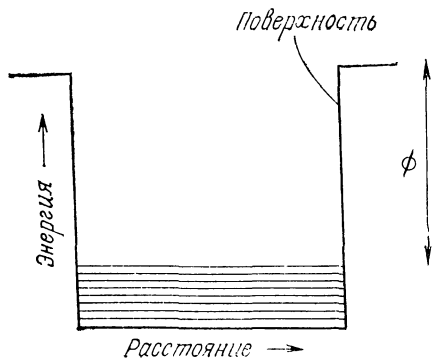


Рис. 3.3. Энергия электрона в твердом теле определяется суммой потенциальной и кинетической энергий, причем вклад последней достаточно велик, только если электрон быстро движется. Даже самый высокоэнергичный, электрон в твердом теле обладает гораздо меньшей *полной* энергией, чем покоящийся электрон в вакууме снаружи металла, поскольку его энергия велика и отрицательна. Электрон, обладающий отрицательным зарядом, может выйти из кристалла в окружающий вакуум, только если он преодолеет поверхностный потенциальный барьер. Высота этого барьера называется *работой выхода* Φ: ее величина обычно составляет несколько эВ на электрон. Способствовать преодолению электронами этого внешнего барьера может тепловое движение (термоэлектронная эмиссия) или поглощение фотона с достаточно большой энергией (внешний фотоэффект или фотоэмиссия).

Хаотическое и упорядоченное движение электронов

Общим для всех теоретических моделей металлической проводимости является то, что движение электронов можно рассматривать как сумму двух составляющих. Первая составляющая — это хаотическое движение, обусловленное тепловой энергией и энергией при $T=0$. (В классических моделях учитывается только тепловое движение. В квантовых моделях «вырожденного» электронного газа преобладающим является движение в нулевой точке). Мгновенная скорость любого электрона довольно велика, однако название «хаотическое» движение подразумевает, что в термодинамическом равновесии в отсутствие каких бы то ни было внешних воздействий средняя по ансамблю скорость равна нулю [2].

[2] Для любого конечного ансамбля электронов средняя дрейфовая скорость $\mathbf{v}_D = \frac{1}{N}\sum_{i=1}^{N} \mathbf{v}_i$, взятая как функция времени, колеблется в термодинамическом равновесии около нуля. Вследствие этого флуктуации напряжения в разомкнутой цепи могут наблюдаться между любыми двумя точками проводника даже в отсутствие макроскопических токов. Этот *тепловой шум* имеет составляющие, обладающие при температуре T одной и той же

Для любого данного электрона *скалярная скорость*, усредненная по достаточно длительному промежутку времени (в течение которого происходит много соударений), отлична от нуля, однако *векторная скорость*, усредненная по большому периоду, в отсутствие какого-либо внешнего поля равна нулю. Это объясняется тем, что соударения, изменяющие направление движения электронов, делают все направления движения равновероятными.

Вторая составляющая движения — упорядоченное дрейфовое движение, обусловленное градиентом электрического поля или температуры в металле. Когда существуют такие градиенты, средняя скорость ансамбля не обращается в нуль и все электроны стремятся перемещаться в одном и том же направлении. Возможность наблюдения такого эффекта зависит от *среднего времени свободного пробега* между соударениями.

3.2. Классическая теория свободных электронов

Распределение скоростей Максвелла—Больцмана

Мы теперь знаем, что на распределение скоростей электронов накладываются квантовые ограничения. Однако некоторые аспекты поведения металлов были объяснены с помощью классических моделей, которые предшествовали появлению статистики Ферми — Дирака. Это были модели, основанные на классическом распределении скоростей Максвелла — Больцмана. Предполагается, что читатель знаком с основами статистической физики и ему известно, что, согласно распределению Больцмана, число электронов в единице объема, скорости которых в *термодинамическом равновесии* лежат в интервале $dv_x dv_y dv_z$ с центром в точке $v_x v_y v_z$, составляет

$$f_0\, dv_x\, dv_y\, dv_z = \\ = n\left(\frac{m}{2\pi k_0 T}\right)^{3/2} \exp\left[\frac{-m\left(v_x^2+v_y^2+v_z^2\right)}{2k_0 T}\right] dv_x\, dv_y\, dv_z, \qquad (3.4)$$

где f_0 — равновесная *функция распределения*, а n — число электронов в единице объема.

амплитудой для всех частот от нуля до (k_0T/h). Тепловой шум, впервые обсуждавшийся Джонсоном [*Johnson J. B.*— Phys. Rev. 32, 97 (1928)] и Найквистом [*Nyquist H.*— Phys. Rev., 32, 110 (1928)], описан в связи с другими флуктуационными явлениями в книгах Киттеля (*Kittel C.* Elementary Statistical Physics, Wiley, 1958) [Имеется перевод: *Киттель С.* Элементарная статистическая физика.— М: ИЛ, 1958] и Мак Дональда (*Mac Donald D. K. C.* Noise and Fluctuations: an Introduction, Wiley, 1962).

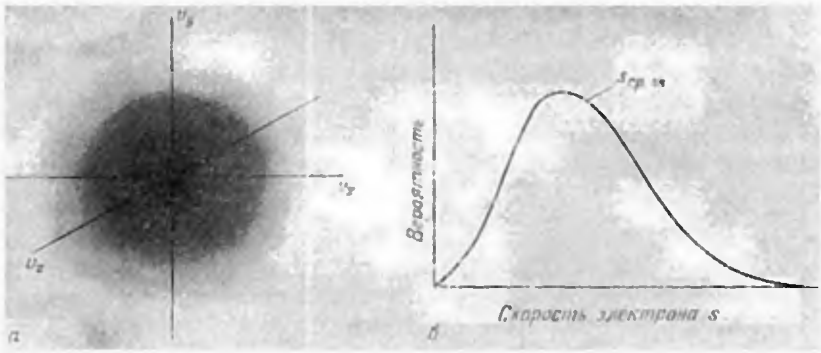


Рис. 3.4. Распределение Максвелла — Больцмана [формула (3.4)]. *а* — распределение векторных скоростей в пространстве скоростей; *б* — равновесная кривая вероятности для скалярной скорости s без учета направлений этой скорости в соответствии с формулой (3.5). Наиболее вероятная скорость равна $(2k_0T/m)^{1/2}$, однако скорость, соответствующая средней энергии, есть $s_{\text{ср.кв}} = (3k_0T/m)^{1/2}$.

Равновесное распределение Больцмана сферически-симметрично в пространстве скоростей, как показано на рис. 3.4, *а*. Таким образом, вероятность того, что скалярная скорость находится в интервале от s до $(s+ds)$, равна

$$\frac{dn}{n} = 4\pi s^2 \left(\frac{m}{2\pi k_0 T}\right)^{3/2} \exp\left[\frac{-ms^2}{2k_0T}\right] ds \qquad (3.5)$$

независимо от направления этой скорости (рис. 3.4, *б*). Средняя кинетическая энергия, приходящаяся на электрон, для распределения Больцмана составляет

$$\frac{3}{2} k_0T = \frac{1}{2} m(s_{\text{ср. кв}})^2 \qquad (3.6)$$

для средней квадратичной скорости (корня из среднего квадрата скорости) $s_{\text{ср. кв}} = (3k_0T/m)^{1/2}$. При обычных температурах эта скорость имеет порядок 10^5 м/с.

Упругое рассеяние и средняя длина свободного пробега

При построении классической теории проводимости свободных электронов кажется совершенно естественным предположение о том, что средняя длина свободного пробега электрона определяется в первую очередь упругими столкновениями

с положительными атомными остатками. В металле, содержащем в единице объема N атомных остатков, каждый из которых обладает эффективным радиусом рассеяния R, средняя длина свободного пробега электронов между двумя рассеивающими соображениям (задача 3.2), должна иметь порядок

$$\lambda \sim (\pi R^2 N)^{-1} \tag{3.7}$$

независимо от скорости рассеиваемого электрона. Такие столкновения могут дать полный спектр углов рассеяния при очень малых изменениях энергии электрона из-за гораздо большей массы иона. По-видимому, средняя длина свободного пробега, определенная выражением (3.7), является той величиной, которую следует подставить в соотношение

$$\tau_m = \frac{\lambda}{s_T}, \tag{3.8}$$

где s_T обозначает тепловую скорость, с которой электрон движется от одного рассеивающего центра к другому, а τ_m — среднее время свободного пробега между столкновениями. Любое внешнее поле должно влиять на изменение скорости электрона в течение коротких интервалов времени, сравнимых с τ_m.

Обсуждение дрейфа электронов, обусловленного полем, производится без учета теплового движения (за исключением введения величины τ_m). Следует всегда помнить, что дрейфовые скорости *малы*, но имеют строго определенное направление. Дрейф накладывается на тепловое движение, которое характеризуется гораздо большими скоростями, обращающимися в нуль в среднем по ансамблю.

Модель Друде

Первая простейшая классическая модель газа свободных электронов в металле была построена Друде в 1900 г. Ее изложение здесь полезно потому, что, несмотря на ограниченность этой модели, она содержит некоторые идеи, которые до сих пор остаются неотъемлемой частью более серьезных теорий.

Друде предположил, что при полной плотности электронов n в единице объема металла каждый его атом должен отдать «электронному газу» не менее одного электрона. Он выяснил, к каким результатам приводит чрезвычайно упрощенная модель, основанная на предположении, что *каждый* электрон обладает кинетической энергией, соответствующей трем классическим степеням свободы поступательного движения, т. е. что каждый электрон обладает кинетической энергией $(3/2)k_0T$. Таким образом, в модели Друде все электроны движутся со

средней квадратичной скоростью, определяемой распределением Больцмана (3.6).

Из этого предположения следует, что электронный газ должен обладать полной энергией $(3/2)nk_0T$ и удельной электронной теплоемкостью $(3/2)nk_0$. Следовательно, в таком металле (например, в натрии или меди), у которого плотность атомов и плотность электронов одинаковы, электроны должны примерно на 50 % увеличивать удельную теплоемкость по сравнению с непроводником. Это предсказание теории для реальных металлов не выполняется. Удельная теплоемкость электронов в металле на самом деле очень мала и линейно зависит от температуры, а не является константой. Таким образом, с самого начала мы можем указать на серьезный недостаток теории Друде (и всех последующих классических теорий). Однако и сейчас для нас остается полезным данное Друде описание дрейфа и электропроводности.

Друде предположил, что движущиеся электроны рассеиваются в результате беспорядочных соударений с атомными остатками. Термин «беспорядочные» указывает на то, что *средняя* скорость непосредственно после любого акта соударения равна нулю. Это означает, что любая дрейфовая скорость в направлении внешнего поля уничтожается в результате рассеивающего столкновения. В промежутках между соударениями поле может изменять скорость движения электрона по величине и направлению, но результат его действия ликвидируется всякий раз, когда происходит соударение. Таким образом, чем больше среднее время свободного пробега между соударениями τ_m, тем сильнее влияние внешнего поля.

Можно определить среднее время свободного пробега τ_m как величину, обратную вероятности столкновения в единицу времени. Рассмотрим группу из n_0 электронов в момент времени $t=0$. Тогда число электронов, не испытавших соударения к моменту времени t, составляет

$$n_t = n_0 \exp(-t/\tau_m). \tag{3.9}$$

Скорость, с которой уменьшается число электронов, не испытавших соударения, равна

$$\frac{-dn}{dt} = \frac{n_t}{\tau_m} = \frac{n_0}{\tau_m} \exp\left(-\frac{t}{\tau_m}\right). \tag{3.10}$$

Предположим теперь, что движение и столкновение электронов происходит в электрическом поле $\mathbf{E}$. По истечении времени t электрон, не испытавший соударения, приобретает *в дополнение* к своей тепловой скорости дрейфовую скорость

$$\Delta \mathbf{v}_t = \frac{-e\mathbf{E}t}{m}. \tag{3.11}$$

Следовательно, расстояние, которое электрон проходит в направлении поля, составляет

$$x_t = \frac{-e\mathbf{E}t^2}{2m}. \tag{3.12}$$

Этот дрейф *накладывается* на беспорядочное тепловое движение. Формула (3.12) является обычным следствием второго закона Ньютона для расстояния, пройденного с постоянным ускорением.

Суммарное перемещение в направлении поля n_0 электронов за один свободный пробег равно

$$\int_0^\infty \mathbf{x}_t \left(\frac{dn}{dt}\right) dt = \left(\frac{-e\mathbf{E}n_0}{2m\tau_m}\right) \int_0^\infty t^2 \exp(-t/\tau_m)\, dt =$$

$$= \left(\frac{-e\mathbf{E}n_0\tau_m^2}{m}\right) \int_0^\infty \frac{1}{2} y^2 e^{-y} dy = \frac{-e\mathbf{E}n_0\tau_m^2}{m}. \tag{3.13}$$

Такой путь прошли бы n_0 частиц, имеющих одинаковое время свободного пробега τ_m и *среднюю дрейфовую скорость*

$$\Delta\bar{\mathbf{v}} = \frac{-e\mathbf{E}\tau_m}{m}. \tag{3.14}$$

Если предположить, что в 1 м^3 металла содержится n электронов, обладающих в электрическом поле $\mathbf{E}$ одной и той же постоянной дрейфовой скоростью $\Delta\bar{\mathbf{v}}$, то плотность электрического тока будет равна

$$\mathbf{J} = -en\Delta\bar{\mathbf{v}} = ne^2\tau_m\mathbf{E}/m = \sigma\mathbf{E}, \tag{3.15}$$

где положительная скалярная величина

$$\sigma = ne^2\tau_m/m \tag{3.16}$$

есть удельная электропроводность. Такой *вид* выражения для электропроводности сохраняется в более сложных квантовых моделях, а также в зонной теории; различие состоит в том, как определены входящие в него величины n, m и τ_m.

Равенство (3.15) представляет собой хорошо известный закон Ома, состоящий в том, что приложенное напряжение и результирующий ток связаны линейным соотношением. При переходе к более сложным ситуациям станет ясно, что в некоторых случаях ток $\mathbf{J}$ может нелинейно зависеть от поля $\mathbf{E}$. Мы увидим также, что σ может быть тензорной величиной, связывающей поле в одном направлении с плотностью тока в другом.

Электропроводность часто выражают не через среднее время свободного пробега, а через другие величины. Обычное выражение имеет вид

$$\sigma = ne\mu, \tag{3.17}$$

где величина

$$\mu = -\Delta \bar{v}/E = e\tau_m/m \tag{3.18}$$

представляет собой *дрейфовую подвижность* электронов (дрейфовую скорость, приобретаемую электроном в поле с единичной напряженностью). Электропроводность также часто выражают через *среднюю длину свободного пробега* $\lambda = s_T\tau_m$ электронов, определяемую как расстояние, проходимое электроном, движущимся с *тепловой* скоростью, за среднее время свободного пробега τ_m. Таким образом, в модели Друде электропроводность σ может быть записана в виде

$$\sigma = \frac{ne^2\tau_m}{m} = \frac{ne^2\lambda}{ms_T} = \frac{ne^2\lambda}{(3mk_0T)^{1/2}}. \tag{3.19}$$

Поскольку почти у всех обычных металлов электропроводность в широком интервале температур изменяется с температурой как T^{-1}, Друде был вынужден предположить, что столкновения электронов с неподвижным распределением ионных остатков приводят каким-то образом к тому, что средняя длина свободного пробега изменяется с температурой как $T^{-1/2}$. Почему должна быть такая зависимость, было совершенно непонятно; это слабое место теории послужило стимулом к созданию новых кинетических моделей электронного газа. При дальнейшем охлаждении электропроводность типичного металла возрастает как T^{-5}, а затем выходит на низкотемпературное насыщение (см. рис. 3.1); эти особенности температурной зависимости σ еще труднее объяснить с точки зрения классических моделей свободных электронов.

С помощью кинетической теории Друде, использованной здесь для вычисления σ, можно аналогичным образом получить электронную теплопроводность

$$\varkappa_e = \frac{2}{3}\tau_m s^2 C_{\text{кл}}. \tag{3.20}$$

Величина $C_{\text{кл}}$ в формуле (3.20) обозначает классическую удельную теплоемкость *электронного газа* при постоянном объеме. Как уже было указано, в модели электронного газа Друде эта удельная теплоемкость должна быть равна $(3/2)k_0n$, так что

$$\varkappa_e = nk_0\tau_m s^2 = \left(3n\tau_m k_0^2 T/m\right). \tag{3.21}$$

Объединяя выражение (3.21) с формулой (3.16) для электропроводности, можно получить число Лоренца

$$L \equiv (\varkappa_e/\sigma T) = 3(k_0/e)^2 = 2{,}2 \cdot 10^{-8}\,(\text{В/К})^2. \qquad (3.22)$$

Значение числа Лоренца, полученное Друде, находится, как можно видеть из табл. 3.1, в удивительно хорошем согласии с его значениями, экспериментально измеренными при комнатной температуре, для многих металлов. Это был один из самых выдающихся успехов теории Друде, послуживший стимулом для развития более сложных моделей.

Отношение теплопроводности к электропроводности для большинства реальных металлов при охлаждении ниже 100 К, как можно видеть из табл. 3.1, существенно отклоняется от закона Видемана — Франца. Однако для многих металлов было установлено, что при очень низких температурах величина $(\varkappa_e/\sigma T)$ снова становится постоянной. (Нечего и говорить, что модель Друде этого не объясняет.) Рис. 3.1 показывает, что для меди отношение $(\varkappa_e/\sigma T)$ становится опять постоянным при температурах столь низких, что величина τ_m как для электропроводности, так и для теплопроводности определяется рассеянием на примесях и дефектах.

Модель Лоренца

Друде использовал грубое упрощение, приписав всем электронам одну и ту же тепловую скорость. От этого упрощения отказался Г. А. Лоренц (1905 г.) (не следует путать с Л. Лоренцем, который публиковал свои работы двадцатью пятью годами раньше). Лоренц[3] исследовал свойства классического распределения электронных скоростей Максвелла — Больцмана. Он рассмотрел возмущение распределения Максвелла в присутствии электрического или температурного градиента, каждый из которых должен приводить как к смещению, так и к искажению симметрии равновесной функции распределения (3.4).

Лоренц воспользовался приближением *кинетического уравнения Больцмана*, с которым мы уже сталкивались в разд. 2.5 в связи с фононной теплопроводностью и диффузией в твердом теле. Тогда было указано, что приближение уравнения Больцмана используется во многих областях физики для описания переноса подвижных частиц путем рассмотрения их статистического распределения. В модели Лоренца речь идет о переносе электрического заряда, связанного с частицами электронного

[3] Лекции Лоренца имеются в английском переводе: *Lorentz H. A.* The Theory of Electrons, Dover, 2nd ed., 1952.

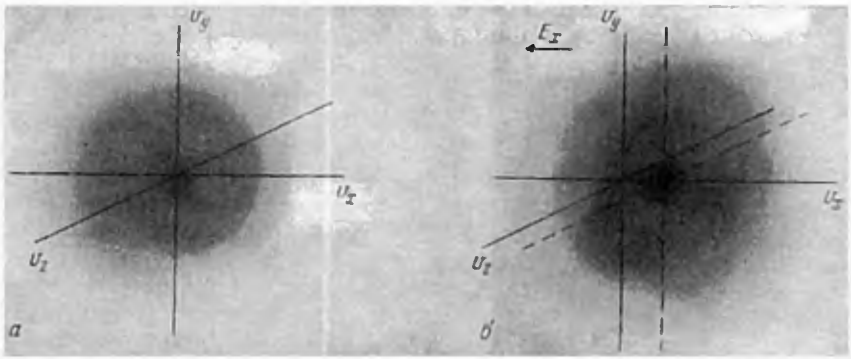


Рис. 3.5. Распределение скоростей в пространстве скоростей для больцмановского электронного газа в термодинамическом равновесии (*а*) и распределение того же ансамбля электронов, когда к нему приложено электрическое поле в *x*-направлении (*б*). Центр тяжести распределения теперь совпадает с точкой пересечения двух штриховых линий Распределение, несомненно, не только сместилось, но и изменило форму

газа (электропроводность), а также с переносом теми же электронами кинетической энергии (электронная теплопроводность).

Строгий анализ электропроводности в больцмановском приближении достаточно сложен, однако намеченное здесь упрощенное рассмотрение позволяет показать все его основные моменты. Более подробное обсуждение читатель найдет в монографиях Зейтца [4] или Займана [5].

Предположим, что распределение скоростей электронного газа в металле в отсутствие каких-либо внешних полей и воздействий соответствует равновесному распределению f_0 [формула (3.4)]. Это распределение было показано на рис. 3.4, *а* и повторно изображено на рис. 3.5, *а*. Чтобы упростить уравнение Больцмана, предположим, что металл однороден, так что величина f_0 не зависит от пространственных координат.

Приложим теперь к металлу электрическое поле, создающее направленный дрейф электронов. Распределение скоростей f в присутствии поля будет, вообще говоря, отличаться от f_0. Для простоты опять рассмотрим случай однородного поля **E**, при котором пространственная производная величины $(f—f_0)$ равна нулю. Тогда для любой точки металла и любого момента

[4] *Seitz F.* Modern Theory of Solids, McGraw-Hill, 1940. [Имеется перевод: *Зейтц Ф.* Современная теория твердого тела.— М.— Л.: Гостехиздат, 1948.]

[5] *Ziman J. M.* Principles of Theory of Solids, Cambridge, 2nd ed., 1972 [Имеется перевод: *Займан Дж.* Принципы теории твердого тела — М · Мир, 1966.]

времени, когда действует поле, скорость изменения функции распределения f со временем представляет собой сумму двух слагаемых

$$\left(\frac{df}{dt}\right)=\left(\frac{\partial f}{\partial t}\right)_{\text{пол}}+\left(\frac{\partial f}{\partial t}\right)_{\text{столкн}}. \tag{3.23}$$

Первый член справа отражает влияние поля на изменение распределения, а второй отражает роль рассеяния в восстановлении равновесного распределения. Уравнение (3.23) должно выполняться в каждой точке пространства скоростей.

Поскольку сила равна скорости изменения импульса, а электрон со скоростью $\mathbf{v}$ обладает импульсом $m\mathbf{v}$ и в электрическом поле $\mathbf{E}$ испытывает силу $-e\mathbf{E}$, первый член в правой части уравнения (3.23) равен

$$\left(\frac{\partial f}{\partial t}\right)_{\text{пол}}=\left(\frac{\partial \mathbf{v}}{\partial t}\right)\cdot\left(\frac{\partial f}{\partial \mathbf{v}}\right)=\left(\frac{-e\,\mathbf{E}}{m}\right)\cdot\left(\frac{\partial f}{\partial \mathbf{v}}\right). \tag{3.24}$$

Таким образом, влияние электрического поля на изменение функции распределения определяется производной этой функции по скорости (которая связана с производной по энергии).

Предполагается, что рассеивающие столкновения полностью уничтожают результаты ускоряющего действия поля на длине пробега, предшествующей столкновению, и восстанавливают распределение скоростей f_0. Поэтому, рассматривая роль столкновений в минимизации величины $(f-f_0)$, Лоренц предположил, что величина $(\partial f/\partial t)_{\text{столкн}}$ должна быть прямо пропорциональна (f_0-f). Эту пропорциональность можно выразить через время релаксации τ_r:

$$(\partial f/\partial t)_{\text{столкн}}=(f_0-f)/\tau_r. \tag{3.25}$$

Заметим, что величина τ_r определена для каждой точки пространства скоростей. Поэтому время релаксации, конечно, связано со средним временем свободного пробега τ_m модели Друде (формулы (3.8) и (3.16)), однако эта связь выражается интегральным соотношением.

С помощью выражений (3.24) и (3.25) уравнение (3.23) можно представить в виде *уравнения непрерывности*

$$\left(\frac{df}{dt}\right)+\frac{e\,\mathbf{E}}{m}\cdot\left(\frac{\partial f}{\partial \mathbf{v}}\right)+\frac{f-f_0}{\tau_r}=0, \tag{3.26}$$

связывающего f_0 с функцией распределения, возмущенной электрическим полем.

На рис. 3.5, *б* показана функция распределения, смещенная электрическим полем в отрицательном направлении оси x. Заметим, что положение «центра тяжести» нового распределения, так же как и любое искажение его формы, определяется кон-

куренцией между действием приложенного поля и процессов рассеяния. Предположение, сделанное Лоренцем и полностью оправданное для не слишком больших полей, состояло в том, что сдвиг величины f как целого, произведенный полем, должен быть мал по сравнению с $s_{\text{ср. кв}}$. Лоренц предположил также, что деформация распределения должна быть менее заметна, чем общий сдвиг.

Если постоянное электрическое поле действует в течение времени, которое велико по сравнению с τ_r, то достигается стационарное состояние и первый член уравнения (3.26) обращается в нуль. Тогда из условия непрерывности следует, что второй и третий члены должны быть равны по величине и противоположны по знаку в каждой точке пространства скоростей, откуда получаем

$$f = f_0 + (\tau_r e\,\mathbf{E}/m)\cdot(\partial f/\partial \mathbf{v}) \tag{3.27}$$

— возмущенное распределение для стационарного состояния. Интегрирование по этому распределению дает для всего электронного газа отличную от нуля дрейфовую скорость, т. е. металл обладает конечной электропроводностью. Если мы предположим, что поле, как на рис. 3.5, *б*, направлено вдоль оси x, то плотность тока равна

$$J_x = -\int e v_x f\, dv_x\, dv_y\, dv_z, \tag{3.28}$$

а поскольку интеграл, содержащий f_0, равен нулю, получаем

$$J_x = \sigma E_x = -\int (E_x e^2/m)\,\tau_r v_x (\partial f/\partial \mathbf{v})\, dv_x\, dv_y\, dv_z. \tag{3.29}$$

Чтобы выполнить это интегрирование, сделаем еще одно упрощение, предположив, что τ_r зависит только от *величины* скорости электрона, но не зависит от *направления* его движения. При решении уравнения Больцмана обычно предполагают, что τ_r зависит от скорости электрона по степенному закону:

$$\tau_r = A s^j. \tag{3.30}$$

Показатель степени j зависит от природы механизма рассеяния. Далее Лоренц предположил, что электроны упруго рассеиваются на неподвижном распределении атомных остатков. Для такого механизма рассеяния средняя длина свободного пробега λ не зависит от скорости электрона, как указано в связи с формулой (3.7) и в задаче 3.2. Это означает, что величина λ должна быть равна величине A в формуле (3.30), а $j = -1$, т. е.

$$\tau_r = \lambda/s. \tag{3.31}$$

Если именно этот механизм рассеяния является определяющим, то поле оказывает наиболее сильное воздействие на медленно движущиеся электроны.

Подстановка (3.31) в (3.29) дает для электропроводности интегральное выражение

$$\sigma = -\int (\lambda e^2 v_x/ms)\cdot(\partial f/\partial v_x)\, dv_x dv_y dv_z, \qquad (3.32)$$

которое можно преобразовать в интеграл по скалярной скорости s, учтя, что после усреднения по всем электронам $v_x^2 = (s^2/3)$ и что объем сферической оболочки в пространстве скоростей, имеющей радиус s и толщину ds, равен $4\pi s^2 ds$. Таким образом, интеграл (3.32) преобразуется в интеграл по скалярной скорости электрона s:

$$\sigma = (4\pi e^2/3m)\int_0^\infty \lambda s^2 (-\partial f_0/\partial s)\, ds. \qquad (3.33)$$

Так же просто электропроводность можно выразить через интеграл по кинетической энергии электрона $\varepsilon = \frac{1}{2} ms^2$.

Следует отметить, что главный вклад в электропроводность в выражении (3.33) дает интервал скоростей электронов, для которых функция распределения имеет максимальную отрицательную производную по скорости[6]. В задаче 3.4 предлагается вычислить интеграл (3.33) (и аналогичный интеграл по энергии), используя выражение (3.4) для f_0, и убедиться в том, что

$$\sigma = \frac{4ne^2\lambda}{3\,(2\pi m k_0 T)^{1/2}}. \qquad (3.34)$$

Отметим, что выражение (3.34) имеет ту же общую форму, что и в модели Друде. Если длина свободного пробега определена так же, то выражения (3.19) и (3.34) различаются только множителем $(3\pi/8)^{1/2} = 1{,}09$. Аналогичным образом, приближение кинетического уравнения Больцмана, использованное Лоренцем, дает для электронной теплопроводности результат, подобный результату Друде, однако уменьшенный примерно на одну треть[7].

---

[6] Для функции f_0, определяемой выражением (3.4), такое утверждение может показаться не столь существенным, однако этот факт приобретает совершенно новый смысл, когда уравнение Больцмана используется в квантовых моделях проводимости, в которых производная функции распределения Ферми — Дирака отлична от нуля только вблизи скорости и энергии Ферми.

[7] Объединяя результаты Лоренца для тепло- и электропроводности, можно получить предсказываемое его теорией значение числа Лоренца $L = (\varkappa_e/\sigma T) = 2(k_0/e)^2$. Напомним, что результат Друде для числа Лоренца (который довольно хорошо соответствовал экспериментальным данным) был на 50 % больше значения, полученного здесь. Может показаться странным, что более реалистическое распределение Больцмана оказывается в этом отношении менее точным, чем модуль Друде, однако этому не нужно придавать большого значения. Ни в модели Друде, ни в модели Лоренца не использовано *правильное* распределение Ферми — Дирака. Приближение Лоренца

Другие эффекты, предсказываемые моделью Лоренца

Согласно модели Лоренца, как и модели Друде, кинетическая энергия электронов должна составлять $(3/2)nk_0T$. Это в свою очередь должно привести к большой электронной теплоемкости, которая на самом деле не наблюдается. Поскольку, как теперь известно, парамагнетизм обусловлен спиновыми магнитными моментами электронов, больцмановский газ свободных электронов должен был бы также обладать и большой магнитной восприимчивостью, по величине обратно пропорциональной абсолютной температуре и с зависимостью от магнитного поля, определяемой функцией Ланжевена[8]. Тот факт, что такая магнитная восприимчивость на опыте отсутствовала, не разочаровал Лоренца, поскольку идея электронного спина появилась гораздо позднее.

Из решения Лоренцем кинетического уравнения Больцмана для комбинации нескольких внешних воздействий возникает ряд интересных следствий. Так, в проводнике, в котором от одного конца к другому меняется температура, возникает электродвижущая сила (эффект Томсона). Выражение для величины этого эффекта может быть получено путем решения кинетического уравнения Больцмана с распределением электронных скоростей, изменяющимся вдоль градиента температуры. Более сложные эффекты возникают при одновременном действии электрического, температурного и магнитного полей[9].

Если на электроны в металле действует магнитное поле **B** и электрическое поле **E**, то, согласно уравнениям Максвелла, полная сила, действующая на электрон, имеющий скорость **v**,

$$\mathbf{F} = -e[\mathbf{E} + (\mathbf{v} \times \mathbf{B})], \tag{3.35}$$

а не просто $-e\mathbf{E}$. Соответственно производную $(df/dt)_{\text{пол}}$

интересно для нас тем, что решение кинетического уравнения Больцмана лежит в основе и современных одноэлектронных теорий переноса. Различные вторичные эффекты (такие, как магнетосопротивление, эффект Нернста и т. п.) в модели Лоренца и всех последующих моделях возникают как естественное следствие теории.

[8] В гл. 5 функция Ланжевена будет обсуждаться для частично упорядоченной системы спинов в магнитном поле в случае, когда тепловые флуктуации препятствуют упорядочению.

[9] Многообразие эффектов, которые можно наблюдать в металлах в присутствии электрического, температурного и магнитного полей, детально рассмотрено в монографии: *Campbell L. L.*, Galvanomagnetic and Thermomagnetic Effects, Longmans, 1923, reprinted by Johnson Reprints, 1960. Теоретическое обсуждение в этой ранней работе проводится на основе классических моделей. Более современное объяснение, основанное на квантовых моделях электронного газа можно найти в книге: *Jan J. P.*, Galvanomagnetic and Thermomagnetic Effects in Metals, Sol. St. Phys., v. 5, ed. F. Seitz, D. Turnbull, Academic Press, 1957.

в уравнении (3.24) следует заменить более общим выражением

$$(\partial f/\partial t)_{\text{пол}} = (-e/m)[\mathbf{E} + (\mathbf{v} \times \mathbf{B})]\,\partial f/\partial \mathbf{v}. \qquad (3.36)$$

Аналогичным образом следует обобщить уравнение (3.26).

Заметим, что магнитное поле непрерывно изменяет направление движения электрона, поскольку магнитная сила Лоренца действует в направлении $\mathbf{v} \times \mathbf{B}$. Это действие магнитного поля, усредненное по распределению скоростей, в термодинамическом равновесии равно нулю, однако результирующий эффект возникает в результате взаимодействия магнитного поля с малыми направленными скоростями дрейфа, обусловленными электрическим полем. Такое взаимодействие приводит к тому, что электроны, движущиеся вдоль металлического прямоугольного стержня, должны одновременно перемещаться к одной из его сторон в направлении $\mathbf{B} \times \mathbf{v}_D$ (см. задачи 3.5 и 3.6). Возникающее при этом небольшое отклонение от электростатической нейтральности приводит к появлению электрического поля, препятствующего поперечному движению, так что в стационарном состоянии плотность тока направлена точно вдоль стержня, однако полное электрическое поле $\mathbf{E}$ имеет как параллельную, так и перпендикулярную току $\mathbf{J}$ компоненты.

При наличии обоих полей стационарное решение кинетического уравнения Больцмана в первом приближении по $\mathbf{B}$ имеет вид

$$\sigma \mathbf{E} = \mathbf{J} + [\pi e\lambda/2(2\pi m k_0 T)^{1/2}]\,\mathbf{J} \times \mathbf{B} = \mathbf{J} - [\sigma R_H]\,\mathbf{J} \times \mathbf{B}. \qquad (3.37)$$

Здесь σ — электропроводность, определяемая выражением (3.34). Параметр R_H называется постоянной *Холла*. Этот параметр определяет величину поперечного электрического поля, возникающего в присутствии магнитного поля (Холл впервые сообщил об экспериментальном наблюдении этого эффекта в 1879 г.)

Сравнение выражений (3.34) и (3.37) показывает, что в модели Лоренца постоянная Холла должна определяться выражением

$$R_H = -(3\pi/8ne), \qquad (3.38)$$

которое зависит от плотности электронов, но не зависит от средней длины свободного пробега. Мы можем по экспериментальным холловским данным рассчитать электронную плотность, как это сделано в табл. 3.2.

В одном из столбцов табл. 3.2 приведены значения произведения $(-\sigma R_H)$, которое называется *холловской подвижностью* носителей заряда. Подвижность измеряется в единицах $м^2/В \cdot с$ (что совпадает с $Тл^{-1}$) и отличается только численным множителем, близким к единице, от дрейфовой подвижно-

Таблица 3.2. **Сравнение экспериментальных данных по эффекту Холла и электропроводности с результатами классической модели Лоренца для переноса электронов**

Металл	Постоянная Холла* R_H, м³/Кл	Электропроводность** при $T = 100$ К σ_{100}, Ом⁻¹·м⁻¹	Концентрация электронов, рассчитанная по коэффициенту Холла $n = -3\pi/8eR_H$, м⁻³	Концентрация атомов, рассчитанная по постоянным решетки, м⁻³	Холловская подвижность при $T = 100$ К $(-\sigma R_H)_{100}$ м²/В·с	Средняя длина свободного пробега, вычисленная по холловской подвижности при $T = 100$ К λ, Å
Натрий	$-2{,}5\cdot10^{-10}$	$8{,}7\cdot10^{7}$	$2{,}9\cdot10^{28}$	$2{,}5\cdot10^{28}$	0,0218	76
Серебро	$-8{,}4\cdot10^{-11}$	$2{,}4\cdot10^{8}$	$8{,}8\cdot10^{28}$	$5{,}9\cdot10^{28}$	0,0202	70
Кадмий	$+6{,}0\cdot10^{-11}$	$4{,}3\cdot10^{7}$	Отрицательная	$4{,}6\cdot10^{28}$	Отрицательная	Отрицательная
Алюминий	$-3{,}0\cdot10^{-11}$	$2{,}1\cdot10^{8}$	$2{,}5\cdot10^{28}$	$6{,}0\cdot10^{28}$	0,0063	22
Железо	$+2{,}5\cdot10^{-11}$	$8{,}0\cdot10^{7}$	Отрицательная	$8{,}5\cdot10^{28}$	Отрицательная	Отрицательная
Никель	$-6{,}0\cdot10^{-10}$	$1{,}0\cdot10^{8}$	$1{,}2\cdot10^{28}$	$9{,}2\cdot10^{28}$	0,0600	210

* Данные по постоянным Холла взяты из книги: Hadbook of Physics, eds. E. U. Condon, H. Odishaw, McGraw-Hill, 1958.

** Данные по электропроводности взяты из книги: *Meaden G. T.* Electrical Resistance of Metals, Plenum Press, 1965.

сти в модели Друде [см. (3.17)]. Средние длины свободного пробега, рассчитанные по данным о холловской подвижности, приведены в последнем столбце табл. 3.2. Таблица составлена так, чтобы отразить переход от простых одновалентных металлов к многовалентным и переходным металлам. Заметим, что для некоторых металлов эффект Холла положителен. Для этих «аномальных» металлов (которые с точки зрения зонной теории не являются аномальными) значения холловской подвижности и длины свободного пробега в таблице опущены.

Как видно из табл. 3.2, плотность электронов, вычисленная по эффекту Холла, довольно хорошо согласуется с атомной плотностью для одновалентного натрия и не слишком хорошо — для одновалентного серебра. Для трехвалентного алюминия холловские данные указывают, что на атом приходится скорее четыре, чем три электрона. Для других элементов, перечисленных в таблице, расхождения еще более серьезны.

Первый член в правой части равенства (3.37) показывает, что присутствие магнитного поля, имеющего компоненту, перпендикулярную **J**, не должно уменьшать электропроводности электронного газа *вдоль* образца. Однако фактически электропроводность любого металла в магнитном поле уменьшается.

Этот эффект *магнетосопротивления* был открыт Томсоном в 1856 г. Если уравнение Больцмана для модели Лоренца решать путем разложения по степеням **B** с учетом более высоких степеней **B**, то из него следует, что должно наблюдаться уменьшение электропроводности, зависящее от квадрата магнитного поля (кроме того, имеются меньшие члены, содержащие более высокие степени **B**). Эта квадратичная зависимость находится в согласии с экспериментом.

В модели, в которой предполагается, что все электроны имеют одинаковую тепловую скорость (как в модели Друде), эффект магнетосопротивления должен отсутствовать. Однако в силу того, что на самом деле *существует* распределение электронных скоростей, магнитное отклонение под действием силы Лоренца точно компенсируется электростатическим полем Холла только для электрона, обладающего некоторой «средней» скоростью. Электрон, имеющий бóльшую или меньшую скорость, должен двигаться от одного конца образца к другому по криволинейной (и, следовательно, более длинной) траектории. Из-за криволинейности траекторий многих электронов макроскопическая электропроводность имеет меньшую величину:

$$\sigma_B = \frac{J^2\sigma}{J^2 + (\sigma R_H)^2 |\mathbf{J} \times \mathbf{B}|^2}. \tag{3.39}$$

Таким образом, магнетосопротивление максимально, когда магнитное поле перпендикулярно направлению тока, и в модели свободных электронов оно должно обращаться в нуль, когда поле **B** параллельно **J**.

В результате осложняющих особенностей, которые можно объяснить только в рамках зонной теории, для некоторых металлов продольное магнетосопротивление (**B** параллельно **J**) *не* обращается в нуль. Более того, для некоторых металлов экспериментальные значения коэффициента Холла оказываются положительными, что нельзя объяснить с помощью выражения (3.38) для свободных электронов. Тем не менее для различных гальваномагнитных эффектов результаты Лоренца были большим шагом вперед. Оказалось, что типичные значения постоянной Холла (см. табл. 3.2) действительно согласуются с представлением об одном или нескольких свободных электронах на атом.

Несостоятельность классических моделей

Мы уделили некоторое внимание обсуждению классических моделей Друде и Лоренца, чтобы подвести фундамент под современные квантовые модели. Наиболее серьезные недостатки классических моделей заключаются не в аномальном знаке по-

стоянной Холла для некоторых металлов и не в более сложном поведении магнетосопротивления. Гораздо важнее тот факт, что металл не обнаруживает ни большой удельной теплоемкости, ни магнитной восприимчивости, которых следовало бы ожидать при наличии совершенно свободных носителей заряда. Обескураживает также и то, что средняя длина свободного пробега, рассчитанная по результатам измерений эффекта Холла и электропроводности, очень велика по сравнению с межатомными расстояниями (см. последний столбец табл. 3.2) и увеличивается при охлаждении. Такое поведение никак не следует из выражения (3.7).

В течение двух десятилетий после выхода работы Лоренца появилось несколько статей, посвященных различным проблемам такого рода. Вин (1913 г.) пришел к выводу, что электроны, участвующие в проводимости, должны вплоть до абсолютного нуля обладать одной и той же скоростью, причем расположение положительных атомных остатков не должно препятствовать их движению. Однако ему самому не было ясно, каким образом можно обосновать эти условия. Дж. Дж. Томсон (1915 г.) предположил, что электрический заряд может распространяться по кристаллу в виде «электрических дублетов» — удивительное предвосхищение «куперовских пар», представление о которых было введено сорока годами позже и легло в основу современной теории сверхпроводимости. Линдеман (1915 г.) постулировал, что электроны могут образовать среду, подобную скорее не свободному газу, а жесткой решетке. Это предположение должно было объяснить отсутствие большого числа степеней свободы, которыми должен обладать газ согласно классической теории теплоемкости. Линдеман полагал, что жесткость такой электронной решетки обеспечивает высокую волновую скорость и, следовательно, хорошие проводящие свойства, несмотря на очень малую величину электронной теплоемкости.

Все эти и многие другие предположения оказались не в состоянии помочь исследователям проникнуть в суть проблемы электронного газа большой плотности. Указанные противоречия удалось в значительной степени разрешить в результате использования принципа Паули, лежащего в основе статистики Ферми — Дирака [10] для частиц с полуцелым спином (к ним относятся и электроны).

---

[10] *Fermi E.* — Z. Physik, **36**, 902 (1926); *Dirac P. A. M.* — Proc. Roy. Soc., A112, 661 (1926). В этих статьях впервые были объяснены следствия принципа Паули для больших систем электронов. Функция распределения Ферми — Дирака описана в любом учебнике статистической или атомной физики, приведенном в общем списке в начале этой книги.

3.3. Квантовая теория свободных электронов

Статистика Ферми—Дирака для электронного газа

Предполагается, что читатель знаком с представлением об электронном спине и с принципом запрета Паули, поэтому приводимое ниже обсуждение статистики Ферми — Дирака будет кратким.

Электроны в металле обычно обладают как кинетической, так и потенциальной энергией, зависящей от пространственного положения электрона. Пока рассмотрим газ *свободных электронов*, потенциальная энергия которого произвольно положена равной нулю. Тогда полная энергия электронного газа совпадает с кинетической и электрон со скоростью v и импульсом $p=mv$ имеет энергию $\varepsilon=(p^2/2m)$. В силу постулированной де Бройлем эквивалентности волновых и корпускулярных представлений этот электрон обладает волновым вектором $\mathbf{k}=(\mathbf{p}/\hbar)$. Электрон может быть представлен вектором, проведенным из начала координат в определенную точку **v**-, **p**- или **k**-пространства, или самой этой точкой. Одна из таких точек **k**-пространства показана на рис. 3.6. Заметим сразу, что все точки **k**-пространства, находящиеся на одинаковом расстоянии от начала координат, будут иметь одну и ту же энергию, поскольку выражение

$$\varepsilon=\frac{\hbar^2}{2m}\left[k_x^2+k_y^2+k_z^2\right]=\frac{\hbar^2k^2}{2m} \tag{3.40}$$

не зависит от направления **k**.

С классической точки зрения допустимо любое значение волнового вектора, и не существует ограничений на число электронов с одним и тем же волновым вектором. Первый из этих двух классических постулатов был опровергнут принципом неопределенности Гейзенберга (1925 г.), второй — принципом запрета Паули (1925 г.). Объединение этих двух принципов приводит к статистике Ферми — Дирака (1926 г.).

Формулируя условия выполнения принципа неопределенности Гейзенберга, потребуем, чтобы волновые функции электронов удовлетворяли уравнению Шредингера с определенными граничными условиями [11].

---

[11] Предполагается, что читатель в какой-то степени знаком с квантовомеханическим языком современной физики. Напомним, что, согласно принципу неопределенности Гейзенберга, существует связь между энергией электрона ε и зависящей от координат функцией, описывающей волну и дающую вероятность обнаружить электрон в некоторой точке пространства. Волновая

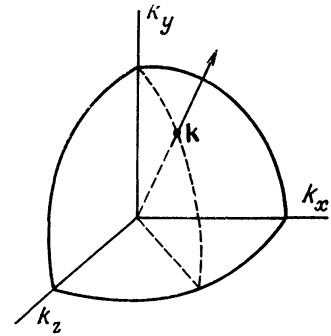


Рис. 3.6. Точка в k-пространстве служит для описания поведения электрона. Компоненты волнового вектора k характеризуют периодичность волновой функции в различных направлениях.

Здесь, так же как и для фононов, используется представление о циклических граничных условиях (волновая функция ψ периодична с периодом L вдоль каждой декартовой оси координат). Решение для электронной волновой функции, имеющее вид плоской волны,

$$\psi = C \exp(i\,\mathbf{k}\cdot\mathbf{r}) \tag{3.41}$$

может существовать только, если выполнены условия

$$\begin{aligned} k_x &= 2\pi n_x/L, \\ k_y &= 2\pi n_y/L, \\ k_z &= 2\pi n_z/L, \end{aligned} \tag{3.42}$$

функция ψ (которая может быть комплексной) принимает такие значения, что величина $-e|\psi|^2$ представляет собой плотность электрического заряда, получающуюся в результате усреднения движения электронов за бесконечно длительный период времени. Величины ε и ψ связаны уравнением Шредингера:

$$[(-\hbar^2/2m)\,\nabla^2 + \mathscr{V}(\mathbf{r})]\,\psi = \varepsilon\psi,$$

где $\mathscr{V}(\mathbf{r})$ — потенциальная энергия электрона, зависящая от координат. Поскольку электрон обладает зарядом $-e$, в электростатическом потенциале $V(\mathbf{r})$ он обладает потенциальной энергией $\mathscr{V}(\mathbf{r}) = -eV(\mathbf{r})$. В зонной теории энергия $\mathscr{V}(\mathbf{r})$ конечна и обладает периодичностью кристаллической решетки, что будет рассмотрено дальше. В квантовой теории свободных электронов предполагается, что всюду $\mathscr{V}(\mathbf{r}) = 0$, так что ε и ψ должны быть связаны уравнением

$$(-\hbar^2/2m)\,\nabla^2\psi = \varepsilon\psi.$$

Решения этого волнового уравнения имеют вид $\psi = A\sin(\mathbf{k}\cdot\mathbf{r}) + B\cos(\mathbf{k}\cdot\mathbf{r})$, или $\psi = C\exp(i\mathbf{k}\cdot\mathbf{r})$, если $\varepsilon = (\hbar^2k^2/2m)$, и подчиняются определенным граничным условиям. Эти граничные условия могут, например, быть поставлены так, что функция ψ должна обращаться в нуль на замкнутой поверхности в координатном пространстве (задача о частице в ящике). Тогда разрешенными становятся только стоячие волны. Используя идею Борна — Кармана о циклических граничных условиях, так же как это делалось для колебательных мод решетки, мы получим при той же плотности состояний бегущие волны, что имеет больший физический смысл.

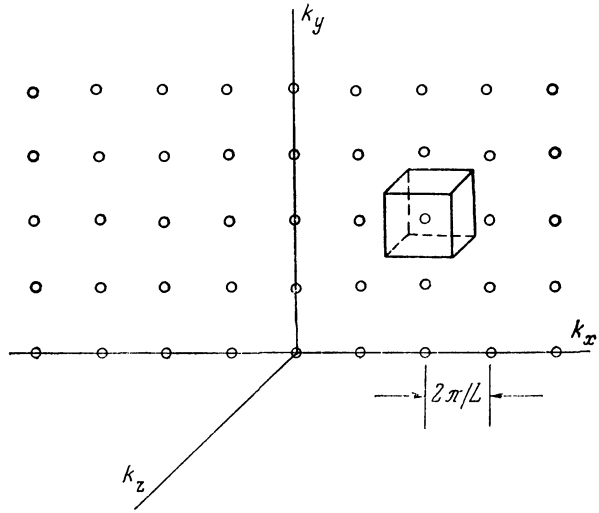


Рис. 3.7. Некоторые из разрешенных точек в **k**-пространстве для циклических граничных условий с периодом L для каждого из направлений в реальном пространстве. Все **k**-пространство будет заполнено, если каждая разрешенная точка окружена кубом объемом $(2\pi/L)^3$. Один такой куб показан на рисунке.

где n_x, n_y и n_z — вещественные целые числа. Теперь допустимым волновым функциям соответствуют только определенные точки **k**-пространства, так что **k**-пространство будет заполнено, если каждая из этих точек окружена кубической ячейкой, имеющей объем $(2\pi/L)^3$ (рис. 3.7).

Для реального металлического кристалла макроскопических размеров эти «ячейки» очень малы, а их число очень велико. Поэтому (как и в случае колебаний решетки), вычисляя число состояний, расположенных в пространстве волновых векторов между сферами радиусами k и $k+dk$, можно записать его как $(k^2L^3/2\pi^2)\,dk$. Тогда *на единицу объема* кристалла приходится $(k^2/2\pi^2)\,dk$ ячеек, попадающих в пределы скалярного интервала dk изменения волнового вектора. Плотность допустимых *электронных состояний* вдвое больше этой величины из-за двух возможных направлений спина электрона. Волновые функции для электронов со спинами $\pm 1/2$ в соответствии с принципом Паули рассматриваются как различные разрешенные состояния, хотя они и обладают одинаковыми значениями k_x, k_y и k_z. Таким образом, число *электронных состояний* на единицу объема в бесконечно малом интервале энергий или интервале изменения волнового вектора равно

$$g(\varepsilon)\,d\varepsilon = g(k)\,dk = (k/\pi)^2\,dk. \qquad (3.43)$$

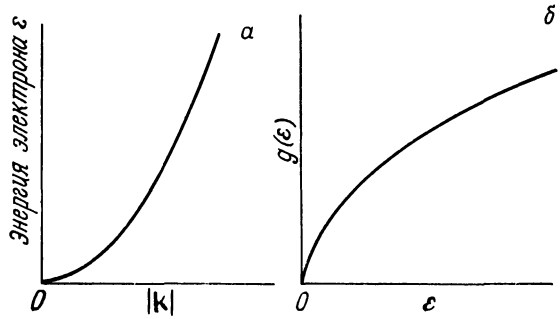


Рис. 3.8. Энергия электрона ε как функция величины его волнового вектора в соответствии с формулой (3.40) (*а*) и плотность различимых квантовых электронных состояний $g(\varepsilon)$ как функция энергии для свободных электронов согласно формуле (3.44) (*б*).

Поскольку энергия и волновой вектор связаны равенством (3.40) (как показано на рис. 3.8, *а*), величину $\mathbf{k}$ можно исключить, получив связь между энергией электрона и плотностью электронных состояний:

$$g(\varepsilon) = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \varepsilon^{1/2}. \qquad (3.44)$$

Этот параболический закон изменения плотности состояний показан на рис. 3.8, *б*.

Рассуждения, с помощью которых в гл. 2 была получена величина $g(\mathbf{k})$ для колебательных состояний, совершенно аналогичны приводимым здесь — следует отметить сходство формул (2.27) и (3.43). Поскольку закон дисперсии для электронов отличен от закона дисперсии для фононов, сходство между функциями $g(\omega)$ для фононов (задача 2.4) и $g(\varepsilon)$ для электронов [формула (3.44)] гораздо меньше.

Однако основное различие между квантовыми распределениями фононов и электронов состоит в том, что фононы подчиняются статистике Бозе — Эйнштейна, которая не налагает ограничений на число возбужденных фононов данной моды, в то время как электроны, будучи фермионами, должны подчиняться принципу запрета Паули. Из-за этого число заполнения электронами любого разрешенного состояния может иметь только два значения: единицу или нуль. В термодинамическом равновесии при температуре T выражение Ферми — Дирака для вероятности заполнения состояния с энергией ε имеет вид

$$f(\varepsilon) = \frac{1}{1 + B\exp(\varepsilon/k_0T)} = \frac{1}{1 + \exp\left(\dfrac{\varepsilon - \varepsilon_F}{k_0T}\right)}. \qquad (3.45)$$

Здесь B — безразмерный параметр нормировки. Нормировку можно выполнить и с помощью параметра $\varepsilon_F = -k_0 T \ln B$, имеющего размерность энергии. Условие нормировки состоит в том, что величина

$$n = \int_0^\infty f(\varepsilon) g(\varepsilon)\, d\varepsilon = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{\varepsilon^{1/2} d\varepsilon}{1 + \exp\left(\dfrac{\varepsilon - \varepsilon_F}{k_0 T}\right)} \qquad (3.46)$$

должна быть равна полной концентрации электронов в единице объема.

Если в единице объема электронного газа содержится очень мало электронов, то условие нормировки (3.46) при обычных температурах выполняется для отрицательного значения параметра ε_F. В этом случае функция распределения является больцмановской для всех реальных значений кинетической энергии, причем $f(\varepsilon) \ll 1$ даже для электронов с самыми малыми энергиями. Однако в реальном металле полная плотность электронов настолько велика, что равенство (3.46) может соответствовать отрицательному значению ε_F только для температур, превышающих 10^5 К. При обычных температурах величина ε_F в металле должна быть положительной; ее значение обычно называют *энергией Ферми* или *уровнем Ферми*.

Представим себе, что металлический «остов», состоящий из периодически расположенных атомных остатков, лишен электронного газа, присутствие которого в реальном металле необходимо для соблюдения условия электростатической нейтральности. Пусть теперь к ионному остову добавляются электроны до тех пор, пока электрическая нейтральность не будет восстановлена. Для простоты предположим, что такой процесс восстановления происходит при температуре абсолютного нуля.

Первые добавленные электроны займут состояния с наименьшей возможной кинетической энергией, которым соответствуют большие длины волн и малые значения волнового вектора, поскольку $\varepsilon = (\hbar^2 k^2/2m)$. Следующие добавленные электроны должны попасть в области $\mathbf{k}$-пространства, отстоящие дальше от начала координат, поскольку области, ближайшие к началу координат, уже заняты. Далее электроны должны последовательно заполнять сферические оболочки $\mathbf{k}$-пространства, соответствующие все большей энергии. Наконец, число добавленных электронов окажется достаточным для восстановления нейтральности, и последние электроны займут состояния, соответствующие низкотемпературной энергии Ферми ε_{F0},

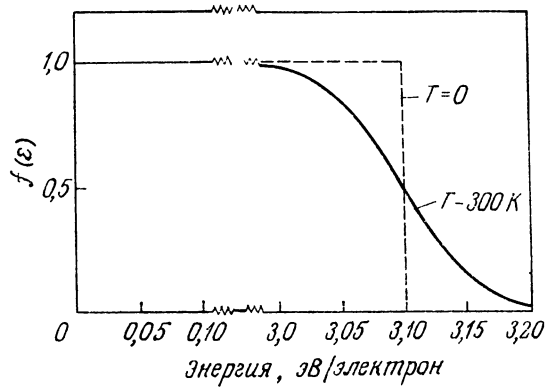


Рис. 3.9. Функция заполнения Ферми—Дирака [формула (3.45)] для плотного электронного газа при обычных температурах. Энергия Ферми, равная 3,1 эВ, соответствует металлическому натрию. Таким образом, видно, что для натрия при комнатной температуре заполнение будет полным вплоть до энергии $0{,}95\varepsilon_F$ и нулевым для энергий выше $1{,}05\varepsilon_F$.

удовлетворяющей уравнению

$$n = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\int_0^{\varepsilon_{F0}} \varepsilon^{1/2}\,d\varepsilon = \frac{1}{3\pi^2}\left(\frac{2m\varepsilon_{F0}}{\hbar^2}\right)^{3/2}, \qquad (3.47)$$

или, в численном виде,

$$n = 4{,}55\cdot 10^{27}\varepsilon_{F0}^{3/2}\ \text{м}^{-3}, \qquad (3.48)$$

где энергии выражены в электронвольтах. Поскольку металлы обычно содержат более 10^{28} свободных электронов в 1 м^3, низкотемпературная энергия Ферми

$$\varepsilon_{F0} = (3\pi^2 n)^{2/3}\hbar^2/2m \qquad (3.49)$$

составляет несколько электронвольт.

Кривые $f(\varepsilon)$ и $f(\varepsilon)g(\varepsilon)$ на рис. 3.9 и 3.10 обнаруживают резкий спад при абсолютном нуле. Это означает, что параметр нормировки, называемый энергией Ферми, представляет собой максимально возможную энергию электрона в электронном газе. Как видно из другой кривой на рис. 3.9, ситуация не слишком сильно изменяется даже при температурах значительно выше абсолютного нуля. Заметим, что функция заполнения $f(\varepsilon)$ падает от единицы до нуля, когда энергия проходит интервал в несколько k_0T с центром в точке ε_F. Поскольку величина k_0T при 300 К составляет всего 0,026 эВ, переходная область очень узка, и при всех температурах, лежащих ниже точки плавления металла, число занятых состояний с энергией выше ε_F очень мало.

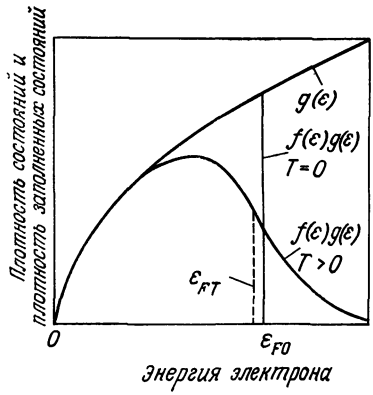


Рис. 3.10. Плотность электронных состояний как функция энергии в модели свободных электронов и плотность заполненных состояний в соответствии с законом Ферми — Дирака. При конечных температурах энергия очень медленно сдвигается вниз от своего положения при абсолютном нуле (см. задачу 3.10). На рисунке этот сдвиг сильно преувеличен. Изображенная здесь кривая для $T>0$ для большинства металлов соответствовала бы температуре нескольких тысяч кельвинов.

Строго говоря, существует различие между параметром ε_{F0}, который определяется полной плотностью свободных электронов согласно равенству (3.49), и параметром ε_F, определяемым из условия нормировки (3.46) для конечной температуры. При любой температуре величина ε_F представляет собой энергию, для которой вероятность заполнения равна 1/2; иногда ее называют *химическим* или *электрохимическим потенциалом*. Как показано на рис. 3.10, величина ε_F слегка падает с повышением температуры. Качественно это легко понять: с повышением температуры в результате теплового возбуждения оказываются заполненными дополнительные состояния выше ε_{F0}, поэтому какое-то число состояний ниже ε_{F0} становится незанятыми. Поскольку $g(\varepsilon)$ с энергией растет, энергия, соответствующая 50 %-ному заполнению, должна снижаться по мере того, как увеличивается тепловое уширение переходной области.

Чтобы выразить температурную зависимость энергии Ферми количественно, обратим внимание на то, что выражение (3.46) содержит интеграл, принадлежащий семейству *интегралов Ферми — Дирака*

$$F_j(y_0)=\int_0^\infty \frac{y^j dy}{1+\exp(y-y_0)}. \qquad (3.50)$$

Интегралы этого семейства не могут быть выражены в элементарных функциях для произвольного значения y_0, однако су-

ществуют их асимптотические выражения для больших отрицательных и больших положительных значений y_0 [12]. Таблицы численных значений для интегралов Ферми — Дирака целого и полуцелого порядков опубликованы [13].

Асимптотически для больших отрицательных y_0 эти интегралы имеют вид

$$F_j(y_0) \approx \Gamma(j+1)\exp(y_0), \quad (-y_0 > 2). \tag{3.51}$$

Это выражение будет использовано в гл. 4 при рассмотрении малых плотностей электронов в полупроводнике. Для электронных плотностей порядка тех, что существуют в металлах, интересен асимптотический вид интеграла Ферми — Дирака для больших положительных y_0, когда решение представляет собой ряд, начинающийся со следующих двух членов:

$$F_j(y_0) \approx \frac{y_0^{j+1}}{(j+1)}\left[1 + \frac{\pi^2 j\,(j+1)}{6y_0^2} + \ldots\right], \quad y_0 \gg 1. \tag{3.52}$$

Последующие члены ряда содержат возрастающие степени y_0^{-2}. Если j — целое число, то ряд конечен, если j — не целое число, то $F_j(y_0)$ представляет собой бесконечный сходящийся ряд. В любом случае, если y_0 достаточно велико, можно ограничиться первыми двумя членами.

В задаче 3.10 читателю предлагается, сравнивая выражения (3.46), (3.49), (3.50) и (3.52), убедиться, что в металле энергия Ферми при любой конечной температуре связана с энергией Ферми при $T=0$ соотношением

$$\varepsilon_F \approx \varepsilon_{F0}\left[1 - \frac{(\pi k_0 T)^2}{12\varepsilon_{F0}^2}\right]. \tag{3.53}$$

---

[12] Асимптотические выражения для больших положительных y_0 впервые обсуждались Зоммерфельдом (*Sommerfeld A. Z.*— Physik, 47, 1 (1928), после чего многократно рассматривались аналитические свойства этих интегралов. Большая часть этого материала приводится в приложении III книги: *Блейкмор Дж.* Статистика электронов в полупроводниках.— М.: Мир, 1964. Для ряда аналитических свойств может оказаться более удобным использование функций $\mathscr{F}_j(y_0)$, которые связаны с $F_j(y_0)$ соотношением

$$F_j(y_0) = \Gamma(j+1)\,\mathscr{F}_j(y_0).$$

[13] Таблицы для $F_{1/2}(y_0)$ и для других интегралов с полуцелым индексом были вычислены, они приведены в работе: *Mc-Dougall J., Stoner E. C.*— Phil. Trans. Roy. Soc., A237, 67 (1938). Имеются таблицы и в ряде современных работ*. Сводку этих данных, полученных для функции $\mathscr{F}_j(y_0)$, см. в приложении II книги: *Блейкмор Дж.* Статистика электронов в полупроводниках.— М.: Мир, 1964.

* Таблицы интегралов $F_j(y_0)$ для различных значений j приведены в монографии: *Фистуль В. И.* Сильно легированные полупроводники.— М.: Наука, 1965.— *Прим. ред.*

Как подтверждает численный пример в этой задаче, изменение энергии ε_F с температурой очень мало, и в дальнейшем мы, как правило, будем говорить об энергии Ферми ε_F, не указывая температуру, к которой она относится.

Удельная теплоемкость вырожденного электронного газа

Заинтересовавшись, как и многие другие авторы, приложением квантовомеханических представлений к макроскопическим системам, Арнольд Зоммерфельд[14] применил к металлу модель свободных электронов. Он использовал при этом статистику Ферми — Дирака, а не Масквелла — Больцмана. Оказалось, что при обычных температурах энергия большей части электронов электронного газа в металле меньше энергии Ферми на несколько k_0T или более. Эти электроны не могут испытать рассеивающего соударения, слабо изменяющего энергию, потому что все состояния с близкой энергией уже заняты, а их дополнительное заполнение запрещено принципом Паули[15]. Таким образом, только небольшая часть самых высокоэнергетичных электронов, составляющая примерно $(4k_0T/\varepsilon_F)$, способна реагировать на приложенный извне электрический или температурный градиент. Отсюда Зоммерфельд сделал вывод, что только электронные состояния, расположенные вблизи ε_F, определяют свойства металла.

Рассмотрим полную энергию электронного газа и ее производную по температуре, т. е. удельную электронную теплоемкость. При абсолютном нуле полная энергия электронного газа, согласно (3.44), минимальна и равна

$$U_0 = \int_0^{\varepsilon_{F0}} \varepsilon g(\varepsilon)\, d\varepsilon = \frac{\varepsilon_{F0}^{5/2}}{5\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}. \qquad (3.54)$$

[14] *Sommerfeld A.*— Naturwiss, 15, 824 (1927); Z. Physik, 47, 1 (1928). Обзор этих работ Зоммерфельда на английском языке см. в статье: *Sommerfeld A., Frank N. H.*— Rev. Mod. Phys, 3, 1 (1931).

[15] Электрон, находящийся в состоянии с малой энергией, лежащей намного ниже уровня Ферми, может быть рассеян в состояние со слабо отличающейся энергией, если другой электрон, первоначально занимавший это последнее состояние, одновременно будет также рассеян в какое-то оказавшееся поблизости незанятое состояние (например, в первоначальное состояние первого электрона). Поскольку электроны неразличимы, такой обмен местами двух электронов, конечно, не дает наблюдаемого результата. Поэтому рассеяние электрона в состоянии с низкой энергией можно обнаружить только в том случае, если он поднимется и займет какое-нибудь пустое состояние вблизи энергии Ферми. Это не так легко осуществить с помощью электрического поля. Из-за неразличимости электронов рассеяние электрона с большим изменением энергии имеет ту же вероятность, что и многоэлектронный процесс, в котором все промежуточные электроны продвигаются на бесконечно малый шаг. Легко видеть, что вероятность такого процесса крайне мала

С помощью выражения (3.47) эту энергию можно записать в виде

$$U_0 = (3n\varepsilon_{F0}/5). \tag{3.55}$$

Следовательно, в отсутствие теплового возбуждения на каждый электрон приходится средняя энергия $(3\varepsilon_{F0}/5)$. Электроны движутся с большими скоростями (как это видно из решения задачи 3.9), а при равновесии любому движущемуся электрону в любой момент времени в том же металлическом кристалле соответствует другой электрон, движущийся в противоположном направлении.

Энергия электронов при любой температуре выше абсолютного нуля несколько превышает U_0, потому что некоторые электроны оказываются термически возбужденными из состояний, лежащих несколько ниже ε_F, в состояния выше ε_F. При любой конечной температуре

$$U = \int_0^\infty \varepsilon g(\varepsilon) f(\varepsilon)\, d\varepsilon = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \int_0^\infty \frac{\varepsilon^{3/2} d\varepsilon}{1+\exp\left(\dfrac{\varepsilon-\varepsilon_F}{k_0T}\right)}; \tag{3.56}$$

в это выражение входит функция $F_{3/2}(y_0)$, принадлежащая семейству функций (3.50). Используя формулы (3.52) и (3.53), получаем

$$\int_0^\infty \frac{\varepsilon^{3/2} d\varepsilon}{1+\exp\left(\dfrac{\varepsilon-\varepsilon_F}{k_0T}\right)} \approx \frac{2\varepsilon_F^{5/2}}{5}\left[1+\frac{5}{8}\left(\frac{\pi k_0 T}{\varepsilon_F}\right)^2\right] \approx \frac{2\varepsilon_{F0}^{5/2}}{5} \times$$

$$\times \left[1-\frac{5}{24}\left(\frac{\pi k_0 T}{\varepsilon_F}\right)^2\right]\left[1+\frac{5}{8}\left(\frac{\pi k_0 T}{\varepsilon_F}\right)^2\right] \approx \frac{2\varepsilon_{F0}^{5/2}}{5} \times$$

$$\times \left[1+\frac{5}{12}\left(\frac{\pi k_0 T}{\varepsilon_F}\right)^2\right], \quad \varepsilon_F \gg k_0 T. \tag{3.57}$$

Таким образом, энергия, определяемая равенством (3.56), равна

$$U \approx \frac{\varepsilon_{F0}^{5/2}}{5\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\left[1+\frac{5}{12}\left(\frac{\pi k_0 T}{\varepsilon_F}\right)^2\right] = \left[U_0 + \frac{n\pi^2 k_0^2 T^2}{4\varepsilon_F}\right], \quad T \ll \frac{\varepsilon_F}{k_0}. \tag{3.58}$$

Отсюда в первом приближении можно получить удельную электронную теплоемкость

$$C_e = (\partial U/\partial T)_e = (\pi^2 n k_0^2 T/2\varepsilon_F). \quad (3.59)$$

Сравним эту величину со значением $C_{кл} = (3nk_0/2)$, ожидаемым для классического электронного газа. Квантовые ограничения изменили вклад электронной теплоемкости в полную теплоемкость:

$$C_e = (\pi^2 k_0 T/3\varepsilon_F) C_{кл}. \quad (3.60)$$

Этот результат часто формулируют в виде утверждения, что электронная теплоемкость в металле *вырождена*, т. е. уменьшена примерно в $(\varepsilon_F/3k_0T)$ раз по сравнению со своим классическим значением. Электронный газ, для которого $k_0T \ll \varepsilon_F$, называют *вырожденным газом* [16].

Результат Зоммерфельда для электронной удельной теплоемкости (3.59) по величине и температурной зависимости согласуется с электронной компонентой удельной теплоемкости, реально наблюдаемой в металлах. Поскольку C_e изменяется как первая степень абсолютной температуры, а решеточная теплоемкость зависит от температуры примерно как T^3, электронный вклад легче всего измерить при очень низких температурах. На рис. 3.11 представлена температурная зависимость полной удельной теплоемкости, имеющей вид суммы вкладов, рассчитанных Зоммерфельдом и Дебаем

$$C_p = AT + BT^3. \quad (3.61)$$

Как было указано в связи с рис. 2.21, каждое из двух слагаемых может быть определено раздельно по пересечению с осью ординат и наклону зависимости (C_p/T) от T^2.

Поскольку на измеряемые характеристики вырожденного электронного газа оказывают влияние только состояния, расположенные на расстояниях до нескольких k_0T от уровня Ферми, представляется логичным выразить C_e через параметры, непо-

[16] Если два состояния системы (например, два возбужденных состояния атома) обладают одинаковой энергией, то их часто называют вырожденными. К сожалению, термин «вырожденные» может иметь два совершенно разных значения. В настоящем разделе он использован в том смысле, что электронная теплоемкость вырождается (деградирует) по сравнению с ее большим значением, ожидаемым на основании классических моделей. Из дальнейшего изложения будет ясно, что ряд других свойств также вырождается в результате квантовых ограничений, поэтому говорят, что в металле имеется «сильно вырожденный газ». В гл. 4 увидим, что электронный газ в полупроводниковом материале может быть как вырожденным, так и невырожденным в зависимости от того, достаточно ли число подвижных электронов для того, чтобы стали существенными квантовые ограничения их движения.

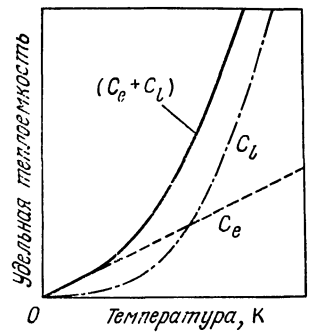


Рис. 3.11. Сложение удельной электронной теплоемкости C_e и удельной теплоемкости решетки C_l металла при низких температурах. Согласно теории Дебая, C_l изменяется как T^3, а C_e для вырожденного электронного газа линейно зависит от T.

средственно связанные именно с этой областью электронных энергий. Так, удельную электронную теплоемкость можно выразить через *плотность состояний* $g(\varepsilon_F)$ при энергии Ферми (см. задачу 3.12):

$$C_e = (\pi^2/3)\, k_0^2 T g(\varepsilon_F). \qquad (3.62)$$

Эта величина эквивалентна классической электронной теплоемкости, которая равна $(3k_0/2)$ на электрон для всех электронов, находящихся в интервале энергий, отличающихся от энергии Ферми на величину, слегка превышающую k_0T.

Парамагнетизм Паули

Квантовая статистика оказалась способной объяснить и другие очевидные несообразности классических моделей. В частности, она объяснила, почему большая система электронов, обладающих спином, имеет довольно малую и не зависящую от температуры магнитную восприимчивость.

После открытия электронного спина стало ясно, что парамагнетизм обусловлен направленной ориентацией магнитных моментов электронов из внешних атомных оболочек. Каждый электрон обладает магнитным моментом, равным одному *магнетону Бора:*

$$\mu_B = e\hbar/2m = 9{,}3 \cdot 10^{-24}\ \text{Дж/Тл} \qquad (3.63)$$

и направленным антипараллельно его угловому моменту (спину). Магнитные моменты электронов, спаренных на внутренних оболочках, взаимно компенсируются, однако во многих твердых телах происходит парамагнитное упорядочение спинов неспаренных электронов из внутренних оболочек[17]. Можно

[17] В гл. 5 будет обсуждаться парамагнитное упорядочение спинов неспаренных электронов из внутренних оболочек, которое происходит во многих материалах, в том числе переходных металлах и редкоземельных элементах. Коллективное упорядочение приводит к ферромагнитному, антиферромагнитному и ферримагнитному поведению.

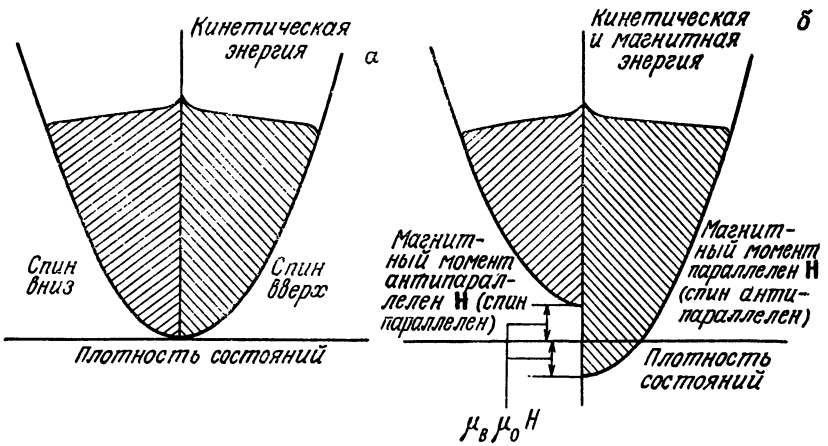


Рис. 3.12. Плотность состояний (ось абсцисс), построенная как функция энергии электрона (ось ординат). Представлены две компоненты плотности состояний, соответствующие противоположным направлениям спина. Заполнение показано для некоторой конечной температуры. *а* — одинаковое заполнение обеих групп в отсутствие магнитного поля; *б* — неодинаковое заполнение в присутствии магнитного поля, приводящее к парамагнетизму Паули.

ожидать, что спины внешних электронов, принадлежащих свободному электронному газу в металле, способны ориентироваться с образованием магнитного момента M, зависящего от приложенного поля.

Для *невырожденного* электронного газа этот магнитный момент должен определяться классической функцией Ланжевена [18]:

$$M = \chi_m H = n\mu_B L(\mu_B \mu_0 H/k_0 T). \qquad (3.64)$$

В магнитном поле, значительно меньшем, чем поле насыщения, $L(\mu_B \mu_0 H/k_0 T) \approx \mu_B \mu_0 H/3k_0 T$, поэтому магнитная восприимчивость χ_m не зависит от поля, однако зависит от температуры:

$$\chi_m = n\mu_B^2 \mu_0/3k_0 T. \qquad (3.65)$$

Такая зависимость, известная под названием *закона Кюри*, фактически ни в одном реальном металле не обнаруживается.

На рис. 3.12 показаны те особенности вырожденного электронного распределения, которые приводят к парамагнетизму, не описываемому законом Кюри. На эти особенности в 1927 г. указал Паули. Паули заметил, что его принцип запрета разрешает двум электронам с противоположно направленными спинами находиться в одной точке **k**-пространства (т. е. иметь

[18] Функция Ланжевена выводится в гл. 5.

одинаковые пространственные волновые функции). Следовательно, полную плотность состояний как функцию энергии можно представить в виде суммы двух компонент, каждая из которых соответствует половине величины $g(\varepsilon)$ [см. формулу (3.44)]. В отсутствие магнитного поля состояния, изображенные с одной стороны рис. 3.12, *а*, очевидно, должны быть заполнены так же, как состояния с другой стороны этого рисунка, соответствующие противоположному направлению спина.

Если теперь приложить магнитное поле, то можно считать, что одно направление спина *параллельно* **H**, а другое — *антипараллельно*. Состояния каждого типа, изображенные на рис. 3.12, сдвигаются соответственно на величину $\pm\mu_B\mu_0 H$ от энергии, соответствующей нулевому полю, как показано на рис. 3.12, *б*. Тогда заполнение состояний с антипараллельным спином (с параллельным магнитным моментом), обладающих меньшей энергией, будет увеличиваться за счет уменьшения заполнения состояний с параллельным спином таким образом, чтобы сохранился общий электрохимический потенциал. В результате возникает положительный магнитный момент, или парамагнетизм.

Следует отметить, что в приближении первого порядка намагниченность пропорциональна приложенному полю. Таким образом, магнитная восприимчивость χ_m является параметром, имеющим совершенно определенный физический смысл. Более того, на переход электронов из состояний с параллельным спином в состояния с антипараллельным спином почти не оказывает влияния небольшое тепловое размытие заполнения около энергии Ферми. Поэтому магнитная восприимчивость почти не зависит от температуры (что хорошо согласуется с данными эксперимента). Следовательно, мы можем просто вычислять χ_m при абсолютном нуле.

Как видно из рис. 3.12, результирующий магнитный момент для этих двух групп электронов равен

$$M = \chi_m H = (\mu_B/2)\int_{-\infty}^{\infty} f(\varepsilon)\,[g(\varepsilon+\mu_B\mu_0 H) - g(\varepsilon-\mu_B\mu_0 H)]\,d\varepsilon =$$

$$= (C\mu_B/2)\left[\int_0^{(\varepsilon_F+\mu_B\mu_0 H)} \varepsilon^{1/2}\,d\varepsilon - \int^{(\varepsilon_F-\mu_B\mu_0 H)} \varepsilon^{1/2}\,d\varepsilon\right] =$$

$$= (C\mu_B/3)\,[(\varepsilon_F+\mu_B\mu_0 H)^{3/2} - (\varepsilon_F-\mu_B\mu_0 H)^{3/2}]. \qquad (3.66)$$

Во второй и третьей строках этого равенства плотность состояний для обоих направлений спинов записана в виде $g(\varepsilon) = C\varepsilon^{1/2}$, где, как видно из формулы (3.44), $C = (1/2\pi^2)\,(2m/\hbar^2)^{3/2}$. Замена пределов интегрирования и возможность опустить $f(\varepsilon)$ под знаком интеграла следуют из того, что при абсолютном

нуле, для которого мы вычисляем интеграл, $f(\varepsilon)$ представляет собой ступенчатую функцию, изменяющуюся от единицы до нуля при энергии Ферми.

Если $\mu_B\mu_0 H \ll \varepsilon_F$ — условие, которое для обычных магнитных полей хорошо выполняется (см. задачу 3.13), результирующая парамагнитная намагниченность (3.66) принимает вид

$$M = \chi_m H \approx H\mu_B^2\mu_0 C\varepsilon_F^{1/2} = H\mu_B^2\mu_0 g(\varepsilon_F). \tag{3.67}$$

Таким образом, парамагнитную восприимчивость можно выразить через величину $g(\varepsilon_F)$ [так же как это было сделано для удельной теплоемкости (3.62)]:

$$\chi_m = \mu_B^2\mu_0 g(\varepsilon_F). \tag{3.68}$$

Вряд ли стоит удивляться, заметив, что эта величина соответствует магнитному действию тех электронов, энергия которых отличается не более чем на $\mu_B\mu_0 H$ от энергии Ферми. Поскольку эти электроны составляют лишь очень малую часть полного числа электронов, модель Паули правильно предсказывает, что магнитная восприимчивость (как и удельная теплоемкость) оказывается вырожденной, т. е. сильно уменьшенной по сравнению с ее классическим значением.

Степень вырождения легко оценить, выразив величину χ_m через отношение n/ε_F. Из формулы (3.44) и (3.49) следует, что

$$g(\varepsilon_F) = (1/2\pi^2)(2m/\hbar^2)^{3/2}\varepsilon_F^{1/2} = 3n/2\varepsilon_F. \tag{3.69}$$

Подставляя эту величину в формулу (3.68), мы получим для парамагнитной восприимчивости Паули выражение

$$\chi_m = 3n\mu_B^2\mu_0/2\varepsilon_F, \tag{3.70}$$

которое вырождено (уменьшено) в $(2\varepsilon_F/9k_0T)^{-1}$ раз по сравнению с результатом, даваемым классической функцией Ланжевена (3.65).

При выводе этого выражения для магнитной восприимчивости не было учтено влияние магнитного поля на поступательное движение свободных электронов. Для модели вырожденных свободных электронов такие эффекты впервые были рассмотрены Ландау в 1930 г. Более подробно они будут обсуждаться в разд. 3.5 в связи с циклотронным резонансом и осцилляционными магнитными эффектами, а здесь отметим только, что магнитное поле накладывает ограничения на движение электронов в направлениях, перпендикулярных магнитному полю (сила Лоренца); это приводит к появлению диамагнитной восприимчивости вида $-\mu_B^2\mu_0 n/2\varepsilon_F$. Таким образом, диамагнитная восприимчивость компенсирует одну треть парамагнитной воспри-

имчивости, и суммарная магнитная восприимчивость, с учетом поправок Паули и Ландау, составляет

$$\chi_m' = n\mu_B^2\mu_0/\varepsilon_F. \tag{3.71}$$

Полученный результат находится в удовлетворительном согласии с не зависящими от температуры значениями магнитной восприимчивости, измеряемыми для простых металлов, например щелочных.

Для большинства металлов для согласия формулы (3.71) с экспериментом требуется, чтобы электроны вели себя как частицы с *эффективной массой*, отличной от обычной массы электрона. Такая же перенормировка массы (которая влияет на плотность состояний, энергию Ферми при данном n и т. д.) обычно требуется и для количественной интерпретации данных по электронной теплоемкости этих металлов. Только такой увеличенной плотностью состояний можно объяснить относительно большую удельную теплоемкость и магнитную восприимчивость переходных элементов. Экспериментальные данные, соответствующие эффективной массе, меньшей, чем m, получены и для некоторых других твердых тел. Все эти аномалии может (в принципе) объяснить зонная теория твердых тел. Понятие эффективной массы будет обсуждаться в разд. 3.5.

Модель металлической проводимости Зоммерфельда

Мы уже рассмотрели вклад, который внес Зоммерфельд в теорию теплоемкости, применив к электронному газу статистику Ферми — Дирака. Теперь обратимся к его анализу проводимости такого газа [19]. Зоммерфельд использовал выражение (3.44) для плотности состояний и выражение (3.45) для функции заполнения. Таким образом, ответственными за проводимость могут быть только электроны с энергией, близкой к уровню Ферми. В остальном Зоммерфельд, рассматривая электронный перенос с помощью кинетического уравнения Больцмана, использовал те же предположения, что и Лоренц. Как и в разд. 3.2, это приближение приводит для электропроводности к выражению в виде интеграла [ср. с (3.33)]:

$$\sigma = \frac{-4\pi e^2}{3m}\int_0^\infty \lambda s^2\left(\frac{\partial f_0}{\partial s}\right)ds = \frac{-8\pi e^2}{3m^2}\int_0^\infty \lambda\varepsilon\left(\frac{\partial f_0}{\partial \varepsilon}\right)d\varepsilon. \tag{3.72}$$

Функция распределения Ферми — Дирака теперь нормирована так, что она дает распределение занятых состояний

[19] *Sommerfeld A.*— Z. Physik, **47**, 1 (1928). См. также: *Sommerfeld A., Frank N.*— Rev. Mod. Phys., **3**, 1 (1931).

в пространстве скоростей [как это было сделано в (3.4) для распределения Больцмана]. Новая функция распределения имеет вид

$$f_0 = \frac{2\,(m/2\pi\hbar)^3}{1+\exp\left(\dfrac{ms^2-2\varepsilon_F}{2k_0T}\right)} = 2\,(m/2\pi\hbar)^3 f(\varepsilon), \tag{3.73}$$

где $f(\varepsilon)$ — вероятность заполнения (3.45). Определение нормирующего множителя в (3.73) предлагается выполнить в задаче 3.14.

Подставив (3.73) в (3.72), запишем электропроводность в виде

$$\sigma = \frac{-2e^2m}{3\pi^2\hbar^3}\int_0^\infty \lambda\varepsilon\left(\frac{\partial f(\varepsilon)}{\partial\varepsilon}\right)d\varepsilon = \frac{-ne^2}{ms(\varepsilon_F)}\int_0^\infty \frac{\lambda\varepsilon}{\varepsilon_F}\left(\frac{\partial f(\varepsilon)}{\partial\varepsilon}\right)d\varepsilon. \tag{3.74}$$

Вторая форма записи в (3.74) приведена для того, чтобы вернуть читателя к отправной точке. Это преобразование было выполнено с помощью равенства (3.49) и выражения для скорости электрона с энергией Ферми, имеющего вид $s(\varepsilon_F) = (2\varepsilon_F/m)^{1/2}$.

Вклад в интеграл (3.74) может дать только энергетический интервал порядка нескольких k_0T вблизи энергии Ферми ε_F, поскольку величина $(df/d\varepsilon)$ как для больших, так и для меньших энергий обращается в нуль. На рис. 3.13 представлены функция $f(\varepsilon)$ и ее производная для конечной температуры. Максимальное значение производной составляет $(-1/4k_0T)$ при $\varepsilon=\varepsilon_F$. Дальнейшее исследование этой интересной функции предлагается в задаче 3.15. Учитывая, что функция $(df/d\varepsilon)$ близка к дельта-функции, можно утверждать, что для *любой* медленно изменяющейся функции энергии $\beta(\varepsilon)$ выполнено равенство

$$-\int_0^\infty \beta(\varepsilon)\left(\frac{\partial f(\varepsilon)}{\partial\varepsilon}\right)d\varepsilon = \beta(\varepsilon_F). \tag{3.75}$$

Следовательно, выражение

$$-\int_0^\infty \frac{\varepsilon\lambda}{\varepsilon_F}\left(\frac{\partial f(\varepsilon)}{\partial\varepsilon}\right)d\varepsilon = \lambda(\varepsilon_F) \tag{3.76}$$

определяет среднюю длину свободного пробега для электрона с энергией, точно равной энергии Ферми. Подставив (3.76) в (3.74), получим электропроводность, выраженную только че-

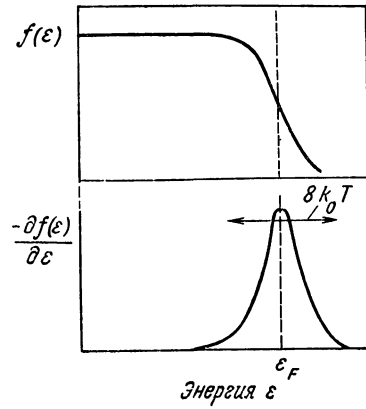


Рис. 3.13. Зависимость от энергии функции заполнения Ферми — Дирака $f(\varepsilon) = [1+\exp(\varepsilon-\varepsilon_F/k_0T)]^{-1}$ и ее производной. Производная $(-df/d\varepsilon)$ симметрична по энергии относительно точки ε_F, ее значение в максимуме равно $(1/4k_0T)$.

рез полную плотность электронов и характеристики электронов с энергией Ферми:

$$\sigma = \frac{ne^2\lambda(\varepsilon_F)}{ms(\varepsilon_F)}. \tag{3.77}$$

Для иллюстрации формул (3.72)—(3.77) сравним в пространстве скоростей или в k-пространстве смещение под действием электрического поля размытого распределения Максвелла (см. рис. 3.5) со смещением довольно резко ограниченной ферми-сферы (рис. 3.14). В последнем случае приложение электрического поля приводит к полному обеднению тонкой изогнутой области v-пространства и полностью заселяет зеркально симметричную ей область с противоположной стороны от заполненной части сферы Ферми. Все состояния с изменяющимся заполнением отвечают энергиям электрона, близким к энергии Ферми. Поэтому электропроводность (3.77) зависит только от $s(\varepsilon_F)$ и $\lambda(\varepsilon_F)$.

Электропроводность можно выразить через среднее время свободного пробега электронов между двумя рассеивающими столкновениями, которое для электронов, движущихся со скоростью Ферми, равно

$$\tau_m = \lambda(\varepsilon_F)/s(\varepsilon_F). \tag{3.78}$$

Электропроводность соответственно будет равна

$$\sigma = ne^2\tau_m/m. \tag{3.79}$$

Заметим, что это выражение по *форме* в точности совпадает с выражением, полученным в самой ранней модели Друде. Если мы займемся выводом из модели Друде формулы (3.16), то заметим, что отличие более поздних теорий заключается только в определении таких величин, как τ_m. Теперь понятно,

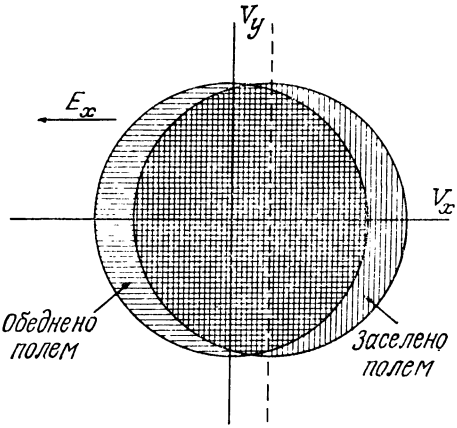


Рис. 3.14. Смещение электрическим полем заполненной области в пространстве скоростей для сильно вырожденного электронного газа. Ситуация для конечной температуры не будет сильно отличаться от изображенной здесь для $T=0$. В отсутствие поля заполненная область представляет собой сферу с центром в начале v-пространства. Внешняя поверхность этого объема называется *сферой Ферми*; в частности, и в том случае, когда заменой координат мы переходим к представлению в k-пространстве. Электрическое поле смещает сферу Ферми как целое на величину, зависящую от $\lambda(\varepsilon_F)$. Согласно зонной теории твердых тел, *поверхность Ферми* может отклоняться от сферической формы, однако смещение распределения Ферми приложенным полем происходит примерно так, как изображено здесь.

что для вырожденного электронного газа имеет значение среднее время свободного пробега только малой части всех электронов. Уже было отмечено, что Зоммерфельд провел существенное различие между большим числом «свободных электронов» и гораздо меньшим числом «проводящих электронов».

Приближение Зоммерфельда дает также выражение для электронной теплопроводности и значение $(\pi^2 k_0^2/3e^2)$ для числа Лоренца, которые при обычных температурах находятся в хорошем согласии с экспериментом. Для вырожденного электронного газа коэффициент Холла равен

$$R_H = -1/ne, \tag{3.80}$$

что на 15 % меньше классического результата (3.38). Подобным же образом Зоммерфельд получил решения уравнения Больцмана для магнетосопротивления и различных гальваномагнитных и термоэлектрических эффектов в сильно вырожденном электронном газе. Эти решения с большим или меньшим успехом могут быть применены к простым и сложным металлическим кристаллам.

Загадка нерассеивающихся электронов

Как было указано выше, Зоммерфельд использовал статистику Ферми — Дирака, но сохранил основные допущения Лоренца при решении кинетического уравнения. Среди них было и допущение о том, что электроны должны испытывать соударения с атомными остатками. Вследствие этого полученное им выражение (3.77) для электропроводности находилось в противоречии с величиной и температурным ходом электропроводности в металлах, известными из эксперимента.

Поскольку результат Зоммерфельда содержит в знаменателе не зависящую от температуры скорость

$$s(\varepsilon_F) = (2\varepsilon_F/m)^{1/2} \sim 10^6 \text{ м/с}, \tag{3.81}$$

единственной величиной в (3.77), которая может зависеть от температуры, оказывается средняя длина свободного пробега. Из экспериментальных значений электропроводности, которые приведены, например, в табл. 3.1 и 3.2, следует, что для хороших проводников величина $\lambda(\varepsilon_F)$ должна составлять при комнатной температуре несколько сотен ангстрем и становится еще больше при низких температурах[20]. Это находится в вопиющем противоречии с величиной ~ 10 Å, которая следует из формулы (3.7) для всех температур. Таким образом, необходимо как-то ответить на два вопроса.

1. Почему ожидаемые упругие соударения электронов с атомными остатками не происходят?

2. Какие процессы рассеяния *на самом деле* определяют среднюю длину свободного пробега?

Обсуждение легче начать со второго вопроса. Электроны, несомненно, должны рассеиваться как фононами, так и остаточными дефектами решетки. Под действием остаточных дефектов понимают рассеяние посторонними атомами, вакансиями, междоузельными атомами, дислокациями, границами зерен и внешними поверхностями[21]. Средняя длина свободного пробега

---

[20] Для данной величины электропроводности модель Зоммерфельда (3.77) требует значительно большей длины свободного пробега, чем классическая модель [выражение (3.19) или (3.34)], поскольку скорость электронов проводимости в вырожденном газе гораздо больше, чем средняя скорость, получающаяся из распределения Максвелла. Следовательно, средние длины свободного пробега, приведенные в табл. 3.2, возрастут примерно на два порядка, если их вычислять по квантовой модели. Эта разница исследуется в задачах 3.3. и 3.16.

[21] Если такие дефекты, как вакансии, междоузельные атомы и т. д., действительно остаточные, то их вклад в электрическое сопротивление может не зависеть от температуры, как предполагается при использовании правила Маттисена. Обычно такое предположение разумно для температур, лежащих значительно ниже точки плавления, когда можно ожидать, что $\lambda_{\text{деф}} \lesssim \lambda_{\text{фон}}$. Конечно, при температурах, близких к точке плавления твердого тела, проис-

$\lambda_{\text{деф}}$, обусловленная только рассеянием на дефектах, не должна зависеть от температуры, в то время как средняя длина свободного пробега для фононного рассеяния $\lambda_{\text{фон}}$ будет уменьшаться с повышением температуры. При одновременном действии этих двух механизмов (и в пренебрежении действием упругого рассеяния на ионах, как если бы его *не было*) средняя длина свободного пробега должна иметь вид

$$1/\lambda = 1/\lambda_{\text{деф}} + 1/\lambda_{\text{фон}}, \tag{3.82}$$

аналогичное соотношение имеет место для электропроводности

$$1/\sigma = 1/\sigma_{\text{деф}} + 1/\sigma_{\text{фон}}, \tag{3.83}$$

как показано на рис. 3.15. Такое сложение компоненты удельного сопротивления, зависящей от температуры и не зависящей от чистоты кристалла, и компоненты, не меняющейся с температурой, но зависящей от чистоты и совершенства[21] кристалла, хорошо известно экспериментаторам и обычно называется *правилом Маттисена* (1860).

Можно ожидать, что средняя длина свободного пробега, определяемая рассеянием на фононах, для температур, значительно превышающих температуру Дебая, должна изменяться как T^{-1}, поскольку для таких высоких температур число фононов всех частот пропорционально T. На противоположном конце температурной шкалы (при $T \ll \theta_D$) можно ожидать, что фононная плотность будет изменяться с температурой, как T^3. Более того, при низких температурах возбуждаются в основном длинноволновые фононы (см. задачу 2.16), которые могут рассеивать электроны только на малые углы. Суммарный эффект от изменения числа фононов и состава их спектра приводит к тому, что обусловленная фононами средняя длина свободного пробега при низких температурах должна изменяться, как T^{-5}. Пытаясь провести интерполяцию от этих двух предельных случаев на промежуточный температурный интервал, Грюнайзен в 1933 г. предложил полуэмпирическое выражение

$$\frac{1}{\lambda_{\text{фон}}} = CT^5 \int_0^{(\theta_D/T)} \frac{x^5 dx}{(e^x - 1)(1 - e^{-x})}, \tag{3.84}$$

которое оказалось в удовлетворительном согласии с данными по температурной зависимости электропроводности многих ме-

---

ходит тепловая генерация вакансий и междоузельных атомов (как показано на рис. 1.54), и, очевидно, правило Маттисена в этих условиях не выполняется.

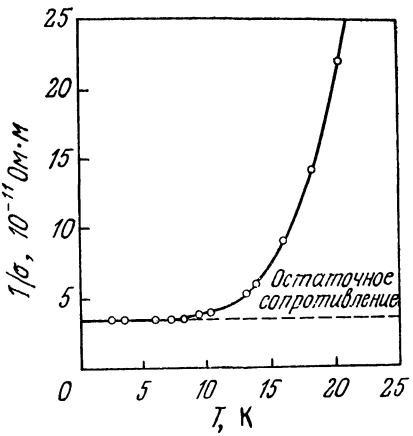


Рис. 3.15. Электрическое удельное сопротивление металла, достигающее при низких температурах постоянного значения, обусловленного рассеянием на дефектах (остаточное сопротивление), и складывающееся с ним сопротивление, обусловленное рассеянием на фононах. В соответствии с правилом Маттиссена эти два удельных сопротивления просто складываются. На рисунке приведены данные для натрия из работы: *MacDonald D. K. C., Mendelssohn K.*— Proc. Roy. Soc., **A202**, 103 (1950).

таллов. Кривая для меди на рис. 3.1 находится в довольно хорошем согласии с выражением Грюнайзена до тех пор, пока не начинает преобладать рассеяние на дефектах (ниже 20 К).

Таким образом, можно существенно продвинуться в объяснении величины и температурной зависимости средней длины свободного пробега и электропроводности, *если* удастся на законном основании отказаться от представления об упругом рассеянии на упорядоченном распределении ионов. *Необходимость* отказа от этого представления оказывается совершенно естественной, если мы рассматриваем электрон в идеальном кристалле как волну, для которой влияние периодического электрического потенциала уже учтено путем периодической пространственной модуляции волновой функции. Как это происходит, мы увидим в разд. 3.4.

Термоэлектронная эмиссия

Явления, для которых не существенно точное знание процессов рассеяния, описываются квантовой моделью свободных электронов достаточно хорошо, по крайней мере для более простых металлов. Мы уже убедились в этом на примере удельной теплоемкости и магнитной восприимчивости. Сейчас полезно отметить, что специального знания внутренних периодических потенциалов не требуется и для общего понимания уравнений, определяющих термоэлектронную эмиссию.

Как видно из рис. 3.3, для участия в термоэлектронной эмиссии из твердого тела электрон должен обладать кинетической энергией, не меньшей, чем $(\varphi+\varepsilon_F)$. Состояния с такими большими энергиями очень слабо заполнены, и для них можно

с уверенностью воспользоваться больцмановской асимптотической формой выражения (3.45), а именно

$$f(\varepsilon_F) \to \exp\left(\frac{\varepsilon_F - \varepsilon}{k_0 T}\right), \quad \varepsilon \gg \varepsilon_F \tag{3.85}$$

Иначе говоря, для столь высоких энергий плотность электронов (с учетом двух направлений спина) в пространстве скоростей определяется асимптотикой формулы (3.73)

$$f_0 dv_x dv_y dv_z = \frac{1}{4}\left(\frac{m}{\pi\hbar}\right)^3 \exp\left(\frac{\varepsilon_F}{k_0 T}\right) \exp\left[\frac{-m\left(v_x^2 + v_y^2 + v_z^2\right)}{2k_0 T}\right] \times \\ \times dv_x dv_y dv_z. \tag{3.86}$$

В качестве дальнейшего ограничения на плотность электронов, способных пересечь поверхность кристалла, укажем, что для того, чтобы выйти за поверхность в x-направлении, электрон должен приближаться к ней с x-компонентой скорости

$$v_x \geqslant [2(\phi + \varepsilon_F)/m]^{1/2} \tag{3.87}$$

независимо от компонент его скорости в плоскости yz. Конечно, коль скоро электрон покинул твердое тело, большую часть своей кинетической энергии он затратил на приобретение потенциальной энергии и потому будет двигаться в x-направлении с гораздо меньшей скоростью

$$u_x = [v_x^2 - 2(\phi + \varepsilon_F)/m]^{1/2}. \tag{3.88}$$

Распределение скоростей и энергий для электронов, вышедших наружу, исследуется в задаче 3.18.

Согласно (3.86),

$$n(v_x)\, dv_x = \frac{1}{4}\left(\frac{m}{\pi\hbar}\right)^3 \exp\left(\frac{\varepsilon_F}{k_0 T}\right) \exp\left(-mv_x^2/2k_0 T\right) \mathcal{J}^2 dv_x, \tag{3.89}$$

где $\mathcal{J}$ обозначает интеграл

$$\mathcal{J} = \int_{-\infty}^{\infty} \exp\left(-mv_y^2/2k_0 T\right) dv_y = (2\pi k_0 T/m)^{1/2} \tag{3.90}$$

или соответствующий интеграл, взятый в бесконечных пределах по v_z. Подставляя (3.90) в (3.89), видим, что число электронов, способных к эмиссии, в бесконечно малом интервале скоростей v_x составляет

$$n(v_x)\, dv_x = \left(m^2 k_0 T/2\pi^2\hbar^3\right) \exp\left(\varepsilon_F/k_0 T\right) \exp\left(-mv_x^2/2k_0 T\right) dv_x. \tag{3.91}$$

Плотность тока, представляющего собой поток высокоэнергетичных электронов в положительном направлении оси x из металла наружу в окружающий вакуум, будет равна

$$\mathbf{J} = \int_{[2(\phi+\varepsilon_F)/m]^{1/2}}^{\infty} ev_x [1-r(v_x)]\, n(v_x)\, dv_x. \qquad (3.92)$$

Здесь $r(v_x)$ — коэффициент отражения, который, согласно квантовомеханическому рассмотрению, должен быть введен даже в том случае, когда электрон обладает достаточной энергией для преодоления барьера. Доля отраженных электронов зависит от того, как именно изменяется поверхностный потенциальный барьер с расстоянием в атомном масштабе, о чем мы еще будем говорить в следующем разделе.

Обычно предполагается, что $r(v_x)$ не слишком сильно меняется на том малом интервале скоростей, который дает основной вклад в интеграл (3.92). Это предположение позволяет вынести множитель $(1-r)$ за знак интеграла. Тогда после подстановки выражения (3.91) для $n(v_x)$ формула для плотности тока эмиссии принимает вид

$$\mathbf{J} = (em^2k_0T/2\pi^2\hbar^3)(1-r)\exp(\varepsilon_F/k_0T)\int_{[2(\phi+\varepsilon_F)/m]^{1/2}}^{\infty} v_x \times$$
$$\times \exp\left[-mv_x^2/2k_0T\right] dv_x. \qquad (3.93)$$

Интеграл в (3.93) может быть легко вычислен путем замены переменной $y = (mv_x^2/2k_0T)$. Тогда

$$\int_{[2(\phi+\varepsilon_F)/m]^{1/2}}^{\infty} v_x \exp\left(-mv_x^2/2k_0T\right) dv_x = \frac{k_0T}{m}\int_{[(\phi+\varepsilon_F)/k_0T]}^{\infty} e^{-y}dy =$$
$$= \frac{k_0T}{m}\exp\left[\frac{-(\phi+\varepsilon_F)}{k_0T}\right]. \qquad (3.94)$$

Этот результат можно подставить в (3.93), что даст плотность тока термоэлектронной эмиссии

$$\mathbf{J} = (emk_0^2T^2/2\pi^2\hbar^3)(1-r)\exp(-\phi/k_0T), \qquad (3.95)$$

которая определяется отношением работы выхода к тепловой энергии k_0T. Формула (3.95) представляет собой закон Ричардсона — Дэшмана для термоэлектронной эмиссии[22], обычно выражаемый в виде

$$\mathbf{J} = AT^2(1-r)\exp(-\phi/k_0T), \qquad (3.96)$$

где

$$A = (emk_0^2/2\pi^2\hbar^3) = 1{,}2\cdot 10^6\ \text{А}/(\text{м}^2\cdot\text{К}^2). \qquad (3.97)$$

Значения величин $A(1-r)$ и ϕ могут быть определены по пересечению с осью ординат и наклону *кривой Ричардсона*, изображающей зависимость $\lg(J/T^2)$ от $1/T$ (рис. 3.16).

[22] *Dushman S.*— Rev. Mod. Phys., 2, 381 (1930).

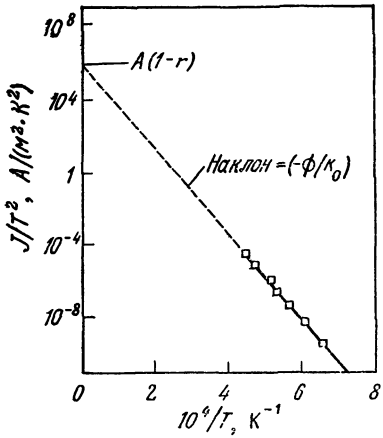


Рис. 3.16. График Ричардсона для вольфрама, по которому (в принципе) можно определить работу выхода и константу Ричардсона. По книге: *Herrmann G., Wagener S.* The Oxide Coated Cathode, Chapman and Hall, 1951. [Имеется перевод: *Герман Г., Вагнер С.* Оксидный катод.— М.: Гостехиздат, 1949.]

Численные данные, представленные в табл. 3.3, показывают, что значения $A(1—r)$, определенные по кривой Ричардсона, для элементов, лишенных покрытия, по порядку величины обычно совпадают с вычисленным значением A. Однако тщательные исследования термоэлектронной эмиссии[23] показы-

Таблица 3.3. **Термоэлектронная работа выхода и константа Ричардсона, определенные по температурной зависимости термоэлектронной эмиссии***

Металл	Работа выхода φ, эВ/электрон	Константа Ричардсона $A(1-r)$, $А/(м^2 \cdot К^2)$
Платина	5,3	$0{,}32 \cdot 10^6$
Вольфрам	4,5	$0{,}72 \cdot 10^6$
Молибден	4,4	$1{,}15 \cdot 10^6$
Таитал	4,1	$0{,}37 \cdot 10^6$
Кальций	3,2	$0{,}60 \cdot 10^6$
Торированный вольфрам	2,6	$0{,}03 \cdot 10^6$
Барий	2,5	$0{,}25 \cdot 10^6$
Цезий	1,8	$1{,}60 \cdot 10^6$
Цезированный вольфрам	1,4	$0{,}03 \cdot 10^6$

* Данные взяты из работы: *Herring C, Nichols M. H.* — Rev Mod Phys., 21, 185 (1949).

[23] См., например, *Herring C., Nichols M. H.*— Rev. Mod. Phys., 21, 185 (1949), а также: *Nottingham W. B.*— In: Handbuch der Physik, v. 21, Springer, 1956, p. 1.

вают, что работа выхода является довольно непостоянным «параметром». Попытки измерить ϕ осложняются несколькими эффектами, к числу которых относятся следующие.

1. Температурная зависимость ϕ, обусловленная тепловым расширением.

2. Небольшое различие в работе выхода для разных кристаллографических плоскостей. Это различие можно понять в свете зонной теории, однако оно совершенно необъяснимо в модели полностью свободных электронов. В поликристаллическом металлическом образце измеряемое значение ϕ обусловливается сложной комбинацией эффектов, однако определяющую роль играют кристаллические плоскости с наименьшей поверхностной работой выхода ϕ.

3. Изменение вероятности эмиссии поверхностным загрязнением.

Последний пункт может оказаться как помехой, так и благом, в зависимости от обстоятельств. Эффективная работа выхода для металла с энергией Ферми, расположенной значительно ниже уровня вакуума, очень чувствительна к присутствию на поверхности адсорбированных газов, примесных атомов и покрытий, а также к условиям заполнения поверхностных состояний, обусловленных нарушением периодичности решетки. Малые изменения величины ϕ вызывают весьма заметный эффект; так, изменение работы выхода, равной 5 эВ, на 1 % вызовет при температуре 1000 К увеличение тока эмиссии на 80 %. Огромное увеличение тока эмиссии наблюдается, когда атомы материала с низкой работой выхода распределены по поверхности металла с большим значением ϕ (задача 3.19 посвящена этому вопросу). Моноатомный и даже меньший слой такого материала способен радикально изменить работу выхода основного металла.

Уменьшение работы выхода путем нанесения поверхностных покрытий имеет большое техническое значение, поскольку делает возможной работу электронных ламп без нагрева катодов до очень высоких температур. Имеется обширная литература о катодах электронных ламп[24]. Несмотря на широкое использование транзисторов, и сейчас во многих случаях применяются электронные лампы с термоэлектронными эмиттерами. К ним относятся лампы высокой мощности для радиопередатчиков и катодно-лучевые трубки.

---

[24] См., например, *Herrmann G., Wagener S.* The Oxide Coated Cathode, Chapman and Hall, 1951. [Имеется перевод предыдущего издания: *Герман Г., Вагнер С.* Оксидный катод.— М.: Гостехиздат, 1949.]

Эммиссия во внешнем поле и автоэлектронная эмиссия

При обсуждении термоэлектронной эмиссии предполагалось, что потенциальная энергия меняется скачком от нуля внутри металла до $(\phi+\varepsilon_F)$ непосредственно у внешней поверхности, как показано на рис. 3.3 или на рис. 3.17 (кривая *1*). Квантовомеханический коэффициент отражения для крутого барьера составляет значительную величину даже для электронов с энергией, превышающей высоту барьера.

На самом деле, как было впервые указано в 1914 г. Шоттки, потенциальный барьер, преодолеваемый электроном, должен быть более пологим. Можно ожидать, что сначала величина $\mathscr{V}(x)$ растет с x линейно, однако, когда электрон выходит из тела на расстояние от поверхности, превышающее несколько ангстрем, он должен испытывать притягивающую силу электростатического изображения заряда $(-e)$ относительно однородной проводящей плоскости. Тогда его потенциальная энергия равна

$$\mathscr{V}(x)=(\phi+\varepsilon_F)-(e^2/16\pi\varepsilon_0 x)\ (\text{большие } x). \tag{3.98}$$

Правильное асимптотическое поведение для очень малых и очень больших x описывается выражениями

$$\mathscr{V}(x)=0, \quad x<0,$$
$$\mathscr{V}(x)=\frac{(\phi+\varepsilon_F)^2}{(\phi+\varepsilon_F)+(e^2/16\pi\varepsilon_0 x)} \quad x>0 \tag{3.99}$$

(см. кривую *2* на рис. 3.17). Для электрона с начальной кинетической энергией, слегка превышающей $(\phi+\varepsilon_F)$, квантовомеханическая вероятность отражения от барьера, описываемого формулами (3.99), значительно меньше, чем от барьера, имеющего вид прямоугольной ступеньки.

Предположим, что в вакууме снаружи от нагретого металлического кристалла в направлении x создано электрическое поле. Это изменит потенциальную энергию для расстояний от внешней поверхности металла, превышающих несколько ангстрем. Теперь имеем

$$\mathscr{V}(x)=(\phi+\varepsilon_F)-(e^2/16\pi\varepsilon_0 x)-exE_x. \tag{3.100}$$

Потенциальная энергия, измененная внешним полем, показана на рис. 3.17 (кривая *3*). Дифференцируя $\mathscr{V}$ по x, находим что существует максимум высоты потенциального барьера, расположенный при

$$x_{\text{макс}}=(e/16\pi\varepsilon_0 E_x)^{1/2},$$
$$\mathscr{V}_{\text{макс}}=(\phi+\varepsilon_F)-(e^3E_x/4\pi\varepsilon_0)^{1/2}. \tag{3.101}$$

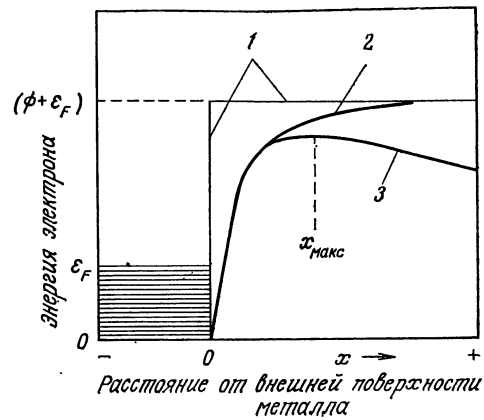


Рис. 3.17. Энергия электрона внутри металла и в вакууме, окружающем металл. За начало отсчета энергии выбрана энергия покоящегося электрона внутри кристалла. *1* — бесконечно крутой потенциальный барьер для термоэлектронов, получающийся если пренебречь силой изображения; *2* — предполагаемый вид барьера с учетом силы электростатического изображения между вылетевшим электроном и металлической поверхностью; *3* — изменение формы и некоторое понижение высоты барьера внешним электрическим полем.

Таким образом, наличие внешнего электрического поля приводит к небольшому понижению работы выхода. Это явление называется эмиссией во внешнем поле, или эмиссией Шоттки (описано в работах Шоттки 1914 и 1923 гг.). Как показывает расчет в задаче 3.20, понижение работы выхода мало для полей, составляющих несколько тысяч В/м, и поэтому максимум потенциала расположен на расстоянии многих ангстрем от поверхности кристалла. Даже очень небольшое уменьшение работы выхода позволяет участвовать в эмиссии многим электронам, которые при нулевом поле не обладают достаточной энергией. Эмиссия Шоттки широко исследована [23].

Когда электрическое поле достигает величины $\sim 10^8$ В/м, происходит другое явление — *автоэлектронная* [25], или *холодная, эмиссия*. Последнее название иногда используется потому,

[25] *Fowler R. H., Nordheim L. W.*— Proc. Roy. Soc., **A119**, 173 (1928). Более современное учебное изложение эмиссии в поле содержится в работах: *Good R. H., Müller E. W.*— In: Handbuch der Physik, v. 21, Springer, 1956, p. 176, а также: *Müller E. W., Tsong T. T.*, Field Jon Microscopy: Principles and Applications, Elsevier, 1969. В этих двух работах эмиссия во внешнем поле рассматривается с упором на микроскопическую структуру металлических решеток в атомном масштабе. В книге: *Duke C. B.* Tunneling in Solids, Academic Press, 1969, наряду с туннелированием, которое может происходить между двумя твердыми телами через узкий зазор вакуума или диэлектрика, обсуждается также и холодная эмиссия.

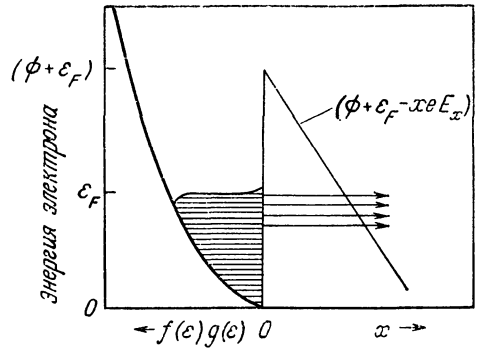


Рис. 3.18. Потенциальный барьер у поверхности металла при наличии очень сильного электрического поля. Если потенциальная энергия снаружи от металла падает ниже уровня Ферми на расстояниях от поверхности, составляющих несколько ангстрем, то может происходить автоэлектронная эмиссия.

что процесс одинаково легко происходит при любой температуре; термическая активация *над* потенциальным барьером перестает играть определяющую роль. На самом деле, как показано на рис. 3.18, поверхностный потенциальный барьер становится таким тонким, что через него может происходить квантовомеханическое туннелирование. При некотором критическом поле толщина барьера становится достаточно мала, чтобы электроны с энергией Ферми обладали конечной вероятностью туннелирования. (Это происходит, когда величина ϕ/eE_x составляет ~ 10 Å.) Для еще больших полей толщина барьера такова, что позволяет туннелировать электронам с еще меньшей энергией. Фаулер и Нордгейм[25] путем анализа квантовомеханической вероятности туннелирования через треугольный потенциальный барьер, показанный на рис. 3.18, получили следующее выражение для зависимости плотности тока от поля:

$$\mathbf{J} = \alpha E^2 \exp\left(-\beta\phi/E\right). \qquad (3.102)$$

Сравнение выражений (3.96) и (3.102) показывает, что в случае холодной эмиссии определяющим параметром становится вместо температуры электрическое поле как под знаком экспоненты, так и в предэкспоненциальном множителе.

Вероятность туннелирования электронов с энергией Ферми должна быть очень малой, пока толщина барьера для них не станет меньше 10 Å. Поэтому можно ожидать (поскольку ϕ обычно составляет ~ 3 эВ), что минимальная напряженность поля, необходимая для холодной эмиссии, должна быть равна примерно $3 \cdot 10^9$ В/м. На самом деле, как уже говорилось, такая эмиссия наблюдается уже при напряжениях, обеспечивающих

в 30 раз меньшее макроскопическое поле. Возможно, что благодаря локальным очень малым неровностям поверхности возникают очень большие электрические поля и почти вся автоэлектронная эмиссия идет из этих точек поверхности. Результаты экспериментальных исследований холодной эмиссии[25] находятся в разумном соответствии с формулой Фаулера — Нордгейма (3.102), содержащей два подгоночных параметра.

Фотоэмиссия

Мы рассмотрели эмиссию электронов из металла под действием тепла или сильного электрического поля. Здесь и в следующем разделе мы укажем на еще два эффекта, которые приводят к эмиссии электронов.

Фотоэмиссия, которую иначе называют фотоэлектрической эмиссией или внешним фотоэлектрическим эффектом,— это эмиссия электронов, происходящая при поглощении металлом фотонов с энергией, превосходящей ϕ. Этот процесс был корректно описан Эйнштейном в 1905 г. Фотон с энергией $\hbar\omega$ переводит электрон с энергией ε в состояние с энергией $(\varepsilon+\hbar\omega)$; этот электрон может выйти наружу, если $(\varepsilon+\hbar\omega)>(\varepsilon_F+\phi)$ и если скорость такого «горячего» электрона направлена к ближайшей поверхности.

Монохроматический поток фотонов обусловливает эмиссию фотоэлектронов, кинетическая энергия которых лежит в пределах от $(\hbar\omega-\phi)$ до нуля, поскольку начальные энергии электронов не превосходят ε_F, а какое-то количество фотоэлектронов потеряет часть своей кинетической энергии в одном или более столкновениях, которые они испытывают до того, как покинут металл. Спектр энергий фотоэлектронов можно исследовать путем приложения тормозящего электрического поля в вакууме вне металла.

Процесс фотоэмиссии имеет три стадии: 1) образование горячего электрона, 2) движение его к поверхности, 3) преодоление поверхностного потенциального барьера. Процесс 1 протекает в слое толщиной в несколько сот ангстрем от поверхности, поскольку оптический коэффициент поглощения металла для фотонов с энергией несколько эВ равен 10^7—10^8 м$^{-1}$. Практически «горячий» электрон, образовавшийся на расстоянии нескольких десятков ангстрем от поверхности, скорее всего достигнет поверхности, не потеряв энергии на соударения с другими электронами. Таким образом, критичными для фотоэмиссии являются свойства металла в пределах нескольких атомных слоев, прилегающих к поверхности.

Фотоэмиссия изучается как из-за тех сведений, которые она может дать о физике твердого тела, так и вследствие прак-

тической важности фотоэмиссионных катодов[26]. Такие катоды широко применяются во многих практически используемых чувствительных приборах, включая вакуумные и газонаполненные фотодиоды, фотоумножители (о которых пойдет речь в следующем разделе), электронно-оптические преобразователи и фотоэлектронные усилители изображения и телевизионные трубки. При малых энергиях фотонов (т. е. при работе в красной или инфракрасной области спектра) одним из возможных вариантов является использование цезия или другого металла с низкой работой выхода в качестве тонкого поверхностного катодного покрытия[26].

Фотоэмиссия широко применяется также для исследования спектра электронных состояний в твердом теле[27]. Необходимость таких сложных исследований станет гораздо более очевидной, когда мы познакомимся с зонной теорией твердых тел. Как мы увидим, «горячий» электрон при фотоэмиссии появляется в результате возбуждения из состояния одной зоны ниже уровня Ферми в состояние [над уровнем вакуума $(\varepsilon_F+\Phi)$] в *другой* зоне. Таким образом, вероятность фотовозбуждения зависит от плотностей состояний в *двух* зонах и от матричного элемента межзонного перехода. Большое число экспериментальных результатов показывает, что часто при таком возбуждении оказывается нарушенным принцип сохранения квазиимпульса[28]. В этой связи заметим, что волновой вектор не может сохраняться, если электрон из свободного квантового электронного газа испытывает фотовозбуждение, а затем испускается. Действительно, если энергия и волновой вектор связаны соотношением (3.40), то электрон с энергией $(\varepsilon+\hbar\omega)$ *должен* иметь волновой вектор, отличающийся от волнового вектора электрона с энергией ε. Изменение волнового вектора во много раз превышает волновой вектор участвующего в возбуждении фотона ($\sim 10^7$ м$^{-1}$). Анализ фотоэмиссии в твердом теле с точки зрения зонной теории показывает, что в одних случаях при фотовозбуждении волновой вектор $\mathbf{k}$ должен сохраняться, а в других — нет.

---

[26] *Sommer A. H.* Photoemissive Materials, Wiley, 1968.

[27] Экспериментальные методы, в которых фотоэмиссия используется для изучения плотности состояний, хорошо описаны в работе: *Spicer W. E.*— Phys. Rev., **112**, 114 (1958). Подробное описание такого метода в применении к меди и серебру, дано в работе: *Spicer W. E., Berglund C. N.*— Phys. Rev., **136**, A1044 (1964).

[28] *Spicer W. E.* Photoelectric Emission — In: Optical Properties of Solids, North-Holland, 1972, p. 755.

Вторичная электронная эмиссия

Электроны могут быть эмиттированы поверхностью твердого тела, если оно подвергается облучению заряженными частицами больших энергий. Впервые такое явление наблюдали Аустин и Штарк в 1902 г. при использовании умеренно быстрых электронов. Они показали, что наблюдается не просто часть падающих электронов, отраженных от металлической поверхности, поскольку в некоторых случаях испущенных электронов было больше, чем падающих. Такой процесс называется *вторичной электронной эмиссией*[29]. Особый интерес представляет случай, когда выход вторичных электронов δ (отношение числа эмиттированных электронов к числу падающих) превышает единицу.

Эмиссионный ток в действительности состоит из трех групп электронов:

1) первичных электронов, которые упруго отразились от поверхности металла;

2) тех первичных электронов, которые испытали неупругое отражение (вероятно, сопровождавшееся появлением одного или более вторичных электронов);

3) истинных вторичных электронов, которые могут потерять а могут и не потерять часть своей энергии в *дальнейших* столкновениях с электронами. Из-за этих дальнейших соударений вторичная эмиссия происходит из приповерхностного слоя толщиной в несколько десятков атомных слоев.

На рис. 3.19 показан энергетический спектр электронов, испущеннных серебром, облучаемым первичными электронами с энергией 155 эВ. Указаны группы *1*, *2* и *3*. Относительный вклад этих трех групп зависит от угла падения, от энергии падающих электронов, а также от свойств облучаемого материала.

Вторичные электроны возникают в результате кулоновского взаимодействия падающего электрона (или другой заряженной частицы) с электроном, лежащим внутри ферми-сферы в $\mathbf{k}$-пространстве. Бэруди[30] рассмотрел этот механизм для неэкранированной кулоновской силы

$$\mathbf{F} = (e^2\,\mathbf{r}/4\pi\varepsilon_0 r^3). \tag{3.103}$$

С такой силой первичный электрон, траектория которого при скорости $\mathbf{v}$ характеризуется параметром соударения $\mathbf{b}$ (расстоянием наибольшего сближения), действует на электрон, нахо-

---

[29] Хороший обзор см. в работе: *Dekker A. J.* Solid State Physics, v. 6, eds. F. Seitz., D. Turnbull, Academic Press, 1958, p. 251.

[30] *Baroody E. M.*— Phys. Rev., 78, 780 (1950).

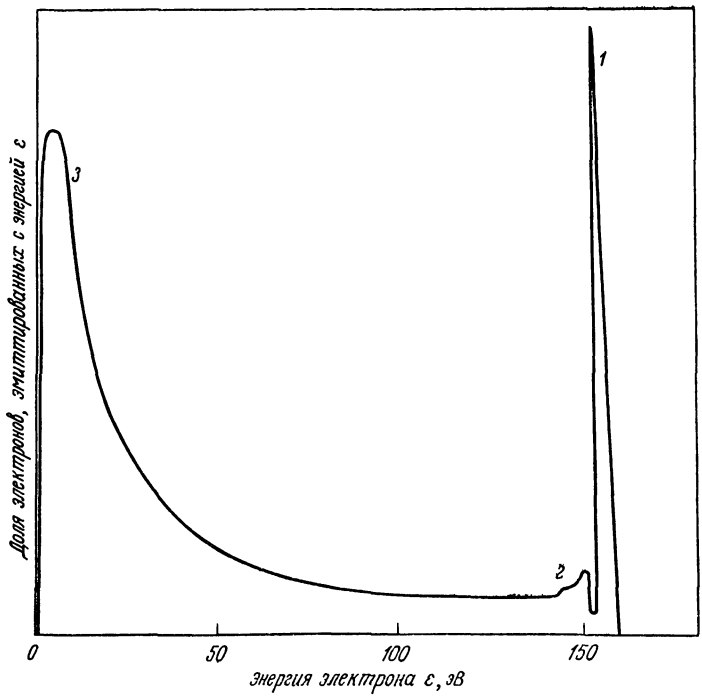


Рис. 3.19. Энергетический спектр вторичных электронов, испускаемых серебром, при облучении его первичными электронами с энергией 155 эВ. Из работы: *Rudberg E.*— Proc. Roy. Soc, **A127**, 111 (1930). Отмечены группы электронов (*1*), (*2*) и (*3*), обсуждаемые в тексте.

дящийся в состоянии с энергией Ферми или меньшей энергией. Он показал (предлагается в задаче 3.21), что удар быстрого электрона приводит к передаче электрону проводимости импульса

$$\Delta \mathbf{p} = e^2 \mathbf{b}/2\pi\varepsilon_0 b^2 \mathbf{v}, \qquad (3.104)$$

который, как будет показано, направлен перпендикулярно $\mathbf{v}$. Таким образом, первичный электрон, падающий нормально к поверхности твердого тела, создает вторичные электроны, первоначально движущиеся параллельно плоскости поверхности; эти электроны должны испытать по крайней мере одно соударение, чтобы получить направление импульса, которое позволит им выйти из кристалла.

Выражение (3.103) может служить разумной аппроксимацией для кулоновской силы, с которой быстрый электрон действует на электрон проводимости, в том случае, когда концент-

рация свободных электронов мала, как это бывает в полупроводнике. Для очень большой плотности электронов (10^{28} м$^{-3}$ и более) в металле кулоновская сила на расстоянии $\mathbf{r}$, превышающем несколько ангстрем, экранируется другими электронами. Поэтому теория Бэруди была модифицирована [31]; было предложено использовать экранированную кулоновскую силу

$$\mathbf{F} = (e^2\,\mathbf{r}/4\pi\varepsilon_0 r^3)\exp(-r/\lambda), \qquad (3.105)$$

где длина экранирования λ обычно составляет 10^{-10} м. Это делает образование вторичных электронов в металле крайне маловероятным, за исключением тех случаев, когда параметр b сравним с длиной λ или меньше ее.

Чрезвычайно привлекательный аспект как теории Бэруди, так и модифицированной теории с экранированной кулоновской силой состоит в том, что согласно им, выход вторичных электронов δ приблизительно пропорционален *квадратному корню* из работы выхода φ. Такое увеличение δ с φ противоречит интуитивным предположениям, однако зависимость $\delta \sim \varphi^{1/2}$ оказывается в полном согласии с экспериментально измеренными результатами в металлах.

Для большинства металлов максимальный выход δ_m обычно имеет порядок единицы. Так, для платины, имеющей одну из самых больших работ выхода ($\varphi \approx 6$ эВ), максимальный выход вторичных электронов $\delta_m \approx 1{,}8$, а в большинстве других металлов он меньше. Кривая выхода для платины показана на рис. 3.20, на котором также изображены данные для диэлектрика MgO, где вторичный выход гораздо больше. Ряд диэлектриков и полупроводников весьма эффективно испускают вторичные электроны. Практически обычно используется вторичная эмиссия либо из неметаллических монокристаллов, либо из тонкого слоя подходящего неметаллического вещества, на поверхности металлического кристалла. Так, оксидированные сплавы Mg—Ag или Be—Cu, интерметаллические соединения, такие, как Cs_3Sb, а также кристаллы фосфида галлия и арсенида галлия с успехом используются в качестве *вторично-электронных катодов* (*динодов*) для *фотоумножителей* [32]. В фотоумножителе фотоизлучающий катод находится внутри вакуумного баллона; фотоэлектроны могут ускоряться электрическим полем. Усиление тока достигается за счет использования ряда вторично излучающих поверхностей (динодов). Хорошо проду-

---

[31] См., например, *Van der Ziel A.*— Phys. Rev., **92**, 35 (1953).

[32] *Zworykin V. K., Ramberg E. G.* Photoelectricity and its Applications, Wiley, 1949.

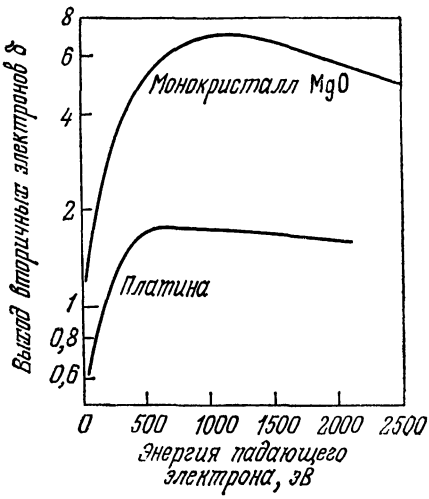


Рис. 3.20. Выход вторичных электронов δ как функция энергии первичных электронов для одного из наиболее эффективных металлических излучателей и очень эффективного диэлектрика. Данные для платины, полученные Коплэндом (Copeland P. E.), взяты из книги: *Dekker*. Solid State Physics, Prentice-Hall, 1957. Данные для MgO — из работы: *Johnson J. B., McKay K. G.*— Phys. Rev., **91**, 582 (1953).

манное устройство позволяет получить и обнаружить ток в 10^9 и более электронов при поглощении на фотоэмиссионном катоде единственного электрона [33].

Плазменные колебания электронного газа

В выражении (3.105) для экранированной кулоновской силы, которое было использовано при обсуждении механизма образования вторичных электронов, множитель $\exp(-\lambda/r)$, отличающий это выражение от обычной кулоновской силы (3.103), возникает из-за коллективного движения газа свободных электронов, который ведет себя как плазма [34], состоящая из подвижных заряженных частиц.

Название «плазма» было предложено Ирвингом Лэнгмюром в 1929 г. для описания коллективных электрических свойств, которые он обнаружил у ионизованного газа. Область физики газовой плазмы хорошо определена. Многие из явлений, наблюдаемых в газовой плазме, имеют место в электронном газе металла или полупроводника, однако при этом возможны некоторые интересные отличия и особенности.

---

[33] *Poultney S. K.* Advances in Electronics and Electron Physics, v. 31, ed. L. Marton, Academic Press, 1972, p. 39.

[34] Элементарное представление о плазме в твердых телах дано в работе: *Hoyax M. F.* Solid State Plasma, Pion, 1970. Более углубленное рассмотрение дано Гликсманом [*Glicksman M.* Solid State Physics, v. 26, Academic Press, 1971, p. 275].

Представим себе электронный газ (классический или квантовый) в твердом теле с фиксированной плотностью заряда атомных остатков (ионов), имеющего противоположный знак. При любом смещении электронного газа относительно системы ионов возникает возвращающее электрическое поле, ничем не связанный электронный газ будет колебаться относительно положения равновесия с характерной *плазменной частотой* ω_p. Для газа, содержащего n электронов на 1 м^3, смещенного как целое на расстояние x, возвращающее поле равно $E_x = (nex/\varepsilon_0)$. Таким образом, уравнение движения для каждого электрона в газе в отсутствие затухания имеет вид

$$m(d^2x/dt^2) = -eE_x = -(ne^2x/\varepsilon_0). \qquad (3.106)$$

Уравнение (3.106) описывает гармоническое колебание с собственной угловой частотой

$$\omega_p = 2\pi\nu_p = (2\pi c/\lambda_p) = (ne^2/m\varepsilon_0)^{1/2}. \qquad (3.107)$$

Эта частота соответствует длине волны, лежащей для газовых плазм в микроволновой области электромагнитного спектра ($\lambda_p = 0{,}1$ м для $n = 10^{17}$ м$^{-3}$), однако для плазмы в полупроводнике ее типичное значение расположено в инфракрасной области и в области ультрафиолета для плазмы в металле ($\lambda_p = 0{,}3$ мкм для $n = 10^{28}$ м$^{-3}$).

Причина, по которой, говоря об условиях плазменного резонанса, мы указываем области электромагнитного спектра, состоит в том, что плазма непрозрачна и полностью отражает электромагнитное излучение с частотой, меньшей, чем ω_p. Электромагнитная волна с напряженностью электрического поля $E_\omega = E_m \exp(i\omega t)$ создает такую электрическую индукцию $D_\omega = D_m \exp(i\omega t)$, что относительная диэлектрическая проницаемость составляет

$$\varkappa_\omega = D_\omega/\varepsilon_0 E_\omega = [1 - (\omega_p^2/\omega^2)]. \qquad (3.108)$$

Таким образом, для электромагнитного излучения с $\omega < \omega_p$, диэлектрическая проницаемость отрицательна (запрещает распространение), но она становится положительной (разрешает распространение) для $\omega > \omega_p$. Простые металлы (такие, как щелочные) действительно непрозрачны и являются полностью отражающими в видимой части спектра, однако прозрачны в ультрафиолетовой области для длин волн, меньших $\lambda_p = (2\pi c/\omega_p)$, что хорошо согласуется со значением, получающимся по формуле (3.108). (В задаче 3.22 дан пример такого соответствия.)

Большая частота плазменного резонанса электронного газа в металле означает, что такой газ быстро реагирует на любое изменение распределения заряда. Перераспределение заряда

электронного газа обеспечивает экранирование действия любого внесенного извне заряда на расстоянии, большем, чем $\lambda = s/\omega_p$. Для холодного электронного газа Ферми — Дирака соответствующая средняя скорость электронов есть $\bar{s} = (\varepsilon_F m)^{1/2}$. Следовательно, если воспользоваться формулой (3.107) для плазменной частоты и выражением (3.49), связывающим плотность электронов с энергией Ферми, то *радиус экранирования Томаса — Ферми* составит

$$\lambda_{\text{ТФ}} = (\bar{s}/\omega_p) = (\varepsilon_0 \varepsilon_F / ne^2)^{1/2} = \left(\frac{3\pi\hbar^2\varepsilon_0}{2me^2}\right)^{1/2}\left(\frac{\pi}{3n}\right)^{1/6}. \quad (3.109)$$

Длина экранирования (3.109) очень близка к 1 Å для плотности электронов 10^{28} м$^{-3}$, типичной для многих металлов. Таким образом, действие заряда, движущегося на расстоянии более одного ангстрема от атома, эффективно экранируется в соответствии с формулой (3.105).

Для электронного газа в полупроводнике (который мы рассмотрим позже, в гл. 4) распределение скоростей может быть не ферми-дираковским, а максвелловским. В таком *невырожденном* электронном газе средняя скорость электрона составляет $\bar{s} = (k_0 T/m)^{1/2}$ и *класссическая длина экранирования* равна

$$\lambda_{\text{кл}} = (\bar{s}/\omega_p) = (\varepsilon_0 k_0 T/ne^2)^{1/2}. \quad (3.110)$$

Для относительно малых плотностей электронов, которые возможны в полупроводнике, $\lambda_{\text{кл}}$ может быть очень большой по сравнению с межатомным расстоянием, тем не менее электронное экранирование играет существенную роль во многих электронных процессах в полупроводнике.

Продольные колебания электронного газа в металле, имеющие плазменную частоту, можно вызвать, направляя поток быстрых электронов через тонкую металлическую фольгу или вынуждая электрон либо фотон испытать неупругое отражение от поверхности металлической пленки. Энергия становится кратной квантованной *энергии плазмона* $\hbar\omega_p$. Потеря энергии электроном или фотоном, вызывающим плазменные колебания, может быть измерена (обычно она составляет от 5 до 15 эВ), что служит дополнительной проверкой значения ω_p, найденного по плотности электронов. Эти два значения ω_p не обязательно совпадают с точностью, лучшей нескольких процентов, из-за неадэкватности модели свободных электронов.

Несмотря на очень большую величину ω_p для электронного газа в металле, в нем могут распространяться волны разных типов с малой скоростью и низкой частотой, если присутствует также и магнитное поле. Поскольку эти волны имеют круговую поляризацию, их называют *геликонными волнами* Эг-

рена [35] *. Геликонные волны интенсивно изучаются как в металлах, так и в полупроводниках. В полупроводниках или полуметаллах, в которых концентрации электронов и дырок равны, вместо геликонной волны возникают две моды, распространяющиеся в магнитном поле: быстрая и медленная волны Альфвена [36].

3.4. Зонная теория твердых тел

Движение электрона в самосогласованном поле

Как уже было сказано, квантовая теория свободных электронов довольно успешно объясняет многие свойства металлов. Трудности возникают, когда нужно делать предположения относительно процесса рассеяния. В традиционной теории свободных электронов средняя длина свободного пробега λ или время τ_m используются при вычислении электропроводности; при этом предполагается, что на каждом ионе может происходить упругое рассеяние. Однако фактические *значения* электропроводности несовместимы с существованием этого процесса (не говоря уже о неправильной температурной зависимости, предсказываемой теорией, основанной на предположении об упругом рассеянии). Хотя мы и можем ввести в модель Зоммерфельда процессы неупругого рассеяния (на фононах и на дефектах), сначала необходимо объяснить, каким образом электроны, находящиеся в определенных состояниях, могут иметь бесконечную длину свободного пробега в отсутствие этих неупругих процессов. Хотелось бы также объяснить, почему одни материалы являются металлами, другие — хорошими изоляторами, а некоторые обладают промежуточной электропроводностью (полупроводники). Рис. 3.21 демонстрирует огромный диапазон изменения значений электропроводности при комнатной температуре у этих трех типов твердых тел. Известно, что отсутствие проводимости не обязательно обусловлено малой величиной подвижности электронов, так как многие диэлектрики являются чувствительными фотопроводниками.

---

[35] *Aigrain P.* Proceedings 5th International Semiconductor Conference, Czechoslovak Academy ox Sciences, 1961, p. 224.

* Существование геликонов впервые было показано Константиновым и Перелем в 1960 г. [*Константинов О. В., Перель В. И.*— ЖЭТФ. 1960, т. 38, с. 161].— *Прим. перев.*

[36] Геликонные волны и волны Альфвена обсуждаются Гликсманом [34]. См. также: *Platzman P. M., Wolff P. A.* Waves and Interactions, in Solid State Plasmas, Academic Press, 1973; *Steele M. C., Vural B.* Wave Interactions in Solid State Plasmas, McGraw-Hill, 1969.

Рис. 3.21. Диапазон изменения электропроводности, измеряемой при комнатной температуре, для некоторых типичных представителей трех классов твердых тел.

В зонной теории твердых тел эти проблемы решают, приписывая в уравнении Шредингера электрону в кристалле отличную от нуля периодическую потенциальную энергию $\mathscr{V}(\mathbf{r})$. При этом используются следующие приближения.

1. Волновая функция ψ для собственного состояния, отвечающего определенной собственной энергии ε, вычисляется для *идеальной* периодической решетки. Рассеяние вводится позже как возмущение. Конечно, существуют ситуации, для которых пренебрежение тепловыми колебаниями решетки до последнего момента представляется довольно некорректной процедурой, однако, как мы увидим, практически такой метод обычно хорошо работает и действительно упрощает расчет.

2. Зонная теория строится как теория для одного электрона. Предполагается, что для любого электрона действие *всего остального кристалла* можно описать с помощью *эффективной потенциальной энергии* $\mathscr{V}(\mathbf{r})$. Величина $\mathscr{V}(\mathbf{r})$ может быть различной для электронов различных энергий.

3. Выбор одноэлектронной модели позволяет использовать одноэлектронное уравнение Шредингера

$$[(-\hbar^2/2m)\nabla^2 + \mathscr{V}^c(r)]\psi = \varepsilon\psi. \qquad (3.111)$$

Собственные состояния, получаемые путем решения этого уравнения, должны быть заполнены в соответствии с законом заполнения Ферми — Дирака (3.45).

Одноэлектронное приближение совместимо с использованием метода самосогласованного поля для нахождения распределения электронов в многоэлектронных атомах, как это было сделано Хартри (1928 г.) и Фоком (1930 г.). В приближении Хартри или Хартри — Фока динамическое взаимодействие электрона с другими электронами в кристалле заменяется средним значением энергии взаимодействия по всем занятым электронным состояниям. Используя методы самосогласованного поля для вычисления электронных спектров в твердом теле, можно (в принципе) исходить только из знания симметрии кристаллической решетки и выбора пробной функции $\mathscr{V}^c(\mathbf{r})$, совместимой с этой симметрией. Найдя решения уравнения Шредингера, можно после нескольких итераций приблизить $(\mathbf{r})$ к его истинному значению и получить такие наборы ψ и ε, которые могут рассматриваться как оптимизированные самосогласованные значения для расчета из первых принципов. Некоторые расчеты электронных зон в твердых телах выполнены этим методом[37]. Существуют, однако, другие теоретические школы, с точки зрения которых расчет упрощается и облегчается, если в него с самого начала заложены и использованы как заданные параметры теории некоторые числа, взятые из эксперимента.

Корреляция между движением одного электрона и движением всех прочих свободных электронов в зонной теории грубо отражена путем учета принципа Паули, заключенного в статистике Ферми — Дирака. С помощью принципа Паули может быть обоснована невозможность электрон-электронных столкновений. Как было указано выше, методы задач многих тел, которые лучше учитывают электрон-электронную корреляцию, очень сложны, и закон распределения Ферми — Дирака в них теряет силу. Обойти эти трудности можно с помощью *метода*

---

[37] Красноречивая аргументация в пользу расчетов энергетических зон из первых принципов дана Германом [*Herman F.* Proceedings of 7th International Conferense, Academic Press, 1964]. Более поздняя точка зрения того же автора на сочетание экспериментальных данных с расчетными зонными параметрами изложена в работе: *Herman F. et al.* Proceedings of the Symposium on Energy Bands in Metals and Alloys, Los Angeles 1967, Gordon and Breach, 1968, p. 19.

элементарных возбуждений, который заменяет систему взаимодействующих электронов суммой энергий для набора «квазичастиц», уже *подчиняющихся* статистике Ферми — Дирака [38].

Функции Блоха

Пространственная зависимость потенциала, действие которого испытывает в кристалле внешний электрон, была рассмотрена Феликсом Блохом [39], сделавшим в результате своего анализа ряд существенных выводов. Полный потенциал $V(\mathbf{r})$ состоит из суммы двух частей.

1. Электростатического потенциала, обусловленного распределением атомных остатков. Для идеальной решетки (т. е. для решетки без фононов) этот вклад в $V(\mathbf{r})$ будет обладать трансляционной периодичностью решетки.

2. Потенциала, обусловленного всеми остальными внешними электронами. Блох предположил, что плотность заряда, создаваемая этими электронами, имеет одно и то же среднее по времени значение в каждой элементарной ячейке кристалла и, следовательно, также является периодической функцией. Такое предположение, разумеется, удовлетворяет требованию электронейтральности и грубо учитывает электрон-электронное отталкивание.

Таким образом, согласно Блоху, полная потенциальная энергия $\mathscr{V}(\mathbf{r}) = -eV(\mathbf{r})$, которая должна быть подставлена в уравнение Шредингера (3.111), обладает периодичностью решетки. Отсюда он заключил, что волновые функции, удовлетворяющие уравнению (3.111) с таким потенциалом, должны иметь вид

$$\psi_k(\mathbf{r}) = U_k(\mathbf{r}) \exp(i\mathbf{k}\cdot\mathbf{r}), \qquad (3.112)$$

где $U_k(\mathbf{r})$ — некоторая функция (зависящая от величины волнового вектора $\mathbf{k}$), *также* обладающая периодичностью решетки.

Уравнение (3.111) принадлежит семейству дифференциальных уравнений, известных в математике как уравнения Хилла. Из периодичности $\mathscr{V}(\mathbf{r})$ следует, что должны существовать решения типа (3.112). Это утверждение математики называют теоремой Флокэ. Однако, когда речь идет о поведении электронов в периодической решетке, теорему, приводящую к функциям вида (3.112), обычно называют теоремой Блоха, а сами функции — функциями Блоха. Они оказываются очень полезными.

Чтобы обсудить теорему Блоха для очень упрощенной гео-

[38] Заинтересованный (и хорошо подготовленный) читатель найдет ссылки на ряд методов задач многих тел в книге Андерсона: [Concepts in Solids, ed. P. W. Andersson, Benjamin, 1964].

[39] *Bloch F.*— Z. Physik, 52, 555 (1928).

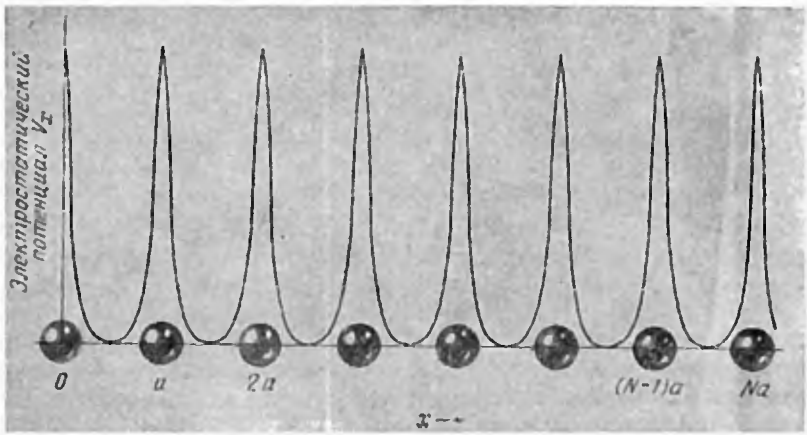


Рис. 3.22. Периодический потенциал линейной моноатомной решетки. Можно предположить, что через некоторое произвольное число N элементарных ячеек волновые функции *всех* собственных состояний будут повторяться.

метрии, представим линейную цепочку идентичных атомов с межатомным расстоянием a. Атомы отдают свои электроны, которые образуют вырожденный одномерный электронный газ, а положительные атомные остатки создают электрический потенциал, представляющий собой ряд периодических максимумов, как показано на рис. 3.22. [Следовательно, потенциальная *энергия* электрона $\mathscr{V}(x) = -eV(x)$ имеет ряд последовательных периодических минимумов.] Поскольку потенциал обладает периодичностью решетки, для любого значения x

$$V(x) = V(x+a). \tag{3.113}$$

Решение одномерного уравнения (3.111) для потенциальной энергии $\mathscr{V}(x)$ существует. Наложим теперь циклическое граничное условие типа Борна — Кармана, состоящее в том, что волновая функция периодически повторяется через каждые N единичных ячеек, где N — большое произвольно выбранное число. Тогда

$$\psi(x) = \psi(x+Na) \tag{3.114}$$

для любого значения x. Ввиду трансляционной симметрии нашей модели представляется вероятным, что волновая функция в одной единичной ячейке связана с волновой функцией в следующей ячейке соотношением типа

$$\psi(x+a) = J\psi(x), \tag{3.115}$$

где J — функция, которую следует определить. Отсюда вытекает, что для точек, разделенных несколькими единичными ячейками, выполняется соотношение

$$\psi(x+ma)=J^m\psi(x) \tag{3.116}$$

и, в частности, что

$$\psi(x+Na)=J^N\psi(x) \tag{3.117}$$

Сравнение выражений (3.114) и (3.117) показывает, что J должно быть одним из N корней N-й степени из единицы:

$$J=\exp(2\pi iM/N), \tag{3.118}$$

где M — целое число.

Очевидно, все наши требования, наложенные на волновую функцию, будут удовлетворяться, если записать $\psi(x)$ в виде произведения двух функций

$$\psi(x)=U_M(x)\exp(2\pi iMx/Na), \tag{3.119}$$

где первый множитель $U_M(x)$ есть функция, которая, подобно $V(x)$, обладает трансляционной периодичностью решетки.

Теперь каждое конкретное значение M может быть сопоставлено с конкретным значением волнового вектора $\mathbf{k}$, определяющего периодичность волновой функции. Поскольку второй множитель в выражении (3.119) на расстоянии Na проходит M полных периодов, мы можем связать k и M соотношением

$$k=2\pi M/Na. \tag{3.120}$$

Воспользовавшись этим определением в формуле (3.119), получаем волновую функцию

$$\psi(x)=U_k(x)\exp(ikx), \tag{3.121}$$

представляющую собой результат Блоха для одномерного случая. Справедливость этого результата может быть показана для двух и трех измерений.

В формуле (3.119) нет необходимости использовать значение M, превышающее $\pm(N/2)$, что соответствует $\mathbf{k}=\pm(\pi/a)$. Любой дополнительный периодический множитель может быть отнесен к амплитуде $U_k(x)$ блоховской функции. Речь идет о том, что интервал изменения волнового вектора от $k=(-\pi/a)$ до $k=(+\pi/a)$ [что соответствует изменению M от $(-N/2)$ до $(+N/2)$] образует первую зону Бриллюэна, как это было показано в гл. 1 и 2. Бóльшие значения M (т. е. бóльшие значения k) соответствуют точкам в зонах Бриллюэна, расположенных дальше от начала координат в $\mathbf{k}$-пространстве.

Выбор наименьшего возможного значения для $\mathbf{k}$ означает, что волновой функции любого электронного состояния соот-

ветствует точка в первой зоне Бриллюэна **k**-пространства с любым числом измерений: одномерного, двумерного и трехмерного. Как и в случае для колебательных состояний решетки, изменение волнового вектора на вектор обратной решетки

$$\mathbf{k}' = \mathbf{k} \pm \mathbf{G} \tag{3.122}$$

приводит с точки зрения наблюдаемого электронного движения к состоянию, эквивалентному **k**. Применяя преобразование (3.122) к волновой функции (3.112), получаем

$$\psi_{k'}(\mathbf{r}) = [U_k(r) \exp(i\mathbf{G}\cdot\mathbf{r})] \exp(i\mathbf{k}'\cdot\mathbf{r}). \tag{3.123}$$

Рассмотрим множитель $\exp(i\mathbf{G}\cdot\mathbf{r})$, который теперь появился в блоховской функции, и покажем, что он обладает периодичностью решетки. При трансляции в реальном пространстве в точку

$$\mathbf{r}' = \mathbf{r} + \mathbf{T} \tag{3.124}$$

выполняется равенство

$$\exp(i\mathbf{G}\cdot\mathbf{r}') = \exp(i\mathbf{G}\cdot\mathbf{r}) \exp(i\mathbf{G}\cdot\mathbf{T}). \tag{3.125}$$

Если **G** и **T** определяются формулами (1.43) и (1.42), то величина $\exp(i\mathbf{G}\cdot\mathbf{T})$ должна быть равна единице. Следовательно, величина $\exp(i\mathbf{G}\cdot\mathbf{r})$ и все выражение в квадратных скобках в (3.123) обладают периодичностью решетки, как и сама величина $U_k(\mathbf{r})$. Поэтому при желании достаточно рассматривать значения **k**, лежащие в первой зоне Бриллюэна.

Для колебаний решетки в одноатомном твердом теле было установлено, что данному волновому вектору соответствует только одно решение (одна частота, одна энергия фонона). Однако в твердом теле с многоатомным базисом данному волновому вектору соответствуют две возможные энергии фонона, принадлежащие оптической и акустической ветви. Для электронных состояний в зонной теории ситуация еще более сложна. Уравнение (3.111) имеет для данного значения **k** бесконечный ряд блоховских решений с последовательно возрастающими энергиями.

Говорят, что решения зонной теории записаны в представлении *приведенных зон*, если энергия электрона задана как многозначная функция **k** в интервале волновых векторов, ограниченном первой зоной Бриллюэна. Такое представление часто используется. Однако иногда полезно использовать представление *расширенных зон*, в котором энергия является однозначной функцией волнового вектора на протяжении интервала, в два или более раз превышающего ширину зоны. Последнее представление обычно используют, когда исследуется зависи-

мость энергии от $\mathbf{k}$ при переходе от одной зоны Бриллюэна к другой и при появлении разрывов энергии на границе зоны, обусловленных конечным периодическим потенциалом.

Разрешенные и запрещенные энергетические зоны

В модели свободных электронов можно было говорить об энергии электрона и его волновом векторе, не рассматривая конкретной кристаллографической структуры твердого тела. Как только мы начинаем учитывать периодический потенциал, становятся важными различные поверхности зоны Бриллюэна в $\mathbf{k}$-пространстве. Особенно важны они в том случае, когда поверхность Ферми в $\mathbf{k}$-пространстве, вычисленная в модели свободных электронов, пересекает границы зоны Бриллюэна кристаллической решетки.

В теории свободных электронов связь между энергией электрона и его импульсом очень проста. Поскольку потенциальная энергия предполагается всюду равной нулю, можно записать

$$\varepsilon = p^2/2m. \tag{3.126}$$

Волна де Бройля, описывающая такой электрон, обладает волновым вектором, связанным с энергией дисперсионным соотношением

$$\varepsilon = \hbar^2 k^2/2m, \tag{3.127}$$

уже приводившемся раньше [см. (3.40)].

Поскольку уравнение (3.111), решения которого должны иметь вид (3.112), было введено в качестве основы зонной теории, имеет смысл установить асимптотическую связь между зонной теорией и теорией свободных электронов. В асимптотическом пределе свободных электронов в случае бесконечно малой периодической потенциальной энергии в уравнении (3.111) блоховские решения представляют собой плоские волны [$\psi \sim \exp(i\mathbf{k}\cdot\mathbf{r})$], причем волновой вектор и энергия связаны соотношением (3.127). В пределе свободных электронов разрешено любое значение энергии, отвечающее некоторому вещественному значению $\mathbf{k}$.

На рис. 3.23 представлена зависимость $\varepsilon(\mathbf{k})$ (3.127) для одномерного твердого тела с постоянной решетки a и бесконечно малым периодическим потенциалом. На рисунке решение показано как в представлении *расширенных зон*, так и в представлении *приведенных зон*. Во втором случае участки решения $\varepsilon(\mathbf{k})$ сдвинуты по горизонтали на вектор обратной решетки, чтобы используемые значения $\mathbf{k}$ лежали в первой зоне Бриллюэна. [Для одномерного твердого тела с постоянной решетки a лю-

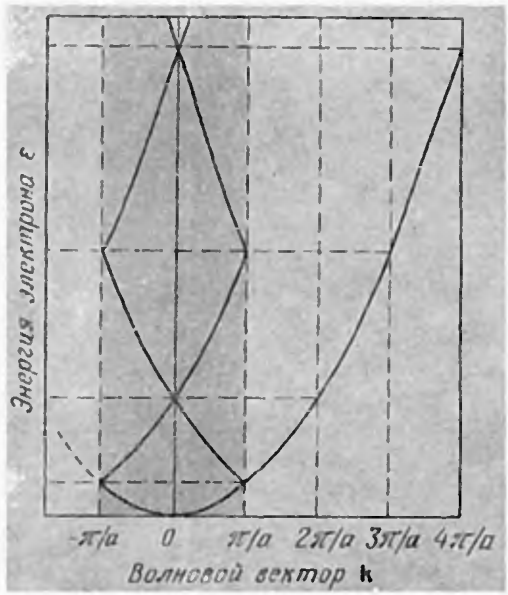


Рис. 3.23. Энергия электрона как функция волнового вектора в одномерном кристалле с постоянной решетки a при равной нулю амплитуде периодического потенциала. В этом случае энергия является непрерывной функцией волнового вектора k. Энергия показана как многозначная функция волнового вектора в выделенной темным цветом первой зоне Бриллюэна (представление приведенной зоны).

бой вектор обратной решетки есть целое кратное величины $(2\pi/a)$.]

Такая приятная простота утрачивается, как только мы переходим к собственным функциям уравнения (3.111), их волновым векторам и разрешенным собственным значениям (энергиям) в случае конечной периодической потенциальной энергии $\mathscr{V}(\mathbf{r})$. Оказывается, что какой бы вид не имела периодическая функция $\mathscr{V}(\mathbf{r})$, одним энергиям могут соответствовать вещественные значения k, а с другими энергиями совместимы только мнимые значения k.

Интервал энергий, для которого блоховские решения (3.112) уравнения (3.111) соответствуют вещественным значениям k, называется *зоной разрешенных энергий* в твердом теле. Имеется ряд интервалов энергий, для которых существуют различные блоховские решения с вещественными k. Интервал энергий, расположенных выше одной разрешенной зоны и ниже другой, называется *запрещенной зоной* (щелью) материала.

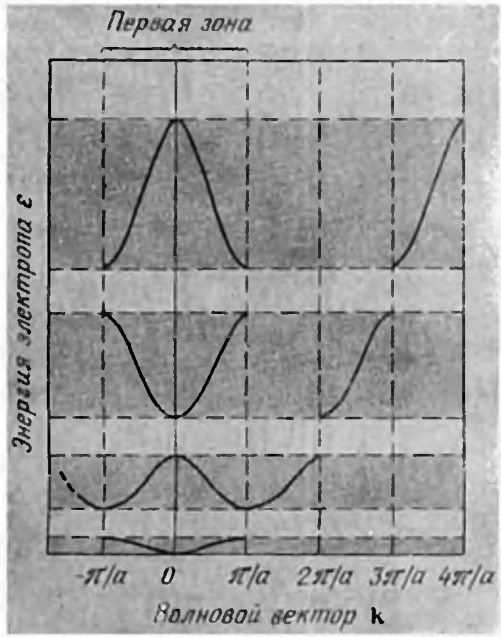


Рис. 3.24. Представление расширенных зон и приведенных зон для энергии электрона как функции волнового вектора в одномерном твердом теле с постоянной решетки a при конечной амплитуде периодического потенциала. Делокализованные ограниченные волновые функции (т. е. волновые функции, соответствующие вещественным $\mathbf{k}$) могут быть построены только внутри темных областей собственных энергий. Это области представляют собой разрешенные электронные зоны, разделенные между собой запрещенными зонами (энергетическими щелями).

На рис. 3.24 схематически изображена предсказываемая картина разрешенных и запрещенных зон для гипотетического одномерного твердого тела с постоянной решетки a. Различие между рис. 3.23 и 3.24 состоит в том, что периодический потенциал предполагается бесконечно малым в первом случае и конечным во втором.

Когда волновой вектор $\mathbf{k}$ веществен (что имеет место в энергетических интервалах, которые были названы разрешенными зонами), функция Блоха (3.112) представляет собой бегущую волну, модулированную с периодичностью решетки. Волна Блоха распространяется по кристаллу без затухания, а средняя по времени плотность электрического заряда $-e|\psi|^2$ имеет одну и ту же величину в каждой элементарной ячейке кристалла.

«Зонный» электрон в состоянии, описываемом функцией Блоха с вещественным $\mathbf{k}$, движется по кристаллу сколь угодно долго; волновая функция не затухает. Следовательно, пока сохраняется идеальная периодичность, зонный электрон обладает бесконечной средней длиной свободного пробега. Таким образом, квантовая механика оказывается в состоянии объяснить полное отсутствие упругого рассеяния в системе положительных ионных остатков. Свободное движение электрона прекращается, когда идеальная периодичность нарушается. При любом таком нарушении функция Блоха уже не удовлетворяет уравнению, и электрон испытывает рассеяние, обычно приводящее к изменению направления движения. Такие нарушения периодичности могут быть обусловлены поверхностью кристалла, статическими точечными дефектами и фононами; из-за них средняя длина свободного пробега становится конечной.

Как можно видеть из (3.112), функции Блоха с мнимым значением $\mathbf{k}$ ведут себя неадекватно. При подстановке мнимого значения $\mathbf{k}$ волновая функция неограниченно возрастает для одного из направлений реального пространства. Это означает, что электрон, энергия которого удовлетворяет уравнению (3.111) с мнимым значением $\mathbf{k}$, не может свободно перемещаться по кристаллу. Следовательно, области энергий электрона, для которых значение $\mathbf{k}$ должно быть мнимым (при точном потенциале) являются запрещенными областями для подвижных электронов в твердом теле. Энергетическая щель в спектре электронных энергий в твердом теле по своей природе соответствует щели между акустической и оптической ветвями колебаний решетки.

Единственный тип волновой функции, который может существовать для энергий, лежащих внутри запрещенного интервала,— это волновая функция, локализованная около какого-то дефекта или у поверхности кристалла. При наличии постороннего атома, вакансии решетки или другого разрыва идеальной периодичности может существовать волновая функция *связанного состояния*, сосредоточенная около точки $\mathbf{r}_0$, в которой расположен дефект,

$$\psi = A \exp[-ik|\mathbf{r}-\mathbf{r}_0|]. \qquad (3.128)$$

Поскольку k — мнимая величина, волновая функция (3.128) монотонно спадает при удалении от точки $\mathbf{r}_0$ в любом направлении, что указывает на локализацию плотности заряда $-e|\psi|^2$. Электрон может быть захвачен локальным электрическим полем дефекта; при этом он будет обладать энергией, расположенной в одной из запрещенных щелей, т. е. энергией, запрещенной для свободно движущихся электронов. Таких локализованных состояний, обусловленных примесями или разупоря-

дочением, гораздо меньше, чем зонных состояний, которые создаются регулярно расположенными атомами. Поэтому в металле влиянием на электроны локализованных состояний часто можно пренебречь. Однако роль локализованных состояний дефектов обычно очень важна в диэлектриках и полупроводниках.

И. Е. Тамм [40] впервые указал на то, что разрыв периодичности решетки на поверхности кристалла приводит к появлению локализованных поверхностных состояний, соответствующих блоховским функциям с мнимыми значениями k и с энергией, лежащей в одной из «запрещенных зон». Интенсивное экспериментальное изучение поверхностей подтвердило, что таммовские поверхностные состояния действительно существуют, однако загрязнение поверхности кристалла (например, адсорбированным кислородом), по-видимому, создает много дополнительных поверхностных состояний. Большая часть литературных данных относится к поверхностным состояниям в полупроводниках [41], а не в металлах, и обсуждать эту тему более подробно мы не будем.

Модель Кронига—Пенни

Для изучения реальных энергетических зон и запрещенных зон в реальных трехмерных твердых телах [для которых точные решения уравнения (3.111) часто невозможно получить] очень много может дать рассмотрение одномерных моделей. В одной из таких моделей, предложенной Кронигом и Пенни [42], используются простые прямоугольные потенциальные барьеры, изображенные на рис. 3.25. Тогда вид зависимости энергии от волнового вектора определяется шириной b и высотой $\mathscr{V}_0$ этих барьеров.

В пределе бесконечно малого произведения $b\mathscr{V}_0$ электроны любой энергии движутся совершенно свободно и обладают тем же законом дисперсии (3.127), что и свободные электроны. В противоположном предельном случае столь широких барьеров, что для электронов с энергией, меньшей $\mathscr{V}_0$, вероятность туннелирования из одного атомного узла в соседний равна нулю, оказываются возможны только значения энергии

$$\varepsilon_j = \pi^2\hbar^2 j^2/2ma^2, \qquad j = 1, 2, 3, \ldots, \tag{3.129}$$

пока ε_j остается меньше $\mathscr{V}_0$. Формула (3.129) представляет собой стандартный квантовомеханический результат для энергии частицы в одномерном ящике ширины a.

[40] *Tamm I.*— J. Physik, **76**, 849 (1932). См.: *Davison S. G., Levine J. D.* Solid State Physics, v. 25, Academic Press, 1970, p. 1.

[41] *Frankl D. R.* Electrical Properties of Semiconductor Surfaces, Pergamon, 1967.

[42] *Kronig R. de L., Penney W. G.*— Proc. Roy. Soc., **A130**, 499 (1930).

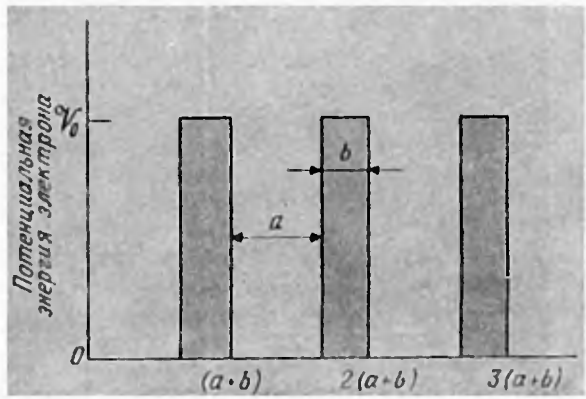


Рис. 3.25. Одномерная периодическая потенциальная энергия электрона, принятая в модели Кронига — Пенни.

Перейдем теперь к случаю больших $\mathscr{V}_0$, однако b, достаточно малых, чтобы существовала конечная вероятность того, что электрон с энергией, меньшей $\mathscr{V}_0$, протуннелирует из одной ячейки в следующую. Крониг и Пенни проанализировали эту промежуточную ситуацию, потребовав, чтобы волновые функции в темных и светлых областях на рис. 3.25 совпадали на границах областей. Подробное решение этой задачи приводится во многих элементарных учебниках квантовой механики. Особенно простую форму оно принимает в предельном случае очень высоких, но очень узких барьеров, когда произведение $b\mathscr{V}_0$ остается конечным. Тогда волновой вектор k и энергия ε связаны соотношением

$$\cos(ka) = \cos(\alpha a) + P\left[\frac{\sin(\alpha a)}{\alpha a}\right], \tag{3.130}$$

где α выражается через энергию следующим образом:

$$\alpha = [2m\varepsilon/\hbar^2]^{1/2}, \tag{3.131}$$

а величина P есть мера непроницаемости барьеров:

$$P = \lim_{\substack{\mathscr{V}_0 \to \infty \\ b \to 0}} [(b\mathscr{V}_0)\, ma/\hbar^2]. \tag{3.132}$$

В пределе $P \to 0$ получаем $\alpha = k$, т. е. дисперсионное соотношение (3.127) для свободных электронов. В пределе $P \to \infty$ получаем набор энергий, даваемых формулой (3.129). Однако решение уравнения (3.130) для конечного P приводит к чередованию областей энергии с вещественным и мнимым k. Величина k должна быть мнимой, когда абсолютная величина

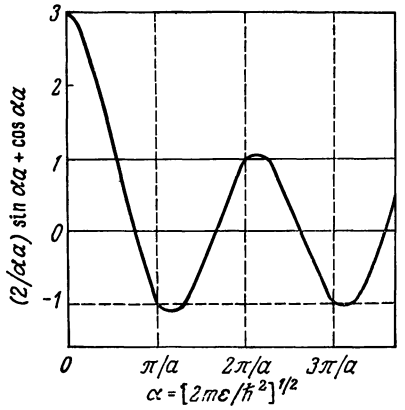


Рис. 3.26. Правая часть уравнения (3.130) для параметра $P=2$, построенная как функция величины α, зависящей в модели Кронига — Пенни от энергии электрона.

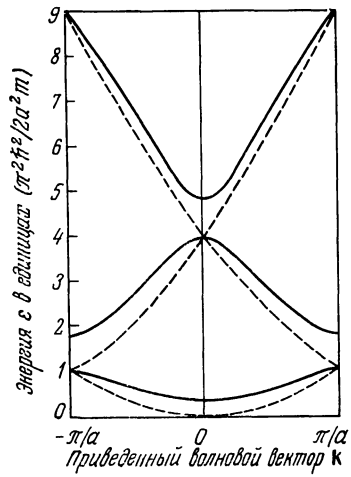


Рис. 3.27. Представление приведенных зон для энергии как функции волнового вектора в модели Кронига — Пенни при $P=2$ (как на рис. 3.26). Соответствующие кривые для $P=0$ (свободные электроны) показаны штриховыми линиями.

правой части уравнения (3.130) больше единицы, и вещественной, когда она меньше единицы. На рис. 3.26 изображена правая часть уравнения (3.130) как функция α для конечного значения параметра P. Заметим, что для $\alpha=0$ ордината равна $(1+P)$, а когда α есть кратное величины (π/a), ордината равна ± 1 независимо от значения P. *Зависящей* от P оказывается ширина интервалов α, для которых k должно быть мнимым, т. е. ширина *запрещенных зон* в нашей модели.

Чередование разрешенных и запрещенных зон показано на рис. 3.27, изображающем зависимость энергии от *приведенного* волнового вектора (представление приведенных зон) для случая $P=2$ (для которого построена зависимость на рис. 3.26). Для сравнения штриховыми кривыми изображена та же зависимость для совершенно свободных электронов ($P=0$). Верхний энергетический предел для каждой разрешенной зоны в этой модели является одним из членов последовательности (3.129), однако нижний предел зависит от заданной величины P. По мере увеличения P каждая зона стягивается во все более узкий интервал энергий, а ширина запрещенной зоны увеличивается.

Модель Кронига — Пенни чрезвычайно идеализирована, поскольку в ней рассматривается одномерный аналог твердого тела и используется искусственно простой потенциал. Тем не менее ее результаты очень поучительны. Они указывают на *принципиальные* стороны более сложных и реалистичных моделей. Одна из особенностей зон, изображенных на рис. 3.27, состоит в следующем: кривизна зависимости $\varepsilon(\mathbf{k})$ у дна и у потолка каждой зоны такова, что $(d^2\varepsilon/dk^2) \neq 0$, а $(d\varepsilon/dk) = 0$. Эта особенность присуща всем более реалистическим моделям.

Переход от изолированных атомов к предельному случаю свободных электронов

Кинетическая энергия электрона в делокализованном состоянии Блоха может быть больше или меньше, чем $(\hbar^2k^2/2m)$ в зависимости от типа атомов, составляющих кристалл, расстояний между ними и симметрии структуры. Множество свойств твердых тел можно объяснить отклонением от «обычной» кинематики (3.126). На первый взгляд это обстоятельство может вызвать недоумение, однако, будучи понятым, оно оказывается чрезвычайно полезным.

Влияние периодического потенциала зависит как от величины этого потенциала, так и от возможности взаимодействия между атомами, которая изменяется с изменением межатомного расстояния. Кривые на рис. 3.28 соответствуют различной величине межатомного взаимодействия: от нуля слева до очень сильного справа. Когда атомы находятся достаточно далеко друг от друга, между ними нет взаимодействия и электроны могут находиться только в ограниченном числе дискретных атомных состояний. Если межатомное расстояние допускает некоторое (но не слишком большое) перекрытие атомных волновых функций, зоны могут быть описаны моделью *сильной связи*, на которой мы остановимся подробнее ниже, рассматривая зоны в реальном трехмерном твердом теле. Приближение сильной связи было использовано Блохом в его первоначальном исследовании энергетических зон в твердом теле. Это приближение называется также методом ЛКАО, поскольку волновые функции электронов в относительно узких зонах модели сильной связи могут быть аппроксимированы Линейной Комбинацией Атомных Орбиталей.

Модель *почти свободных электронов* для электронных состояний в твердом теле применима, когда периодический потенциал сравнительно мал (и/или межатомное расстояние достаточно мало, чтобы обеспечить значительное перекрытие волновых функций соседних атомов). Отклонения от дисперсионного соотношения для свободных электронов (3.127) заметны только

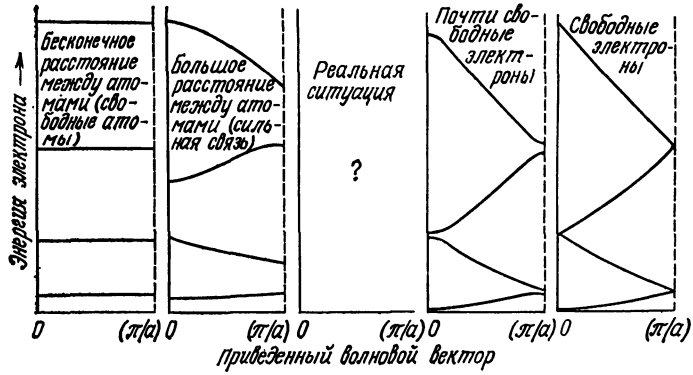


Рис. 3.28. Одномерное изображение зависимости энергии от волнового вектора для электронных состояний при изменении взаимодействия между соседними атомами от нуля (атомы бесконечно удалены друг от друга) до полного снятия межатомных барьеров.

в областях k-пространства вблизи границ зоны Бриллюэна. При сравнении рис. 3.28 и рис. 3.24 становится ясно, что второй соответствует случаю «почти свободных» электронов: для областей, расположенных достаточно глубоко внутри зон, дисперсионное соотношение $\varepsilon(\mathbf{k})$ остается практически невозмущенным.

В следующем разделе мы обсудим применимость модели почти свободных электронов на примере алюминия. Однако у большинства твердых тел эффективный межатомный потенциал, определяющий форму зон, лежит между значениями, отвечающими случаю сильной связи и случаю почти свободных электронов. Поэтому расчеты зонной структуры реальных твердых тел очень сложны, на что и указывает вопросительный знак в средней части рис. 3.28.

Однако независимо от того, является ли периодический потенциал большим или малым, можно наглядно представить себе, что зоны и запрещенные зоны между ними образуются в результате уширения дискретных атомных уровней, как показано на рис. 3.29. Система атомов, достаточно удаленных друг от друга, обладает большим числом состояний с одной и той же энергией. Если межатомное расстояние монотонно уменьшать до его равновесного значения в твердом теле, то вследствие взаимодействия между атомами каждый атомный уровень размазывается по интервалу энергий. Каждая «зона» вмещает столько электронных состояний, сколько их было в первоначальном атомном уровне. Как мы увидим дальше (и исследуем в задачах 3.23 и 3.24), для элементар-

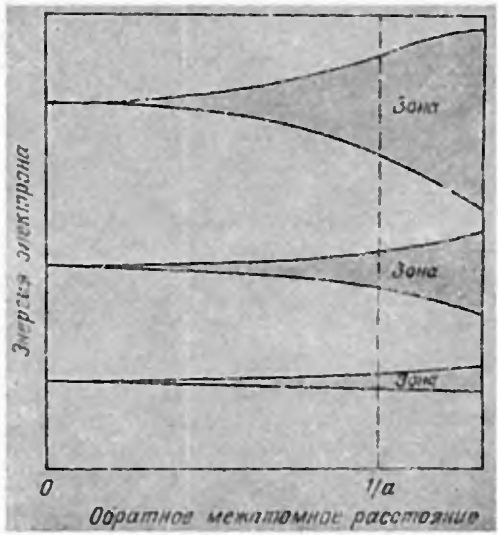


Рис. 3.29. Уширение каждого атомного энергетического уровня в зону конечной ширины по мере того, как удаленные атомы сближаются, образуя твердое тело. Уширение обычно менее выражено для глубоко лежащих уровней, соответствующих внутренним электронным оболочкам.

ного твердого тела, состоящего из N атомов, в объеме k-пространства, занимаемом каждой зоной Бриллюэна, может разместиться $2N$ электронов.

Отражение Брэгга и энергетическая щель

Рассмотрим электронное состояние, для которого k находится на границе зоны, так что уровень энергии оказывается непосредственно над одной из щелей или под ней. Для одномерной ситуации, изображенной на рис. 3.24, 3.27 и 3.28, это означает состояние, для которого $k = \pm\pi/a$. Вспомним, что для колебательных мод решетки на самой вершине акустической и на дне оптической ветви групповая скорость волн становится равной нулю; такие моды описываются стоячими волнами. В двухатомной решетке один вид атомов участвует в колебаниях той моды, которая лежит непосредственно под щелью, а другой — в колебаниях с самой низкой энергией оптической моды.

Аналогичная ситуация имеет место для электронных состояний у края щели. Функция Блоха для электронного состоя-

ния, лежащего у края зоны, представляет собой не бегущую, а стоячую волну, потому что электрон с таким волновым вектором, движущийся (в реальном пространстве) с компонентой импульса, перпендикулярной границе зоны (в k-пространстве), испытывает дифракцию Брэгга в соответствии с условием

$$2\mathbf{k}\cdot\mathbf{G}+G^2=0 \qquad (3.133)$$

[оно уже приводилось ранее, см. (1.54)]. Это условие выглядит сложным, будучи сформулировано для многомерного случая, однако для простого одномерного приближения его смысл ясен:

$$k=\pm G/2=\pm\pi/a. \qquad (3.134)$$

Для такого значения волнового вектора сдвиг фаз между волнами, отраженными от двух соседних атомов, равен π. Следовательно, решение для этого значения k содержит две равные компоненты, представляющие собой волну, распространяющуюся влево $k=(-\pi/a)$ и вправо $k=(+\pi/a)$. Это и есть стоячая волна.

Волны, распространяющиеся в противоположных направлениях, могут складываться двумя способами. Один из них представляет собой четную комбинацию:

$$\psi_e(x)=U_e(x)\,[\exp(i\pi x/a)+\exp(-i\pi x/a)]=2U_e(x)\cos(\pi x/a), \qquad (3.135)$$

где $U_e(x)$ обладает периодичностью решетки, а косинус имеет период $2a$. Волновая функция $\psi_e(x)$ не изменяется при замене x на $-x$.

Другим способом стоячая волна может быть составлена из нечетной (антисимметричной) комбинации

$$\psi_0(x)=U_0(x)\,[\exp(i\pi x/a)-\exp(-i\pi x/a)]=2iU_0(x)\sin(\pi x/a), \qquad (3.136)$$

которая меняет знак при изменении знака x. Нет ничего страшного в том, что $\psi_0(x)$ оказывается мнимой величиной, поскольку плотность электрического заряда, которая связана с волновой функцией, равна $-e|\psi|^2$ и в обоих случаях она представляет собой вещественную отрицательную величину. Однако величина $-e|\psi_e|^2$ проходит через ряд максимумов для x, кратных a, в точках, где расположены атомные остатки. Плотность заряда, связанная с $\psi_0(x)$, обращается в этих точках в нуль, поскольку при этих значениях x синус обращается в нуль.

Таким образом, электрон с волновым вектором $k=\pm\pi/a$ может быть описан *либо* волновой функцией, отвечающей состоянию, в котором он большую часть времени проводит вблизи атомных остатков, *либо* волновой функцией такого состояния, в котором электрон должен оставаться в пространстве между

атомами. Поскольку в каждой точке, в которой находится атом, электростатический потенциал велик и положителен, и поскольку электрон обладает отрицательным зарядом, энергия состояния с волновой функцией $\psi_e(x)$ меньше энергии состояния $\psi_0(x)$. Наличие периодического потенциала привело к появлению энергетической щели, которой не существует в теории свободных электронов.

Для одномерного случая n-ю зону свободных электронов в приведенной зоне Бриллюэна получаем, взяв часть параболы (3.127) для свободных электронов с центром в начале координат и сместив ее на вектор обратной решетки, достаточно большой, чтобы обеспечить перенос из n-й зоны Бриллюэна в первую зону. Вообще, всякий раз, когда значение волнового вектора $\mathbf{k}$ свободного электрона пересекает границу зоны Бриллюэна, для конечного межатомного потенциала возникает щель.

Теперь рассмотрим применение этого принципа к трехмерным твердым телам.

Зоны Бриллюэна для многомерных твердых тел

Теперь необходимо рассмотреть зоны Бриллюэна еще раз для структур, имеющих два и три измерения. Двумерная зона Бриллюэна для косоугольной решетки была показана на рис. 1.52. По определению ее центр размещен в одном из узлов обратной решетки, и она ограничена набором бисекторов* векторов обратной решетки, выходящих из центра. Эта зона имеет ту же площадь, что и элементарная ячейка, определяемая базисными векторами обратной решетки. Обобщая, можно сказать, что для твердого тела, имеющего m измерений, объем единичной ячейки в обратном пространстве, как и объем каждой из зон Бриллюэна, равен $(2\pi)^m$, поделенному на объем элементарной ячейки в реальном пространстве. Разрешенную плотность электронных состояний в обратном пространстве можно определить, обратившись к рис. 3.7; это позволяет понять высказанное выше утверждение, что для элементарного твердого тела, содержащего N атомов, зона Бриллюэна вмещает $2N$ электронных состояний (множитель 2 возникает из-за спинового вырождения). Следует подчеркнуть, что не только *первая* зона Бриллюэна может вместить $2N$ состояний, но и

* Бисектором вектора автор называет геометрическое место точек, равноотстоящих от его концов. Таким образом, в одномерном пространстве — это точка, делящая вектор пополам, в двумерном — прямая линия, проведенная перпендикулярно вектору через его середину, а в трехмерном — плоскость, перпендикулярная вектору и проходящая через его середину.— *Прим. перев.*

любая зона более высокого порядка в представлении расширенных зон должна обладать (в обратном пространстве) таким же объемом. Это должно быть так, поскольку любая зона более высокого порядка путем трансляции на соответствующий вектор обратной решетки может быть переведена в приведенную зону.

Легко видеть, что все зоны одномерной решетки имеют одну и ту же длину в пространстве k_x. Из рис. 3.24 видно, что m-я зона простирается от $k_x=[-m\pi/a]$ до $[-(m-1)\pi/a]$ и затем от $k_x=[(m-1)\pi/a]$ до $[m\pi/a]$. Следовательно, каждая зона имеет полный объем (т. е. длину), равный $[2\pi/a]$. Ниже мы убедимся, что все зоны двух и трех измерений также имеют одинаковый объем, однако сначала необходимо выяснить, каким образом строятся зоны выше первого порядка. Правила построения были сформулированы Бриллюэном [43].

Если связанная с электроном волна де Бройля (или в нашем случае другая волна) распространяется по периодической решетке, причем ее волновой вектор, измеренный от начала координат в k-пространстве, заканчивается на бисекторе вектора обратной решетки, то зависимость $\varepsilon(\mathbf{k})$ имеет разрыв. Такие значения $\mathbf{k}$ удовлетворяют условию дифракции Брэгга (1.54) или (3.133). Бриллюэн замечает, что зоны вводятся в целью исключить разрывы в дисперсионном соотношении $\varepsilon(\mathbf{k})$ везде, *кроме* зонных границ.

Каждая зона должна быть ограничена перпендикулярными бисекторами векторов обратной решетки, исходящих из определенной ее точки, выбранной за начало координат k-пространства. Поскольку обратное пространство для идеальной решетки бесконечно, за начало координат может быть выбрана *любая* точка обратной решетки. Для двумерной решетки перпендикулярными бисекторами являются прямые линии, как было показано раньше (см. рис. 1.52). Для трехмерной решетки каждый бисектор вектора обратной решетки есть плоскость. Ни один бисектор не может проходить через внутреннюю часть зоны. Поэтому объем, соответствующий n-й зоне Бриллюэна,— это объем, заключенный между n-й и $(n-1)$-й такими поверхностями.

Границы зон для гипотетической двумерной квадратной решетки с постоянной решетки a определяются решениями уравнения

$$mk_x+nk_y=(\pi/a)(m^2+n^2), \qquad (3.137)$$

---

[43] *Brillouin L.*— J. Phys. Radium, 1, 377 (1930); *Brillouin L.* Wave Propagation in Periodic Structures, Dover, 2nd ed., 1953. [Имеется перевод: *Бриллюэн Л., Пароди М.* Распространение волн в периодических структурах.— М.: ИЛ, 1959.]

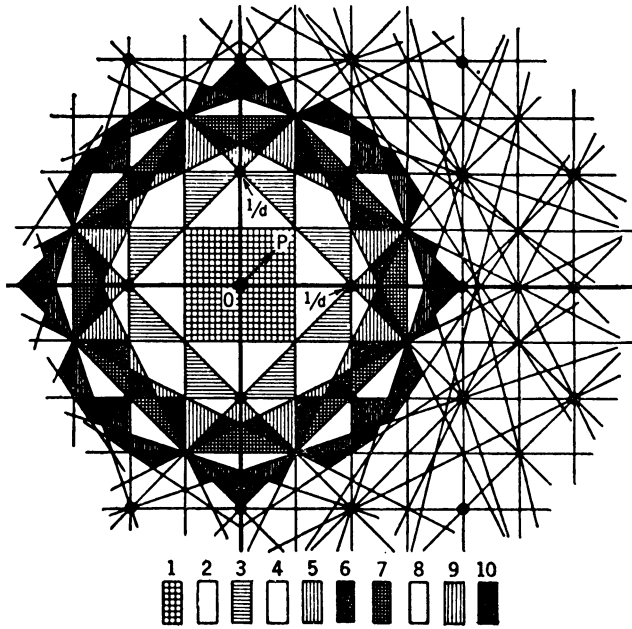


Рис. 3.30. Первые десять зон двумерной квадратной решетки (*Brillouin L.* Wave Propagation in Periodic Structures, Dover, 2nd ed., 1953. [Имеется перевод: *Бриллюэн Л., Пароди М.* Распространение волн в периодических структурах.— М.: ИЛ, 1959.])

где m и n — целые числа. Линии, удовлетворяющие этому уравнению, ограничивают первые десять зон для квадратной решетки, изображенных на рис. 3.30. Обратите внимание на разбиение **k**-пространств зонами высоких порядков. Тем не менее каждая такая зона может быть сведена по форме и объему к первой (приведенной) зоне путем трансляций на соответствующий вектор **G**. На рис. 3.31 показаны результаты составления такой «разрезной головоломки» для второй — девятой зон квадратной решетки. Отсюда видно, что каждая зона действительно имеет площадь $(2\pi/a)^2$.

Возможно теперь читатель захочет узнать, когда лучшим способом описания энергетических зон в твердом теле служит представление приведенных зон, а когда лучше использовать представление расширенных зон. На рис. 3.32 дано трехмерное изображение зависимости энергии от волнового вектора. На этой трехмерной диаграмме энергия отложена по вертикальной оси, а векторы k_x и k_y — в горизонтальной плоскости для первой зоны и для соответствующих областей второй, третьей и

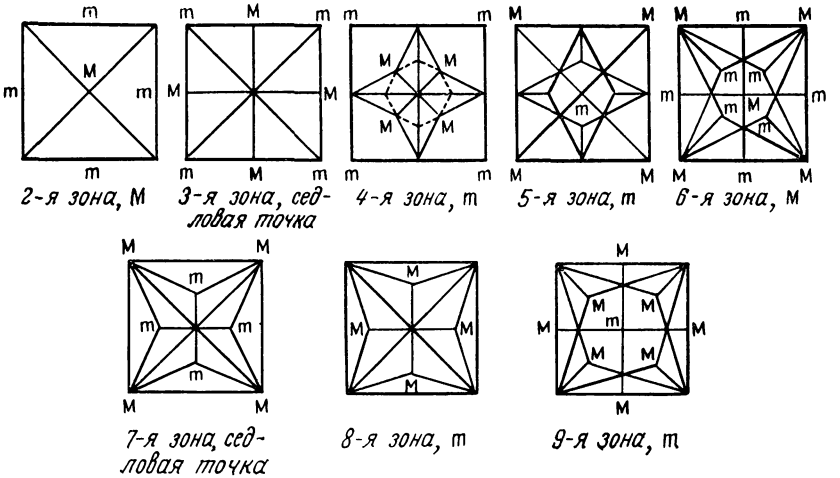


Рис. 3.31. Приведение зон со второй по девятую к первой зоне для двумерной квадратной решетки (*Brillouin L.* Wave Propagation in Periodic Structures, Dover, 2nd ed., 1953. [Имеется перевод: *Бриллюэн Л., Пароди М.* Распространение волн в периодических структурах.— М.: ИЛ, 1959.])

четвертой зон. Разрывы в энергии, соответствующие переходу **k** через зонную границу, изображены как вертикальные уступы. Соответствующая фигура для модели свободных электронов с пренебрежимо малым периодическим потенциалом будет характеризоваться таким же общим возрастанием ε с **k**, но не будет иметь разрывов.

Таким образом, метод расширенных зон полезен для передачи информации, содержащейся в рис. 3.32. Представление приведенных зон оказывается более полезным, когда мы хотим узнать последствия изменения волнового вектора электрона, находящегося в зоне, выше первой. В задаче 3.23 исследуется вид траектории в **k**-пространстве для четвертой зоны квадратной решетки.

Геометрические принципы, лежащие в основе построения зон, следующих за первой, для квадратной решетки, остаются в силе и для других двумерных и трехмерных решеток. На рис. 3.33 показаны первые пять зон для двумерной гексагональной решетки, а в задаче 3.24 предлагается привести зоны более высокого порядка к первой. В задаче 3.25 рассматриваются первые четыре зоны прямоугольной решетки.

В трехмерном случае неизбежно геометрическое усложнение топологии зонной поверхности, если рассматривается любая структура, более сложная, чем простая кубическая (для

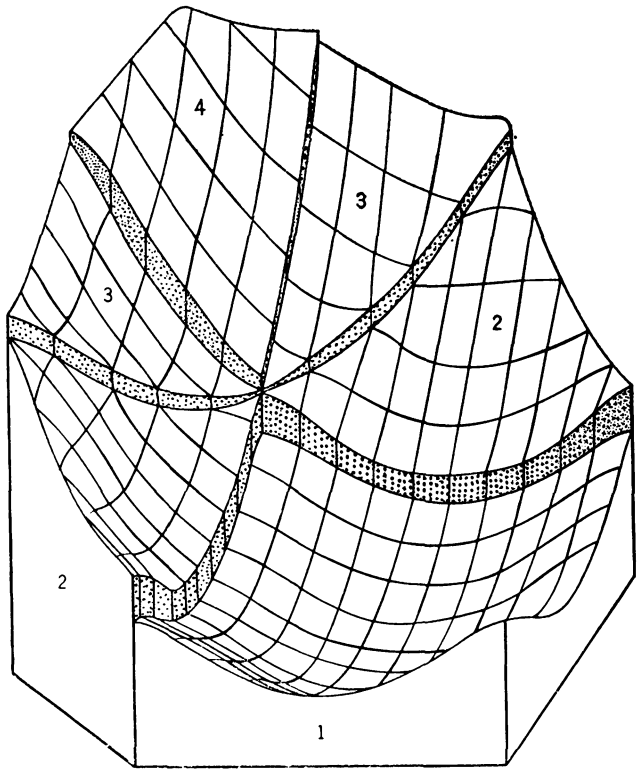


Рис. 3.32. Объемное изображение зависимости ε (k) для гипотетической двумерной квадратной решетки; в представлении расширенных зон. По оси ординат отложена энергия, k_x и k_y — оси в горизонтальной плоскости. Изображены части зон *1*, *2*, *3* и *4* для конечного периодического потенциала, создающего разрывы энергии конечной величины на зонных границах [*Slater J. C.* Quantum Theory of Molecules and Solids, v. 2., McGraw-Hill, 1965].

последней все будет точно так же, как в одно- и двумерном случае). На рис. 3.34 изображены первые четыре зоны для простой кубической решетки, а на рис. 3.35 и 3.36 показаны границы нескольких первых зон для о. ц. к.- и г. ц. к.-решеток. О. ц. к.-решетка имеет простую первую зону, однако внешняя поверхность второй зоны оказывается разбитой на фрагменты в k-пространстве в силу того, что области, принадлежащие соседним зонам, должны быть разделены векторами обратной решетки.

Для г. ц. к.-кристалла с ребром куба a обратная решетка обладает о.ц.к.-симметрией с ребром куба $(4\pi/a)$. Используя

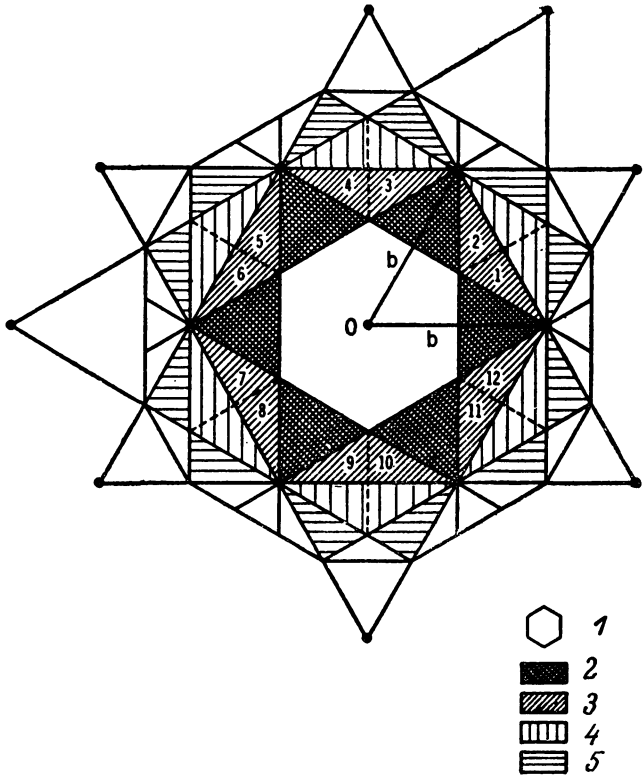
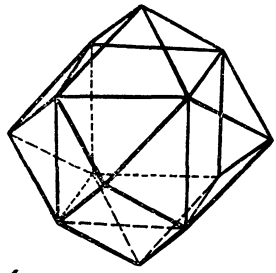
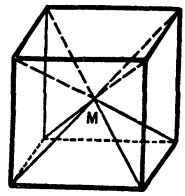


Рис. 3.33. Первые пять зон Бриллюэна для двумерной гексагональной решетки (*Brillouin L.* Wawe Propagation in Periodic Structures, Dover, 2nd ed., 1953. [Имеется перевод: *Бриллюэн Л., Пароди С.* Распространение волн в периодических структурах.— М.: ИЛ, 1959.])

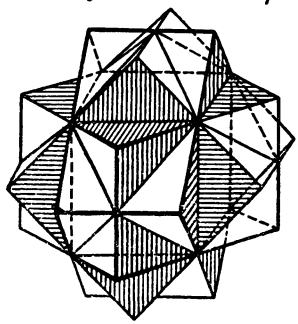
процедуру построения бисекторов для векторов обратной решетки, получаем, что первая зона имеет вид усеченного октаэдра с объемом $4(2\pi/a)^2$, т. е. в четыре раза большим, чем для простого кубического кристалла с таким же a. Дело здесь в том, что ребро куба в г. ц. к.-кристалле больше длины вектора элементарной трансляции для ромбоэдрической примитивной ячейки. Гранецентрированный куб содержит четыре атома, и зона Бриллюэна опять имеет объем, который может вместить два состояния с противоположными спинами на каждый атом. Это общее правило для элементарного твердого тела независимо от его структуры [см. задачу 3.26, где рассматривается гексагональная плотноупакованная (г. п. у.) решетка]. В химическом соединении имеется серия зон для каждого компонента.



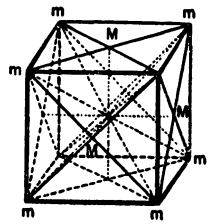
1-я зона – центральный куб
2-я зона – между центральным кубом и додекаэдром



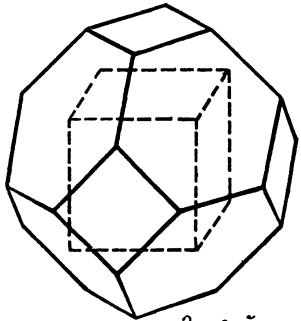
Приведение 2-й зоны



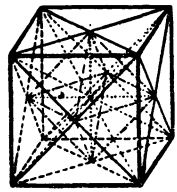
3-я зона – между додекаэдром и внешней поверхностью



Приведение 3-й зоны



4-я зона – между 3-й зоной и новой внешней поверхностью



Приведение 4-й зоны

Рис. 3.34. Первые четыре зоны Бриллюэна и их приведение к размеру и форме первой зоны для простой кубической решетки (*Brillouin L.* Wave Propagation in Periodic Structures, Dover, 2nd ed, 1953. [Имеется перевод: *Бриллюэн Л., Пароди М.* Распространение волн в периодических структурах.— М.: ИЛ, 1959.])

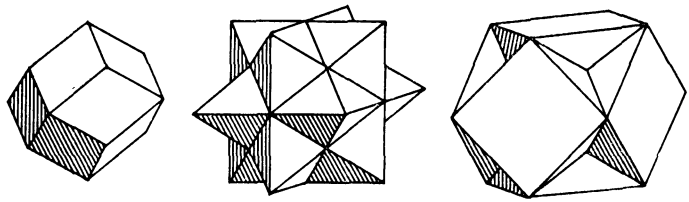


Рис. 3.35. Первые три зоны Бриллюэна для твердого тела, кристаллизующегося в о. ц. к.-структуре (и, следовательно, обладающего обратным пространством с г. ц. к.-симметрией). Грани первой зоны параллельны плоскостям (110) [*Slater J. C.*— Rev. Mod. Phys., **6**, 212 (1934)].

Поскольку имеется соответствие между дискретными состояниями изолированного атома и зонами Бриллюэна в твердом теле, можно полагать, что

1 зона происходит от 1*s*-состояний;
1 зона происходит от 2*s*-состояний;
3 зоны происходят от 2*p*-состояний;
1 зона происходит от 3*s*-состояний;
3 зоны происходят от 3*p*-состояний;
5 зон происходят от 3*d*-состояний и т. д.

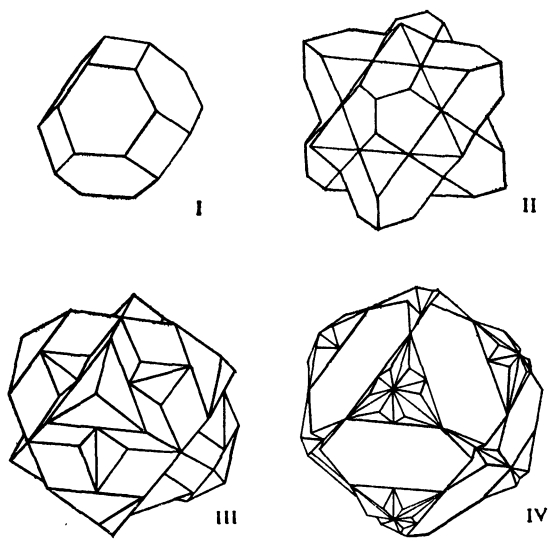


Рис. 3.36. Первые четыре зоны Бриллюэна для твердого тела, кристаллизующегося в г. ц. к.-структуре. В этом случае обратное пространство обладает о. ц. к.-симметрией. Зоны Бриллюэна такой же формы были получены для структур, производных от г. ц. к., включая структуру алмаза и цинковой обманки (*Harrison W A.* Solid State Theory, McGraw-Hill, 1970. [Имеется перевод: *Харрисон У.* Теория твердого тела.— М.: Мир, 1972.])

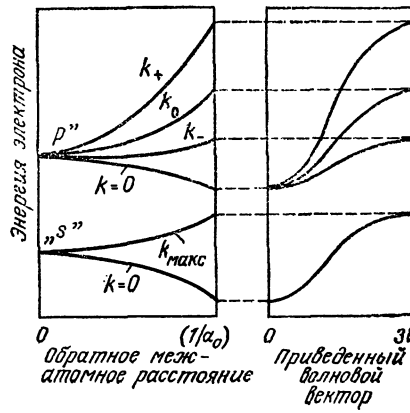


Рис. 3.37. Гипотетический пример того, что может произойти с компонентами атомного p-состояния, когда атомы сближаются и образуют твердое тело. В данном случае p-состояния образуют зону, имеющую в целом $6N$ состояний, с максимальной плотностью состояний в верхней ее части.

В кристаллическом поле атомные p- или d-состояния с шестью или десятью электронами на атом могут оставаться вовсе нерасщепленными или по крайней мере вырожденными при некотором значении $\mathbf{k}$, как показано на рис. 3.37. В результате возникает *энергетическая* зона, содержащая более одной зоны Бриллюэна и имеющая более чем $2N$ электронных состояний. Наоборот, различные компоненты атомных p- и d-состояний

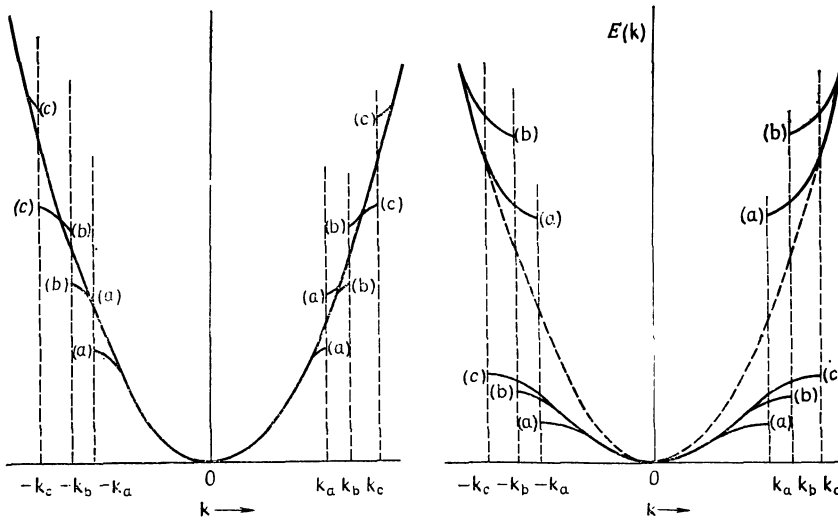


Рис. 3.38. Зависимость энергии от волнового вектора для двух соседних зон в трехмерном твердом теле. k_a, k_b и k_c — координаты границы зоны для различных направлений. Перекрытие происходит для ситуации, изображенной слева, поскольку самая большая энергия зоны *1* превышает самую низкую энергию зоны *2* (для другого направления в $\mathbf{k}$-пространстве). По книге: *Seitz F.* Modern Theory of Solids, McGraw-Hill, 1940.

могут разделиться на неперекрывающиеся полосы энергий. Таким образом, реальные зоны могут иметь $2N$, $4N$, $6N$ и т. д. состояний.

В реальном (трехмерном) твердом теле перекрытие по энергии может иметь место для состояний, которые произошли от атомных уровней совершенно различных типов. Эту возможность иллюстрирует рис. 3.38, на котором показана зависимость $\varepsilon(\mathbf{k})$ для трех различных направлений в k-пространстве. Поскольку предполагается, что периодический потенциал отличен от нуля, энергия должна увеличиваться по мере движения k к границе зоны по любому из направлений k_a, k_b или k_c. Для ситуации, изображенной в правой части рисунка, *любая* энергия во второй зоне больше *любой* энергии в первой зоне. Однако для ситуации, изображенной в левой части рисунка, самая большая энергия для направления k_c в зоне 1 больше самой низкой энергии в зоне 2 для любого другого направления в k-пространстве. Следовательно, электрон из зоны 1 может быть переброшен в зону 2 без изменения энергии; таким образом, эти зоны образуют одну составную зону, содержащую $4N$ состояний. Очевидно, перекрытие происходит и между зонами, изображенными на объемном рис. 3.32.

Металлы, диэлектрики и полупроводники

Ранние попытки объяснить существование таких проводников тока, как металлы, и хороших диэлектриков были основаны на представлении о том, что электроны в диэлектрике должны иметь очень низкую *подвижность*. Подвижность $\mu = (\sigma/ne)$ выражается в единицах $м^2/(В \cdot с)$ (отношение дрейфовой скорости к напряженности приложенного электрического поля); ее значение для типичного металла при комнатной температуре составляет $\mu \approx 0{,}1$ $м^2(В \cdot с)$. В *некоторых* диэлектриках подвижность электронов действительно очень мала[44], однако в целом эта гипотеза оказывается не в состоянии объяснить низкую проводимость многих тел с достаточно большой подвижностью. Известно, что многие диэлектрики под действием освещения приобретают значительную проводимость[45]; параметры этой

[44] В частности, низкие подвижности имеют место в тонких пленках и аморфных диэлектриках, а также в молекулярных твердых телах, например в органических полупроводниках. Как указывал А. Ф. Иоффе [*Иоффе А. Ф.* Физика полупроводников.— М.: Изд-во АН СССР, 1960], существует ряд нерешенных проблем в описании динамики движения электронов в твердых телах с очень малой подвижностью электронов.

[45] Полное обсуждение свойств диэлектрических и полупроводниковых фотопроводников как с теоретической, так и с экспериментальной точек зрения дается в книге: *Bude R. H.* Photoconductivity of Solids, Wiley, 1960. [Имеется перевод: *Бьюб Р.* Фотопроводимость твердых тел.— М.: Мир, 1969.]

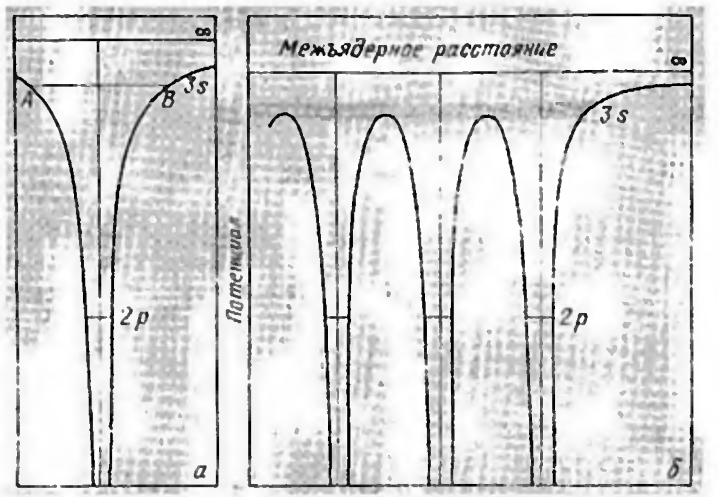


Рис. 3.39. Самосогласованная модель атома натрия (слева) и металлического натрия (справа) [*Slater J. C.*— Physics Today, 21 (No. 1), 43 (1968).] 3*s*-состояние изолированного атома уширяется в зону, которая заполнена наполовину. Для данного частного случая полузаполненная зона, в которой расположен уровень Ферми, совпадает с максимумами периодической потенциальной энергии. Это следует из расчета, на котором основан данный рисунок. Отметим, что 3*s*-зона расположена ниже по энергии, чем изолированное 3*s*-состояние. Это и приводит к связыванию атомов в твердое тело; уменьшение энергии равно энергии сцепления.

фотопроводимости ясно показывают, что и в диэлектриках *могут быть* электроны с очень высокой подвижностью.

С точки зрения зонной теории твердого тела отсутствие металлической проводимости указывает лишь на отсутствие частично заполненных зон. В диэлектрике каждая зона либо полностью заполнена, либо полностью пуста. В натрии, например, имеется 11 электронов на атом. В изолированном атоме эти электроны имеют следующую конфигурацию

$$\mathrm{Na} = [1s^2 2s^2 2p^6]\, 3s^1. \qquad (3.138)$$

Десять внутренних электронов образуют в атоме натрия замкнутые оболочки, и можно ожидать, что в твердом теле они образуют очень узкие зоны, заполняя первые пять зон в расширенном k-пространстве. Единственный внешний электрон, приходящийся на атом, заполняет следующую зону Бриллюэна наполовину, откуда следует, что натрий должен быть металлом. На рис. 3.39 показаны 2*p*- и 3*s*-состояния натрия, наглядно представленные Слэтером.

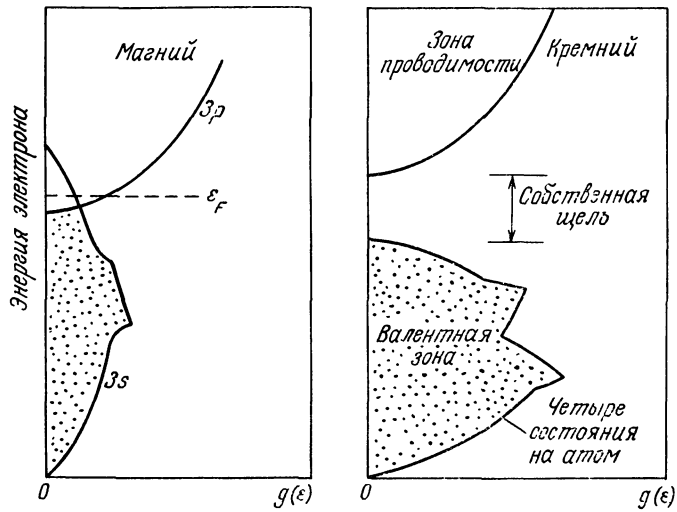


Рис. 3.40. Различие в распределении по энергиям электронных состояний, возникших из 3*s*- и 3*p*-состояний, для металлического магния и неметаллического кремния. (Зоны показаны не в масштабе, и форма зон не является точной). Магний представляет собой металл благодаря перекрытию 3*s*- и 3*p*-зон. В кремнии валентная зона является *s*—*p*-гибридной; она содержит как раз столько мест, сколько нужно для размещения четырех валентных электронов на атом, так что выше остается отдельная пустая зона. На каждой кривой плотности состояний изображены резкие изменения наклона при энергиях, соответствующих сингулярностям Ван Хова.

Следующий элемент периодической системы

$$\mathrm{Mg} = [1s^2 2s^2 2p^6]\, 3s^2, \tag{3.139}$$

очевидно, мог бы быть диэлектриком, *если бы* не было перекрытия по энергии с состояниями 3*p*-зон. Поскольку известно, что магний представляет собой металл, какая-то особенность периодического потенциала должна приводить к перекрытию, схематически изображенному в левой части рис. 3.40. Энергия Ферми в магнии такова, что 3*s*-зона заполнена примерно на 90 %, а заполнение перекрывающейся с ней 3*p*-зоны составляет всего несколько процентов.

Кривые для плотности состояний на рис. 3.40 не претендуют на точность; скорее они иллюстрируют те особенности, которые делают магний металлом, а кремний диэлектриком. В разд. 3.5 выполнено подробное вычисление скалярной величины $g(\varepsilon)$ из дисперсионного соотношения $\varepsilon(\mathbf{k})$. На одной из особенностей рис. 3.40 следует остановить внимание прежде, чем двигаться дальше. Это *сингулярности Ван Хова*, показанные при некоторых энергиях. О них уже шла речь выше в гл. 2, поскольку

такие же сингулярности или *критические точки* имеются в спектре колебаний решетки (см. рис. 2.7). Особенность Ван Хова[46] имеет место при энергии, для которой поверхность постоянной энергии в k-пространстве либо касается в существенном направлении границы зоны, либо изменяет знак кривизны по отношению к зонной границе.

В правой части рис. 3.40 схематически изображен спектр разрешенных и заполненных состояний в кремнии. Этот элемент с четырьмя электронами на атом, которые должны быть распределены по 3*s*- и 3*p*-состояниям, тем не менее оказывается неметаллом. Это происходит потому, что симметрия структуры алмаза, в которой кристаллизуется кремний, расщепляет шесть 3*p*-состояний на атом таким образом, что самая верхняя заполненная зона представляет собой гибридную (*s*—*p*)-зону с четырьмя состояниями на атом. Это именно то число, которое обеспечивает полное размещение всех электронов, внешних по отношению к атомным остаткам с замкнутыми оболочками. Следующая разрешенная зона начинается с энергии, расположенной на 1 эВ выше, но в чистом кремнии при низких температурах эта зона совершенно пуста. Соответственно чистый кремний при низких температурах представляет собой отличный диэлектрик. Суммарное действие дефектов, примесей и тепла приводит к тому, что кремний ведет себя, как *полупроводник*. Таким образом, различие между диэлектриком и полупроводником не принципиальное, а только количественное.

Вильсон[47] отметил, что тепловое возбуждение может опустошить небольшую часть состояний в верхней заполненной зоне диэлектрика (валентной зоне), одновременно заполнив некоторое количество нижних состояний в следующей зоне (зоне проводимости). Проводимость, обусловленная электронами в зоне проводимости и таким же числом *положительных дырок* в валентной зоне, характерна для поведения *собственного полупроводника*, а энергетическое расстояние между валентной зоной и зоной проводимости есть *собственная щель* (ширина запрещенной зоны). Далее Вильсон[48] предположил, что при температурах, слишком низких для значительной генерации собственных электронно-дырочных пар, могут быть термически ионизованы локализованные электронные состояния внутри запрещенной зоны, связанные с дефектами и примесями. При этом будут генерироваться *либо* свободные электроны, *либо* свободные дырки. Такой второй вид поведения по-

---

[46] *Van Hove L.*— Phys. Rev., **89**, 1189 (1953).
[47] *Wilson A. H.*— Proc. Roy. Soc., **A133**, 458 (1931).
[48] *Wilson A. H.*— Proc. Roy. Soc., **A134**, 277 (1931).

лупроводника называется *примесным*, поскольку он зависит от несобственных свойств проводящей среды.

Более подробно диэлектрики и полупроводники будут рассмотрены в гл. 4. Здесь же можно отметить, что неметаллические твердые тела, в которых ширина запрещенной зоны меньше 2* эВ или в которых может быть создана большая плотность локализованных состояний дефектов возле края либо валентной зоны, либо зоны проводимости, считаются полупроводниками. Те твердые тела, в которых запрещенная зона широкая, а уровни дефектов расположены далеко от краев обеих зон, являются при обычных температурах диэлектриками.

Как было указано, распределение $3s$- и $3p$-состояний в кремнии, которое делает этот элемент неметаллом, является следствием симметрии алмазной решетки. Аналогичные ситуации обнаружены и в других элементарных твердых телах и в большом числе соединений. В материале, который существует в двух или большем числе аллотропических модификаций, может иметь место характерное для металла перекрытие энергетических интервалов зон Бриллюэна в одной из возможных кристаллических структур и не иметь места в другой**.

Расчет формы энергетических зон

Исчерпывающее обсуждение методов, применяемых для расчета энергетических зон и щелей в трехмерных твердых телах, не может быть дано в этой книге как из-за обширности вопроса, так и из-за его сложности. Достаточно упомянуть основные теоретические приемы, указав, что в работах, перечисленных в конце главы, содержится подробное их описание.

Одним из первых теоретических методов был *метод ячеек* Вигнера и Зейтца, предложенный в 1933 г. В этом методе исходят из волновой функции, являющейся решением уравнения Шредингера с выбранным электростатическим потенциалом. Затем эта функция должна быть продолжена так, чтобы она была непрерывной и периодической при переходе от области одного атома в область соседнего. Вигнер и Зейтц рассматривали многогранник, построенный в реальном пространстве вокруг каждого атома кристалла таким образом, чтобы совокупность этих многогранников, или *ячеек*, заполняла все пространство. Затем они подбирали волновые функции для электронов, непрерывные вместе со своими производными в важ-

* Это в значительной мере условно. Например, у соединений SiC и GaP, рассматриваемых как полупроводники, ширина запрещенной зоны при комнатной температуре равна $\sim 2{,}2$ эВ.— *Прим. ред.*

** Примером могут служить модификации олова. Обычное белое олово — металл, а серое олово — полупроводник.— *Прим. ред.*

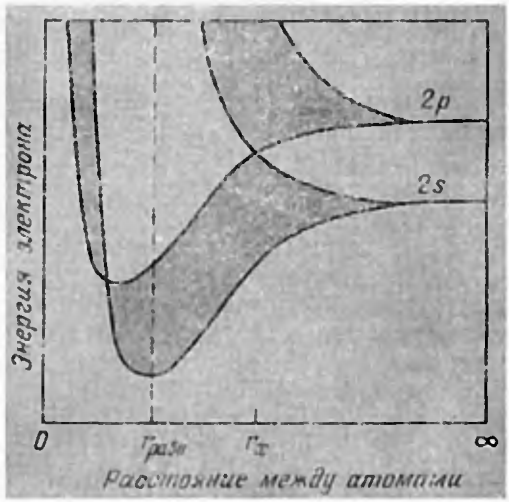


Рис. 3.41. Энергия электрона как функция межатомного расстояния для алмаза, вычисленная Кимбаллом с помощью метода ячеек [*Kimball G. E.*— J. Chem. Phys., 3, 560 (1935)]. Для расстояний, превышающих r_x, зоны представляют собой уширенные атомные $2s$- и $2p$-состояния и содержат соответственно по два и по шесть электронов на атом. (Состояния $1s$, принадлежащие атомным остовам, находятся при гораздо меньшей энергии и здесь опущены.) При расстояниях, меньших r_x, зоны представляют собой s—p-гибриды. Нижняя зона (четыре электрона на атом) заполнена, а верхняя пуста. При равновесном межатомном расстоянии $r_{равн}$ ширина собственной энергетической щели, согласно Кимбаллу, равна 7 эВ.

ных точках границ, отделяющих одну ячейку от другой. Чтобы решить задачу методом Вигнера — Зейтца, часто бывает удобно заменить многогранную ячейку сферической ячейкой того же объема.

Метод ячеек с успехом применялся для простых щелочных металлов. На рис. 3.39 показан пример использования этого метода для вычисления энергетических зон натрия. Другим примером одного из первых успешных описаний зонных энергий с помощью метода Вигнера — Зейтца может служить расчет Кимбалла энергетических зон, образующихся из $2s$- и $2p$-состояний углерода в решетке алмаза. Кимбалл выполнил эти расчеты для различных межатомных расстояний. Результаты его вычислений, представленные на рис. 3.41, правильно предсказывают, что при межатомном расстоянии, меньшем некоторой критической величины r_X, два нижних и шесть верхних состояний на атом (для изолированных атомов) заменяются четырьмя нижними и четырьмя верхними состояниями на атом.

Таким превращением объясняется неметаллическая природа алмаза и гомологичного ему ряда твердых тел (кремний, германий, SiC, GaAs и т. д.).

Несмотря на успехи метода ячеек, в последние годы для неметаллических твердых тел в основном используются другие методы расчета, которые дают лучшую количественную точность. В большинстве этих методов требование непрерывности и периодичности налагается на волновую функцию с самого начала, даже если это требование означает, что исходная волновая функция не удовлетворяет уравнению Шредингера, пока член, описывающий потенциальную энергию, не будет каким-то образом с ней согласован.

Одним предельным случаем таких процедур является *метод сильной связи*, введенный в 1928 г. Блохом в его первоначальном исследовании энергетических зон[39]. На рис. 3.28 сильная связь представлена как самый слабый тип взаимодействия между соседними атомами. В этом приближении из атомных орбиталей (каждая из которых локализована на определенном атоме) составляется линейная комбинация, которая удовлетворяет требованию теоремы Блоха — Флокэ и описывает состояние, бегущее по кристаллу. Структура образующейся при этом зоны зависит от интегралов перекрытия, которые чрезвычайно чувствительны к особенностям поведения атомных орбиталей снаружи от атомных остатков и к межатомному расстоянию в решетке, поскольку в приближении сильной связи действие на электрон атомов, расположенных дальше одного межатомного расстояния, сказывается мало.

Простейшее приближение в методе сильной связи[49] приводит к следующему виду зависимости энергии электрона от волнового вектора для простой кубической решетки:

$$\varepsilon(\mathbf{k}) = \varepsilon_0 - \varepsilon_1 [\cos(ak_x) + \cos(ak_y) + \cos(ak_z) - 3]. \qquad (3.140)$$

Это решение лежит в интервале энергий от ε_0 до $(\varepsilon_0 + 6\varepsilon_1)$. На рис. 3.42 показана зависимость $\varepsilon(\mathbf{k})$ в направлении k_x и ряд кривых постоянной энергии для плоскости $k_x k_y$ в k-пространстве. Только для самых низких энергий (вблизи центра зоны) наблюдается качественное сходство со спектром свободных электронов. Такое описание зоны должно быть правильным, когда межатомное взаимодействие мало и зона представляет собой слегка уширенный атомный уровень.

---

[49] Заинтересованный читатель, имеющий определенную подготовку по квантовой механике, может найти обсуждение этой модели на не слишком сложном уровне в книгах: *Ziman J.* Principles of the Theory of Solids, Cambridge, 2nd ed, 1972. [Имеется перевод: *Займан Дж.* Принципы теории твердого тела — М.: Мир, 1966.] *Dekker A. J.* Solid State Physics, Prentice-Hall, 1957.

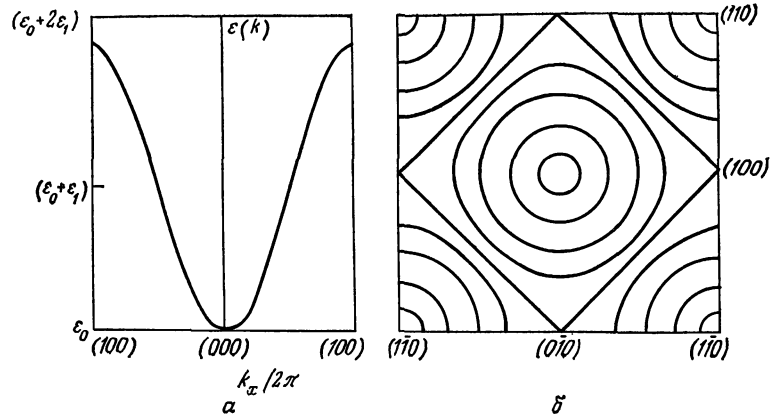


Рис. 3.42. *а* — зависимость энергии электрона от волнового вектора в k_x-направлении для простой кубической решетки в модели сильной связи (формула (3.140)); *б* — линии постоянной энергии в плоскости $k_z=0$ зоны Бриллюэна для этой модели.

Когда имеется значительное межатомное взаимодействие, в приближении сильной связи следует использовать *линейную комбинацию атомных орбиталей* (метод ЛКАО). При этом применяется квантовомеханическая вариационная процедура, позволяющая найти такую комбинацию из *s*-, *p*- и *d*-орбиталей, которая соответствует наименьшей энергии системы. В одной из модификаций этого метода (Левдин, 1950 г.) атомные орбитали выбраны таким образом, что функции с центром на различных атомах автоматически ортогональны.

Методом ЛКАО было выполнено много зонных расчетов, как из первых принципов, так и интерполяционных, в которых используется определенная информация относительно зон в твердом теле, полученная из других источников. Методы сильной связи, естественно, с наибольшим успехом применяются, когда влияние периодического потенциала атомных остатков достаточно велико. Это справедливо, например, для ионных кристаллов, таких, как NaCl, а также для зоны, образованной из 3*d*-состояний 1-й группы переходных элементов. Напомним, что изолированные атомы первой группы переходных элементов имеют один или два электрона в 4*s*-подоболочке и частично заполненную 3*d*-оболочку. Когда эти атомы сближаются и образуют твердое тело, перекрытие 4*s*-состояний очень велико, но перекрытие состояний незаполненного внутреннего уровня довольно слабое. Типичный вид результирующей плотности состояний как функции энергии представлен на рис. 3.43 для никеля. Перекрытие по энергии между 3*d*- и 4*s*-зонами для

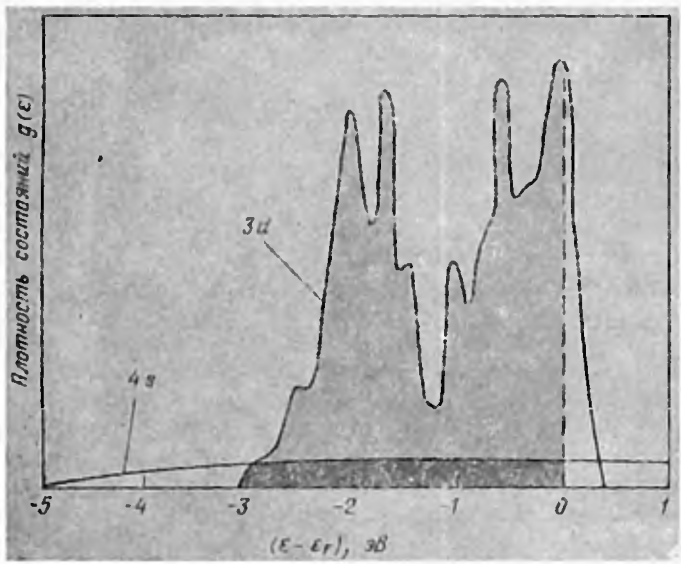


Рис. 3.43. Плотность состояний в никеле, обладающем г. ц. к.-структурой, для зон, расположенных близко к энергии Ферми, вычисленная с помощью метода сильной связи [*Koster G. F.*— Phys. Rev. **98**, 901 (1955)]. Число электронов на атом в никеле как раз достаточно для того, чтобы полностью заполнить 3*d*-зону, если бы 4*s*-зона целиком находилась в области более высоких энергий. Однако 4*s*-зона характеризуется очень низкой плотностью состояний и занимает широкий энергетический интервал; ее нижняя граница лежит ниже дна 3*d*-зоны, а верхняя на 5 эВ выше потолка той же зоны. В результате энергия Ферми такова, что 3*d*-зона заполнена на 94 %, поэтому 4*s*-зона содержит по 0,6 электрона на атом.

переходных элементов приводит к частичному заполнению 4*s*-зоны. Проводимость такого металла почти полностью зависит от небольшого числа высокоподвижных электронов в 4*s*-зоне, однако при энергии Ферми имеется большая плотность 3*d*-состояний, которая определяет электронную теплоемкость и магнитные свойства.

Противоположным методу сильной связи предельным приближением является *метод почти свободных электронов* (метод ПСЭ), который применим для ситуаций с сильным перекрытием. В методе ПСЭ предполагается, что периодический потенциал создает только малое возмущение в спектре состояний свободных электронов. В качестве обоснования такого приближения для электронного газа в металле вспомним, что атомные остатки представляют собой области, где потенциальная энергия электрона отрицательна по отношению к межатомным об-

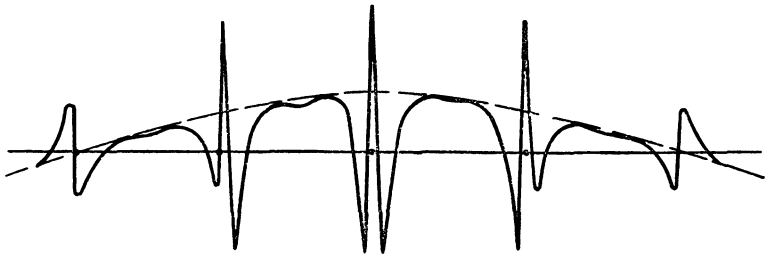


Рис. 3.44. Координатная зависимость волновой функции электронов в 3*s*-зоне металлического натрия для собственного значения энергии, соответствующего $k=\pi/4a$ [*Slater J. C.* — Rev. Mod. Phys., 6, 209 (1934)]. Штриховая кривая изображает простую плоскую волну, которая достаточно хорошо аппроксимирует волновую функцию в областях пространства, внешних по отношению к атомным остовам.

ластям. Из условия сохранения полной энергии следует, что электрон должен двигаться быстрее (обладать большей кинетической энергией), когда ему приходится пересекать область атомного остатка. Поскольку объем, внешний по отношению к атомным остаткам, больше объема внутри них (см. задачу 3.1) электрон, очевидно, почти все время проводит во внешних областях, относительно свободных от действия поля, и его состояние почти не зависит от точного вида потенциала атомного остатка. Поэтому в условиях сильного перекрытия удовлетворительным первым приближением для волновой функции электрона должна быть волновая функция *свободного* электрона (плоская волна). Это приближение показано пунктирной кривой на рис. 3.44 для электрона связи натрия.

Вспомним, что поверхность постоянной энергии в **k**-пространстве для *совершенно* свободных электронов представляет собой сферу с центром в точке $\mathbf{k}=0$. Энергетическая поверхность для $\varepsilon=\varepsilon_F$ (поверхность Ферми) для таких электронов — это идеальная сфера (рис. 3.14). В ПСЭ методе принимается, что *малый* периодический потенциал не будет сильно отклонять форму поверхностей постоянной энергии (таких, как поверхность Ферми) от сферической. Это позволяет установить, какие части **k**-пространства будут заняты, а какие останутся пустыми.

Метод ПСЭ проиллюстрирован на рис. 3.45 для гипотетической квадратной решетки. В схеме расширенных зон все электронные состояния, расположенные внутри сферы некоторого радиуса с центром в начале координат **k**-пространства, показаны заполненными. В правой части рис. 3.45 показано электронное распределение в схеме приведенных зон. Можно видеть, что имеются «дырки» в углах первой зоны и карманы, запол-

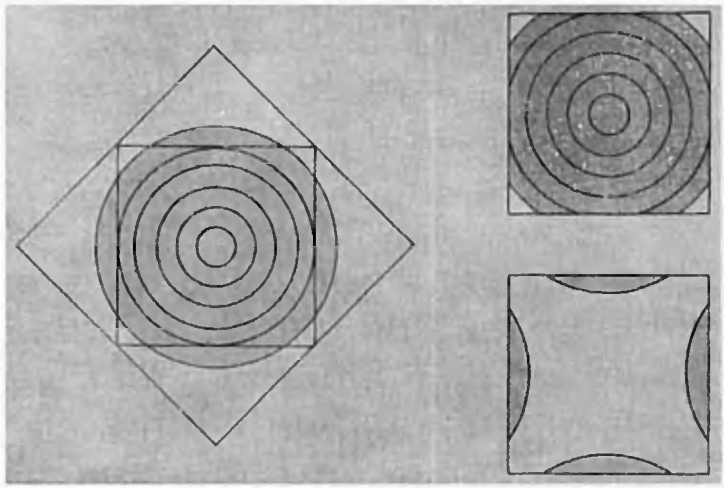


Рис. 3.45. Первые две зоны простой квадратной решетки, показанные в представлении расширенных зон и в представлении приведенных зон. Число электронов достаточно для того, чтобы почти целиком заполнить одну зону, однако если периодический потенциал очень мал (модель почти свободных электронов), то углы первой зоны, соответствующие в ней наибольшим значениям энергии, будут оставаться пустыми, а во второй зоне будут области (карманы), занятые электронами. При увеличении периодического потенциала кривые постоянной энергии в k-пространстве будут деформироваться, и в первой зоне, по-видимому, останется больше электронов.

ненные электронами — во второй зоне. Увеличение амплитуды периодического потенциала деформирует форму кривых постоянной энергии таким образом, что степень заполнения первой зоны увеличивается за счет второй.

Поведение трех приходящихся на атом электронов связи, в алюминии хорошо объясняется моделью ПСЭ[50]. На рис. 3.46 показано сечение поверхности Ферми алюминия, наложенное на границы зон Бриллюэна. Можно видеть, что заметное искажение сферической формы имеет место только вблизи зонных границ. Первая зона полностью заполнена; каждая из следующих трех зон заполнена частично. Харрисон[50] построил объем-

[50] *Harrison W.*— Phys. Rev., **118**, 1183—1190 (1960). Прекрасный обзор теоретических приближений (как ПСЭ, так и многих других) в сочетании с экспериментальными результатами, касающимися формы зон и формы поверхностей Ферми в металлах, см. в книге: The Fermi Surface, eds W. A. Harrison, M. B. Webb, Wiley, 1960.

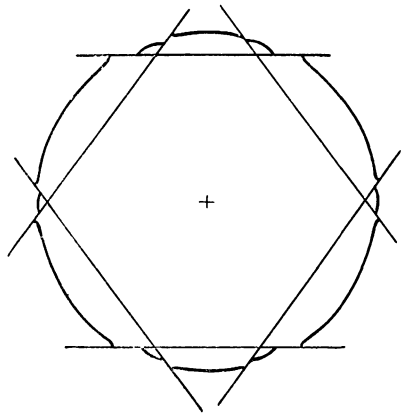


Рис. 3.46. Сечение обратного пространства плоскостью (110) для твердого тела с г. ц. к-структурой. Показаны первая зона и части второй, третьей и четвертой зон. Здесь же изображен контур поверхности Ферми алюминия (три внешних электрона на атом), полученный в предположении, что электроны почти свободны [*Harrison W.*— Phys. Rev., 118, 1190 (1960)].

ные изображения заполненных областей первых четырех зон (на рис. 3.47); некоторые особенности этой картины подтверждаются экспериментом[51].

Модель Кронига — Пенни можно рассматривать как одномерную форму приближения ПСЭ, когда параметр P [см. (3.132)] мал. Для бóльших значений P модель Кронига — Пенни становится одномерной формой метода ППВ (присоединенных плоских волн), о котором пойдет речь ниже.

Существует несколько способов, позволяющих улучшить решения метода ПСЭ, имеющие вид плоских волн. Следует иметь в виду, что одиночная плоская волна достаточно хорошо описывает волновую функцию на рис. 3.44 в области *вне* атомных остатков, но что необходима комбинация огромного числа плоских волн для описания полной волновой функции, так как она меняется очень быстро вблизи атомных остатков. Однако разность между отдельной плоской волной (штриховая линия) и полной волновой функцией может быть описана достаточно просто с помощью *атомных орбиталей* на данном узле. На этом основании Херринг в 1940 г. предложил метод *ортогонализованных плоских волн* (ОПВ), т. е. функций, которые представляют собой линейные комбинации плоских волн и смеси атомных волновых функций для заполненных состояний атомных остатков. При этом свойства электрона вполне удовлетворительно описываются как внутри, так и вне областей атомных остатков. Метод ОПВ был использован во многих расчетах

---

[51] *Kamm G. N., Bohm H. V.*— Phys. Rev., 131, 111 (1963); *Spong F. W., Kip A. F.*— Phys. Rev., 137, A431 (1965).

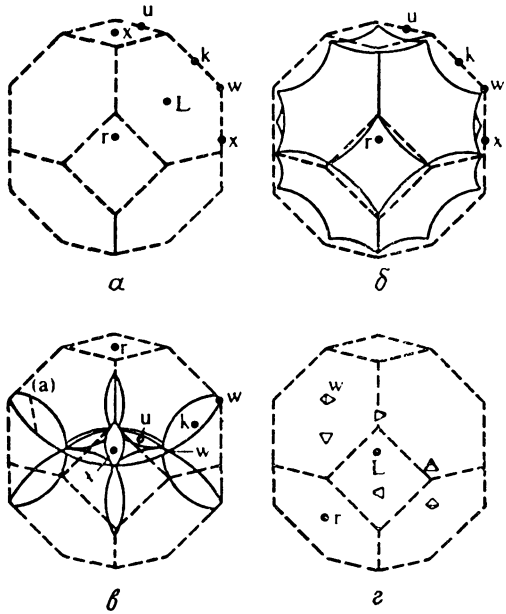


Рис. 3.47. Объемное изображение первых четырех зон в k-пространстве для алюминия (имеющего г. ц. к.-структуру). Показаны заполненные области второй, третьей и четвертой зон, которые все вместе образуют поверхность Ферми. Шесть небольших электронных карманов в четвертой зоне исчезают при увеличении потенциала атомного остова. *а* — первая зона полностью заполнена; *б* — дырочный карман второй зоны; *в* — электронные области третьей зоны; *г* — электронные карманы четвертой зоны [*Harrison W. A.*— Phys. Rev., **118**, 1183 (1960)].

формы зон как для металлических[52], так и неметаллических[53] твердых тел.

Другое приближение, улучшающее метод ПСЭ, возникло в результате анализа самого́ периодического потенциала. Слэтер[54]

[52] Впервые метод был применен для бериллия: *Herring C., Hill A. G.*— Phys. Rev., **58**, 132 (1940). Среди многочисленных последующих расчетов для металлов методом ОПВ можно отметить расчет для алюминия: *Heine V.*— Proc. Roy. Soc., **A240**, 340 (1957), в котором был обоснован подход Харрисона.

[53] Херман [*Herman F.*— Phys. Rev., **93**, 1214 (1954)] приводит результаты расчета методом ОПВ для алмаза и германия. Его расчет для германия — полупроводника, в котором энергия зоны проводимости имеет наименьшее значение, а энергия в валентной зоне достигает своего наивысшего значения в разных частях зоны Бриллюэна, открыл новую эру в понимании полупроводников (см. примечание 37 этой главы относительно более тонких расчетов методом ОПВ для германия, проделанных Херманом).

[54] *Slater J. S.*— Phys. Rev., **51**, 846 (1937). Трудности использования метода ППВ не столь непреодолимы, как это представляется на первый

привел доводы в пользу того, что разумной аппроксимацией для потенциала должен быть постоянный потенциал вне атомных остатков и обычный атомный потенциал внутри некоторой сферы, окружающей каждый атомный остаток. Уравнение Шредингера в этом случае решается для областей каждого типа отдельно и решения сшиваются на сферических границах между областями. Внутри каждого атомного остатка волновую функцию можно разложить в ряд по сферическим функциям. Вне атомных остатков решение, конечно, представляет собой суперпозицию плоских волн. Использование такой комбинации известно под названием метода *присоединенных плоских волн* (ППВ). Метод ППВ представляется очень сложным, однако (с помощью больших вычислительных машин) он был успешно применен как в расчетах для металлических твердых тел, так и для полупроводников.

Для расчета энергетических зон был использован ряд других методов, которых слишком много, чтобы перечислять их здесь. Характерная особенность многих из них состоит в том (как и в методе ППВ), что уравнение Шредингера решается для потенциального распределения, представляющего собой измененный по сравнению с истинным периодический потенциал. Использование такого *псевдопотенциала* может привести к неправильным волновым функциям, однако часто позволяет с достаточной точностью установить, как зависит энергия от волнового вектора внутри различных зон Бриллюэна.

3.5. Динамика движения электронов

Фазовая и групповая скорости

Тот факт, что при описании электронов в реальных кристаллах с помощью волновых функций фазовая скорость волны не всегда совпадает как по величине, так и по направлению с фактической скоростью движения электрона, может вызывать недоумение. Для иллюстрации рассмотрим точку **K** в k-пространстве, изображенном на рис. 3.48. Волны с волновым фактором, близким к **K**, движутся в направлении **OK** с фазовой скоростью $\mathbf{v}_p = (\omega/\mathbf{k})$. Казалось бы, отсюда следует, что существенно полное значение **k**, которое мы исключили из рассмотрения, введя представление о приведенной зоне.

взгляд. Это было показано тем же автором в работе: Phys. Rev., **92**, 603 (1953). Метод был успешно использован как для металлов, так и для полупроводников.

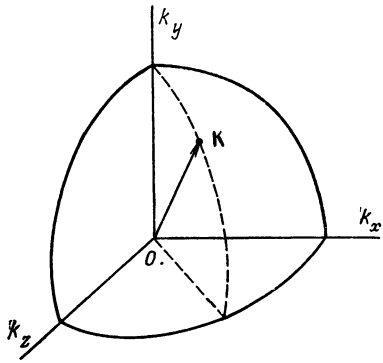


Рис. 3.48. Для описания состояния движения электрона используется соответствующая ему точка в k-пространстве.

Однако на самом деле фазовая скорость не имеет реального физического смысла. *Волновой пакет* с центром в точке **K**, описывающий электрон, движется по решетке в реальном пространстве с групповой скоростью $\mathbf{v}_g = \partial\omega/\partial\mathbf{k}$. Если состояние электрона характеризуется волновым вектором **K** и соответствующей ему энергией ε_K, то его угловая частота (как следует из зависящего от времени уравнения Шредингера) равна $\omega = (\varepsilon_K/\hbar)$. Тогда групповая скорость составляет

$$\mathbf{v}_g = \frac{\partial\omega}{\partial\mathbf{k}} = \frac{\partial}{\partial\mathbf{k}}\left(\frac{\varepsilon_k}{\hbar}\right) = \frac{1}{\hbar}\nabla_k\varepsilon. \qquad (3.141)$$

В этой формуле достаточно знать величину k, отсчитанную либо от границы зоны, либо от центра приведенной зоны.

Чтобы понять это утверждение, вспомним, что волна вероятности

$$\psi = U_k(\mathbf{r})\exp[i(\mathbf{k}\cdot\mathbf{r} - \omega t)] \qquad (3.142)$$

распространяется по твердому телу таким образом, что плотность заряда $-e|\psi|^2$ в каждой ячейке имеет одно и то же среднее по времени значение. Фазовые скорости отдельных волн представляют собой математическую абстракцию, и только движение центра заряда для волнового пакета описывает реальную траекторию электрона. Следовательно, величина k, отсчитанная от края зоны, определяет скорость, направление движения и энергию электрона.

Даже для совершенно свободного электронного газа фазовая и групповая скорости различны по *величине* (см. задачу 3.28), хотя в этом случае групповая скорость и направлена по **OK** для электрона с волновым вектором **K**. Однако для распределения электронных состояний, обусловленного конечным периодическим потенциалом, поверхности постоянной энергии не обязательно сферические. На рис. 3.49 показано поперечное сечение такой деформированной поверхности. Волновой пакет (в котором каждая отдельная волна обладает фазовой скоростью, параллельной **OK**) характеризует состояние электрона в точке **K**, однако любой электрон, занимающий это состояние, будет двигаться в реальном пространстве со скоростью $\mathbf{v}_g = (1/\hbar)\nabla_k\varepsilon$. Это движение происходит в направлении нормали

к поверхности постоянной энергии. Как показано на рис. 3.49, $\mathbf{v}_g$, очевидно имеет направление, совершенно отличное от **OK**.

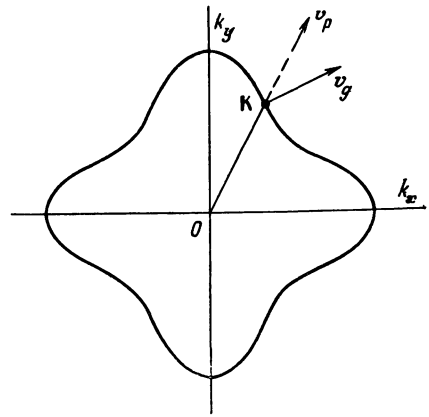


Рис. 3.49. Различие между направлениями фазовой и групповой скорости для деформированных поверхностей постоянной энергии.

Динамика электронов сильно зависит от формы поверхностей постоянной энергии. Если на электрон не действуют дополнительные силы, то в реальном пространстве он остается в состоянии равномерного прямолинейного движения со скоростью $\mathbf{v}_g$. При этом он может быть представлен в **k**-пространстве одной и той же точкой, не зависящей от времени. Такую ситуацию могут изменить три типа воздействий.

1. Электрон может испытать рассеяние. Обычно это приводит к изменению как энергии, так и волнового вектора. В результате электрон исчезает из одной точки **k**-пространства и появляется в другой.

2. Электрон может испытывать действие электрического поля. При этом энергия электрона должна изменяться. Перемещение электрона в **k**-пространстве и в реальном пространстве определяется скоростью изменения энергии.

3. Электрон может испытывать действие магнитного поля. При наличии только одного этого поля требуется, чтобы направление движения (и положение в **k**-пространстве) изменялось без изменения энергии электрона. Следовательно, под действием магнитного поля электрон движется в **k**-пространстве по поверхности постоянной энергии.

Обсудим действие электрического и магнитного полей более подробно. При этом необходимо ввести *эффективную массу*, характеризующую инерцию электрона, находящегося в поле периодического потенциала. Хотя формально полупроводники рассматриваются в гл. 4, в данном разделе удобно исследовать динамику электронов применительно как к металлам, так и полупроводникам.

Эффективная масса

Рассмотрим в периодическом твердом теле электрон, который находится под действием внешнего электрического поля $\mathbf{E}$. Сила, действующая на электрон, равна $-e\mathbf{E}$. Следовательно, скорость изменения энергии электрона составляет

$$\frac{d\varepsilon}{dt} = -e\mathbf{E}\cdot\mathbf{v}, \tag{3.143}$$

где $\mathbf{v}$ — скорость движения самого электрона, т. е. величина v_g, а не v_p, о которой шла речь в предыдущем разделе. Бесконечно малое изменение энергии можно записать как $d\varepsilon = \nabla_k \varepsilon d\mathbf{k}$. Подставим выражение для скорости (3.141) в уравнение (3.143). В результате этих двух подстановок получим

$$\hbar\,(d\mathbf{k}/dt) = -e\,\mathbf{E}. \tag{3.144}$$

Вспомнив что сила есть скорость изменения импульса, убеждаемся в том, что величина $\hbar\mathbf{k}$ по отношению к внешним силам ведет себя как импульс электрона.

Ускорение электрона под действием приложенного внешнего поля равно

$$\mathbf{a} \equiv \left(\frac{d\mathbf{v}}{dt}\right) = \frac{1}{\hbar}\nabla_k\left(\frac{d\varepsilon}{dt}\right) = \frac{1}{\hbar}\nabla_k(-e\,\mathbf{E}\cdot\mathbf{v}) =$$

$$= \frac{1}{\hbar}\nabla_k\left[\frac{(-e\,\mathbf{E})}{\hbar}\nabla_k\varepsilon\right] = \frac{-1}{\hbar^2}\nabla_k\nabla_k\varepsilon\,(e\,\mathbf{E}). \tag{3.145}$$

Таким образом, ускорение имеет компоненты

$$a_i = \frac{-1}{\hbar^2}\sum_j \frac{\partial^2\varepsilon}{\partial k_i \partial k_j}(eE_j). \tag{3.146}$$

Сравнение выражений (3.145) и (3.146) с законами движения Ньютона показывает, что тензорная величина $1/\hbar^2\nabla_k\nabla_k\varepsilon$ имеет размерность (масса)$^{-1}$. Поэтому величину

$$[m_{ij}] = \hbar^2\,[\partial^2\varepsilon/\partial k_i \partial k_j]^{-1} \tag{3.147}$$

называют *тензором эффективной массы* для электрона, находящегося в периодическом поле.

В некоторых твердых телах с достаточно сложной структурой недиагональные компоненты этого тензора велики и электрическое поле, приложенное в одном направлении, вызывает ускорение в другом направлении. Так происходит, когда поверхности постоянной энергии в $\mathbf{k}$-пространстве заметно отличаются от сферических, как, например, показано на рис. 3.49. Для идеально простого и изотропного твердого тела все недиагональные компоненты тензора эффективной массы обращаются в нуль, а три диагональные компоненты, допускаемые видом

периодического потенциала, одинаковы. В таком идеальном твердом теле электрон имеет *скалярную эффективную массу*

$$m^* = \frac{\hbar^2}{d^2\varepsilon/dk^2} \tag{3.148}$$

по крайней мере в части области изменения своей энергии (т. е. по крайней мере в некоторых областях зоны Бриллюэна).

Когда зависимость энергии от волнового вектора квадратична, то можно утверждать, что электроны обладают не зависящей от энергии эффективной массой, поскольку

$$\varepsilon = \frac{\hbar^2 k^2}{2m^*}. \tag{3.149}$$

Однако в широком энергетическом интервале зависимость ε от $\mathbf{k}$ неизбежно отклоняется от простого вида (3.149). Такое отклонение зависимости $\varepsilon(\mathbf{k})$ от квадратичной часто описывают как зависимость эффективной массы от энергии, хотя его можно с равным успехом описывать и другими способами. Эффективную массу следует рассматривать как удобный способ описания многих сложных явлений в твердом теле.

В качестве примера укажем, что в твердом теле с сильно анизотропной кристаллической структурой тензор эффективной массы может иметь положительные компоненты для одних направлений и отрицательные — для других.

Для любого периодического потенциала в зоне Бриллюэна найдутся точки, в которых уравнение Шредингера дает наименьшее возможное значение энергии электрона. Это может быть единственная точка (в центре зоны), несколько точек (эквивалентных вследствие симметрии) на поверхности зоны или ряд точек (также эквивалентных вследствие симметрии) внутри зоны. Рис. 3.50 может служить примером последней ситуации, когда энергия минимальна в шести эквивалентных точках внутри зоны Бриллюэна на направлениях [100], а поверхность постоянной энергии для несколько большей энергии представляет собой шесть вытянутых эллипсоидов с центрами в этих точках.

В любом случае будем считать, что энергия электрона минимальна, если $\mathbf{k}=\mathbf{k}_1$ (и в соответствующем числе точек, эквивалентных $\mathbf{k}_1$). Тогда все три главные компоненты тензора массы m_{ij}, *как правило*, положительны в окрестности точки $\mathbf{k}_1$ (и во всех эквивалентных областях $\mathbf{k}$-пространства). Более того, для энергий, отличающихся от ε_1 на доли электронвольта, эффективная масса *обычно* практически не зависит от энергии:

$$\varepsilon \approx \varepsilon_1 + A_1(\mathbf{k}-\mathbf{k}_1)^2. \tag{3.150}$$

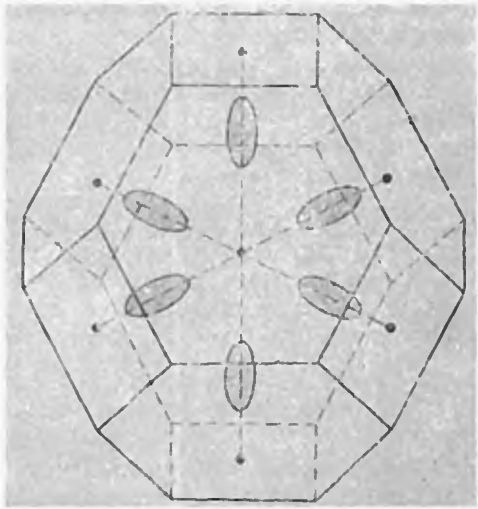


Рис. 3.50. Пример зоны, для которой минимум электронной энергии не расположен ни в центре, ни на одной из внешних границ зоны Бриллюэна. Минимум энергии находится в точке, где волновой вектор равен $(a00)$, и еще в пяти точках, эквивалентных данной в силу симметрии. Для несколько большей энергии условие постоянства энергии выполнено на поверхности шести эллипсоидов. Такой вид имеет самая нижняя часть зоны проводимости полупроводникового кремния, а также вторая снизу зона проводимости арсенида галлия.

Однако в этом выражении A_1 обычно зависит от направления вектора $(\mathbf{k}—\mathbf{k}_1)$. Таким образом, для электронов, расположенных у дна зоны разрешенных энергий, можно утверждать, что

$$\varepsilon = \varepsilon_1 + \frac{\hbar^2}{2m_1}(\mathbf{k}—\mathbf{k}_1)^2, \tag{3.151}$$

где m_1 — эффективная масса *для данного направления* в k-пространстве, отсчитываемого от зонного минимума.

В тех частях зоны Бриллюэна, где энергия электрона достигает максимума, по-видимому, все три компоненты $(\partial^2\varepsilon/\partial k_j^2)$ *отрицательны*, так что следует записать

$$\varepsilon = \varepsilon_2 — \frac{\hbar^2}{2m_2}(\mathbf{k}—\mathbf{k}_2)^2. \tag{3.152}$$

Электрон, принадлежащий к такой группе энергетических состояний, реагирует на поле так, как если бы он обладал отрицательной массой. Таким образом, точка реального пространства, в которой находится *незаполненное* состояние возле потолка зоны, перемещаясь по кристаллу, ведет себя как частица,

которую обычно называют *положительной дыркой*. Дырка имеет положительный заряд (по сравнению с зарядом заполненного состояния) и положительную массу m_2. Для зоны, которая заполнена почти до потолка (как в случае валентной зоны полупроводника или для не полностью заполненных зон в некоторых металлах), более удобно говорить о движении дырок, а не электронов.

Реальность положительных дырок

Положительные дырки будут рассмотрены ниже, в гл. 4, поскольку электроны и дырки играют одинаковую роль в проводимости полупроводников. В металлах и полуметаллах их роль достаточно важна, поэтому рассмотрим их несколько подробнее.

Проводимость может иметь место в зоне, заполнение которой меняется в широких пределах: от почти пустой до почти заполненной, однако она невозможна в зоне, которая совершенно пуста или целиком заполнена. Электроны в целиком заполненной зоне могут обмениваться состояниями, но это ни к чему не приводит, поскольку они неразличимы. Проводимость в почти пустой зоне осуществляется электронами с положительной массой и отрицательным зарядом. Явления, связанные с электропроводностью, такие, как эффект Холла, зависящие от первой степени заряда подвижных носителей, имеют соответственно отрицательный знак.

Говорят, что проводимость почти заполненной зоны обусловлена движением небольшого числа незанятых состояний — положительных дырок, каждая из которых имеет положительную массу и положительный заряд. Шокли[55] детально проанализировал волновую функцию для зоны, имеющей единственное незанятое состояние, и показал, что электрическое и магнитное поле обусловливают следующую зависимость импульса от времени:

$$d\,\mathbf{p}/dt = +e\,[\mathbf{E} + (\mathbf{v} \times \mathbf{B})] \qquad (3.153)$$

[ср. со знаком в формуле (3.35)], что соответствует проводимости дырки.

Представление о дырочной проводимости объясняет положительные коэффициенты Холла у кадмия и железа (см. табл. 3.2). Зоны в кадмии подобны зонам магния, показанным на рис. 3.40, с малой степенью перекрытия между (почти за-

[55] *Shockley W.* Electrons and Holes in Semiconductors, Van Nostrond, 1950, sect 15.8. [Имеется перевод: *Шокли В.* Теория электронных полупроводников.— М.: ИЛ, 1953.]

полненной) зоной, происходящей от 5*s*-состояний и почти пустой 5*p*-зоной. Дырки из 5*s*-зоны стремятся сделать коэффициент Холла положительным, а электроны 5*p*-зоны — отрицательным, и в этом состязании победа остается за дырками. В железе положительный коэффициент Холла обусловлен 80 %-ным заполнением 3*d*-зоны.

Если у читателя еще остаются сомнения в правомочности рассматривать *отсутствие* частицы в качестве подвижного носителя заряда, то постараемся облегчить задачу с помощью следующей очень грубой аналогии. Рассмотрим четыре стеклянные трубки, запаянные с обоих концов. Пусть одна из них будет пустой, другая — целиком заполнена водой, еще одна — почти целиком заполнена водой, а четвертая содержит небольшое количество воды. Если один из концов каждой трубки поднять и опустить, то, очевидно, в пустой трубке не произойдет никакого переноса. Также не произойдет никакого переноса жидкости в заполненной трубке, поскольку молекулы воды (подобно электронам) могут только обмениваться местами. Однако можно наблюдать, как вода перемещается от одного конца к другому в трубке, содержащей небольшое количество жидкости, в то время как *воздушные пузыри* движутся в противоположном направлении в почти полной трубке. Воздушные пузыри в трубке, заполненной водой, ведут себя таким же образом, как дырки в заполненной зоне. Дырки поднимаются к состояниям с максимально допустимой энергией электрона и движутся к отрицательному контакту в электрическом поле.

Сложный вид эффективной массы

Теперь вернемся к свойствам тензора эффективной массы во всем интервале энергий зоны. Следует отдавать себе отчет, что массы m_1 и m_2 в формулах (3.151) и (3.152) не равны друг другу (если они не совпали случайно) и ни одна из них, вообще говоря, не равна обычной массе электронов. Величины m_1 и m_2 отражают способность электрона у дна зоны или дырки у ее потолка реагировать на внешнее поле, *действующее совместно* с межатомным периодическим потенциалом. В зависимости от вида этого потенциала эффективная масса численно может быть больше или меньше обычной массы изолированного электрона.

Для промежуточных областей энергетического интервала внутри зоны, расположенных достаточно высоко по отношению к ε_1 и достаточно низко по отношению к ε_2 тензор эффективной массы неизбежно имеет более сложный вид. Характер сложностей определяется симметрией кристалла и величиной периодического потенциала. В тех случаях, когда этот потенциал

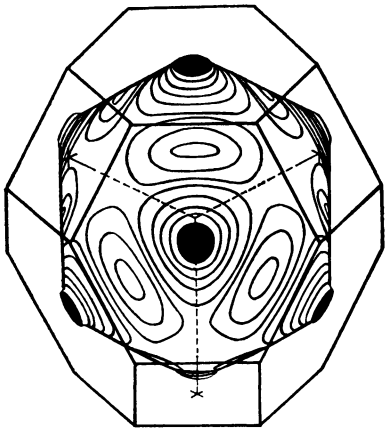


Рис. 3.51. Поверхность Ферми для меди. Поверхность Ферми в этом металле формируется электронами, расположенными в заполненной наполовину 4*s*-зоне. Эта форма поверхности была получена Пиппардом (1957 г.) из результатов измерений анизотропии аномального скин-эффекта в меди [*Smith G. E.* The Fermi Surface, eds. W. A. Harrison, M. B. Webb, Wiley, 1960].

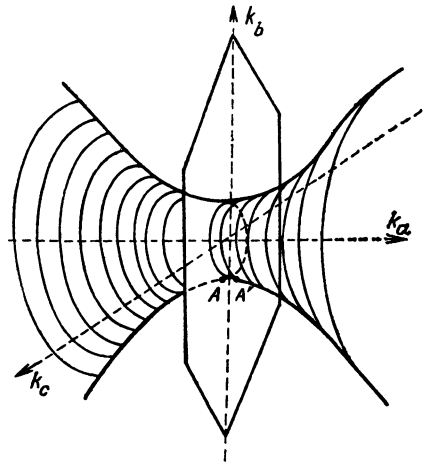


Рис. 3.52. Часть ферми-поверхности меди, показанная в представлении *повторяющихся зон*. Для энергетических состояний на границе зоны эффективная масса положительна в направлении k_b и k_c, но отрицательна в направлении k_a, перпендикулярном плоскости зонной границы. Часть ферми-поверхности, имеющая форму такого типа, известна в литературе под названием «шейки». В магнитном поле электрон можно заставить прецессировать вокруг такой «шеечной орбиты» постоянной энергии.

мал (как предполагалось при построении кривых постоянной энергии на рис. 3.45) или когда он велик (как на рис. 3.42), поверхности постоянной энергии должны отклоняться от сферической формы и не описываться квадратичными зависимостями (3.151) и (3.152). В качестве примера формы поверхности промежуточной энергии на рис. 3.51 дано объемное изображение поверхности Ферми в k-пространстве для меди. Эту поверхность можно получить из сферы, вытянув ее наружу вдоль направлений [111] (таким образом, чтобы вдоль этих направлений она достигала границы зоны Бриллюэна), и вдавив по различным направлениям [110]. В результате «эффективная масса», соответствующая состоянию с энергией Ферми, в большой степени зависит от направления соответствующего волнового вектора.

Поверхность Ферми, изображенная на рис. 3.51, была получена в результате большого числа чрезвычайно сложных экспе-

риментов[56], в которых использовались методы, развивающиеся с середины 50-х годов. Некоторые из этих средств экспериментального исследования будут рассмотрены в настоящем разделе. При изучении динамики электронов можно наблюдать движение электрона, который имеет как положительные, так и отрицательные компоненты тензора эффективной массы, как у электронов на поверхности Ферми меди вблизи областей, примыкающих к границе зоны.

Особенности движения электрона в этих областях k-пространства можно более ясно себе представить, обратившись к рис. 3.52. На этом рисунке использовано *представление повторяющихся зон*, в котором ситуация, существующая с одной стороны от зонной границы, повторяется как зеркальное отражение с другой стороны. (Такой способ представления правомерен, поскольку состояния, разделенные вектором обратной решетки, одинаковы. Поэтому правая часть рис. 3.52 изображает состояния, которые равным образом могли быть представлены на расстоянии одной ширины зоны слева.) Для гиперболической поверхности энергии на рис. 3.52 эффективная масса должна быть отрицательна в направлении $\mathbf{k}_a$, перпендикулярном к плоскости зонной границы, и положительна в направлениях, ортогональных $\mathbf{k}_a$. Динамика движения электрона на такой гиперболической поверхности энергии отражает наличие положительных и отрицательных компонент массы, которые могут быть определены экспериментально.

Прежде чем закончить обсуждение рис. 3.52, необходимо отметить, что поверхность энергии всегда пересекает границу зоны под прямым углом. Это происходит потому, что электрон

---

[56] Одним из чрезвычайно эффективных методов исследования поверхностей Ферми (поверхностей постоянной энергии) в металлах является измерение аномального скин-эффекта, т. е. высокочастотной проводимости чистого металлического образца на высоких частотах, на которых глубина проникновения переменного поля в металл гораздо меньше средней длины свободного пробега электрона. В ряде других методов использовано квантование электронных орбит в магнитном поле, в результате которого движение электрона в плоскости, перпендикулярной магнитному полю, периодично с циклотронной частотой $\omega_c = (eB/m^*)$. С этим квантованием, которое обсуждается ниже в данном разделе, связано большое число эффектов, включая циклотронный резонанс для соответствующей комбинации высокочастотного поля и постоянного магнитного поля, осциллирующую зависимость электропроводности и магнитной восприимчивости от постоянного магнитного поля и магнитную модуляцию коэффициента поглощения ультразвуковых волн В различных статьях, содержащихся в сб. The Fermi Surface, eds. W. A. Harrison, M B. Webb, Wiley, 1960, сообщаются результаты, полученные этими методами. Исчерпывающий обзор используемых экспериментальных методов дан в работе: *Pippard A. B* — Rep Physics, 23, 176 (1961). См. также: *Ziman J.* Electrons in Metals, Taylor and Francis, 1963 [Имеется перевод: *Займан Дж.* Электроны в металлах.— М : Мир, 1962.]

на границе зоны не имеет (в реальном пространстве) компоненты скорости, перпендикулярной граничной плоскости зоны Бриллюэна (в **k**-пространстве). Как мы видели при одномерном анализе границ зон, на границе электронное состояние представляет собой не *бегущую*, а *стоячую* волну, поскольку здесь реализуются условия брэгговского отражения. В случае нескольких измерений разрешено движение, параллельное зонной границе, но движение, перпендикулярное границе, запрещено. По этой причине состояния A и A' на рис. 3.52 отвечают одинаковым условиям движения.

В связи с рис. 3.27, построенным для одномерной модели Кронига — Пенни, было отмечено, что, как правильно показано на рисунке, $(d\varepsilon/dk)$ обращается в нуль у дна и потолка каждой зоны, а величина $(d^2\varepsilon/dk^2)$ остается в этих точках конечной. Конечные вторые производные определяют величину электронной и дырочной эффективных масс, а обращающиеся в нуль первые производные подтверждают [см. формулу (3.141)], что достигается брэгговское состояние стоячей волны с нулевой групповой скоростью. По тем же причинам такую же кривизну имеют кривые, изображающие зависимость $\varepsilon(\mathbf{k})$ на рис. 3.24, 3.28, 3.37 и 3.38.

Следует ли придерживаться такого же правила при изображении поверхностей постоянной энергии в многомерном **k**-пространстве? Только что в связи с рис. 3.52 было указано, что в соответствии с условием Брэгга [формула (3.133)] электрон на границе зоны в реальном пространстве должен иметь нулевую компоненту скорости, перпендикулярную к плоскости границы зоны Бриллюэна. Поэтому поверхность постоянной энергии в **k**-пространстве должна пересекать плоскость границы зоны под прямым углом. Обращаясь к приведенным ранее рисункам, мы видим, что на рис. 3.42, *б* кривые постоянной энергии в углах зоны подходят к зонной границе под прямыми углами. Прямые линии, соответствующие промежуточной энергии в зоне, должны слегка искривляться в направлении нормального падения на очень малом расстоянии от границы зоны.

Энергетические кривые, разрезающие зонную границу на рис. 3.45 изображены *не перпендикулярными* к ней. Теперь мы понимаем, что это следовало бы сделать если бы периодический потенциал был конечным. Соответствующее искривление участков сечения поверхности Ферми на рис. 3.46 для алюминия получено в приближении ПСЭ.

Плотность состояний

Независимо от того, насколько сложной и разнообразной может быть форма поверхностей постоянной энергии, для каждой энергии, лежащей в пределах данной зоны, всегда можно определить скалярную плотность состояний $g(\varepsilon)$ таким образом, чтобы выполнялось равенство

$$2N = \int_{\varepsilon_1}^{\varepsilon_2} g(\varepsilon)\, d\varepsilon. \tag{3.154}$$

Типичный вид такого распределения состояний, соответствующего двум состояниям на каждый атом твердого тела, показан на рис. 3.53. Такая кривая может быть описана с помощью (зависящей от энергии) эффективной массы по плотности состояний.

Изобразим поверхность постоянной энергии в приведенном k-пространстве первой зоны для энергии электрона ε. (Это будет набор эквивалентных поверхностей, если ситуация подобна изображенной на рис. 3.50). Пусть полная площадь этой поверхности (или поверхностей) равна S, тогда объем k-пространства, заключенного между поверхностями, соответствующими энергиям ε и $(\varepsilon + d\varepsilon)$, равен

$$\int \left(\frac{dS}{\nabla_k \varepsilon}\right) d\varepsilon. \tag{3.155}$$

Поскольку объем k-пространства, приходящийся на одно электронное состояние, для кристалла с объемом V составляет $(4\pi^3/V)$, плотность состояний (для единичного объема) можно определить равенством

$$g(\varepsilon)\, d\varepsilon = \left[\int \frac{dS}{4\pi^3 \nabla_k \varepsilon}\right] d\varepsilon. \tag{3.156}$$

В простейшем случае

$$\varepsilon = \varepsilon_1 + \frac{\hbar^2}{2m_1}(\mathbf{k} - \mathbf{k}_1)^2, \tag{3.157}$$

причем m_1 не зависит от направления $(\mathbf{k}-\mathbf{k}_1)$. В этой удачной ситуации $d\varepsilon/dk = \hbar^2 k/m_1$, а $\int dS = 4\pi k^2$, поэтому

$$g(\varepsilon) = \frac{1}{2\pi^2}\left(\frac{2m_1}{\hbar^2}\right)^{3/2} (\varepsilon - \varepsilon_1)^{1/2}. \tag{3.158}$$

Рис. 3.53. Типичная кривая зависимости плотности состояний от энергии внутри зоны. Полная площадь, заключенная под кривой, дает полное число состояний в зоне, равное для простой зоны удвоенному числу атомов в кристалле. Для энергий, не слишком превышающих ε_1, кривизна определяется параметром массы m_1. «Дырочная» масса m_2 определяет кривизну кривой для самой верхней части зоны. Для промежуточных энергий $g(\varepsilon)$ не изменяется с энергией ни как $(\varepsilon-\varepsilon_1)^{1/2}$, ни как $(\varepsilon_2-\varepsilon)^{1/2}$. Особенности Ван Хова на кривой плотности состояний должны возникнуть при энергиях ε_a и ε_b. (Через ε_a обозначена наименьшая энергия, при которой поверхность постоянной энергии соприкасается с границей зоны.)

Это выражение отличается от формулы (3.44) для модели свободных электронов Зоммерфельда только использованием величины m_1 вместо обычной массы электрона [57].

---

[57] Если энергия растет как квадрат волнового вектора (т. е. эффективная масса m^* не зависит от энергии), то решения уравнения Шредингера

$$[(-\hbar^2/2m)\nabla^2 + \mathscr{V}(r)]\psi = \varepsilon\psi$$

в соответствующем интервале энергий совпадают с решениями уравнения

$$(-\hbar^2/2m^*)\nabla^2\psi = \varepsilon\psi.$$

Другими словами, «свобода» модели Зоммерфельда вновь получена перенормировкой массы, учитывающей влияние периодического потенциала. (Можно провести аналогию с перенормировкой массы в квантовой электродинамике.)

Предположим, что полученные таким образом состояния лежат вблизи дна зоны. В этом случае часто возникает вопрос: можно ли пользоваться тем же самым значением m^* для описания состояний *ниже* дна зоны, т. е. связанных состояний, соответствующих некоторому локализованному дефекту. Уравнение Шредингера для электрона, находящегося в поле неподвижного положительного заряда, имеет вид

$$[(-\hbar^2/2m^*)\nabla^2 - (e^2/4\pi K\varepsilon_0 r)]\psi = \varepsilon\psi,$$

где, как можно надеяться, $\varkappa$ — макроскопическая диэлектрическая проницаемость. Это уравнение дает состояния вида (3.128) с энергетическим спектром, отличающимся от спектра атома водорода множителем $(m^*/\varkappa^2 m)$. Такая «водородоподобная модель» или «модель эффективной массы» может успешно применяться при описании связанных состояний с малой энергией связи (состояний, которые в реальном пространстве занимают область, значительно превышающую элементарную ячейку). Для связанных состояний с большой

Независимо от того, насколько сложными могут быть поверхности постоянной энергии в **k**-пространстве (многосвязанные деформированные, незамкнутые поверхности и т. д.), во всех случаях, когда величина $g(\varepsilon)$ такова, что

$$\frac{d}{d\varepsilon}[g(\varepsilon)]^2 \to \text{const}, \tag{3.159}$$

можно считать, что плотность состояний характеризуется *эффективной массой плотности состояний* m^*. Дело в том, что если величина $g(\varepsilon)$ конечна и удовлетворяет условию (3.159), ее всегда можно привести к виду

$$g(\varepsilon) = \frac{1}{2\pi^2}\left(\frac{2m^*}{\hbar^2}\right)^{3/2}(\varepsilon - \text{const})^{1/2}. \tag{3.160}$$

В тех случаях, когда условие (3.159) не выполняется, требуется введение массы плотности состояний, «зависящей от энергии». Трудность в понимании электронных явлений в металлах по сравнению с полупроводниками отчасти связана с тем, что электроны в полупроводнике, находятся в состояниях, близких к краю зоны (где велика вероятность того, что равенство (3.160) является хорошим приближением), в то время как электроны, обладающие энергией Ферми в металле, занимают такие состояния из промежуточного интервала энергий, для которых условие (3.159) обычно плохо выполняется[58].

энергией связи и малым боровским радиусом более существенную роль играет фактическая природа дефекта, ответственного за связанное состояние, поскольку в этом случае становится сомнительной обоснованность понятия диэлектрической проницаемости. Подобное рассмотрение состояний дефектов, конечно, более важно для диэлектриков и примесных полупроводников, чем для металлов.

[58] Свойства металлов, которые определяются электронами проводимости, позволяют выяснить ситуацию для электронов с энергией Ферми, но они не могут дать информации о состояниях с более высокой или более низкой энергией. Однако электрон из области зоны, лежащей ниже энергии Ферми, может быть возбужден в незанятое состояние с более высокой энергией посредством поглощения фотона. Интенсивность такого оптического перехода зависит от природы волновых функций, описывающих верхнее и нижнее состояния, и от плотности состояний в этих двух зонах. Важной характеристикой при этом является *комбинированная плотность состояний*, связывающая верхнее и нижнее энергетические состояния для данной разности энергий (т. е. энергии фотона). Таким образом, интерпретация экспериментальных данных далеко не проста. Сингулярность в оптическом отражении можно ожидать, когда энергия фотона проходит через значение, для которого имеется особенность Ван Хова (см. примечание [46]) у плотности состояний какой-либо из зон, или у комбинированной плотности состояний. Кроме отражения, оптическую связь между состояниями, лежащими ниже ε_F, и состояниями выше $\varepsilon_F + \varphi$, можно также изучать посредством фотоэмиссии (которая обсуждается в разд. 3.3). Фотоэмиссия и оптическое отражение (включая такие их варианты, как пьезоотражение и отражение, модулированное электрическим полем) являются полезными диагностическими средствами изуче-

Для того чтобы свести воедино различные эффекты, связанные с понятием «плотности состояний» и «эффективной массы», следует отметить, что если полный интервал энергий, который перекрывает зона, мал, то:

1) поверхности постоянной энергии должны отстоять в **k**-пространстве далеко друг от друга, т. е. градиент $\nabla_k \varepsilon$ должен быть мал;

2) среднее значение $g(\varepsilon)$ по энергетическому интервалу зоны велико;

3) в такой зоне электроны обладают большой эффективной массой m^*;

4) для таких «тяжелых» электронов ускорение в приложенном извне электрическом поле мало;

5) электронная подвижность в этой узкой зоне должна быть поэтому достаточно малой (подвижность, вероятнее всего, изменяется обратно пропорционально первой или второй степени эффективной массы, показатель степени определяется основным механизмом рассеяния);

6) электроны в такой зоне будут обладать большой электронной теплоемкостью $C_v \sim m^*$;

7) они будут также обладать большой парамагнитной восприимчивостью Паули ($\chi_m \sim m^*$).

В силу таких же рассуждений зона, перекрывающая большой интервал энергий, содержит электронные состояния с малой эффективной массой, и электроны в таких состояниях будут иметь высокую подвижность и т. д.

Скорость в электрическом поле

Если все состояния в энергетической зоне заполнены, то принцип запрета Паули не допускает никакого отклика на внешнее электрическое поле. Для зоны, которая заполнена наполовину, реакция на внешнее воздействие, каким является электрическое поле, определяется состояниями, близкими по энергии к энергии Ферми. Для выяснения того, что *должно* происходить с электроном в электрическом поле, представим себе пустую зону, в которой имеется только один электрон (или заполненную зону с единственной дыркой). Представим себе, что для этого электрона не существует рассеяния, так что

ния металлов, поскольку они могут дать некоторую (по-видимому, не непосредственную) информацию относительно зон, лежащих *выше* и *ниже* энергии Ферми. Кардона [*Cardona M.* Modulation Spectroscopy, Academic Press, 1969] обсуждает какую дополнительную информацию об электронных состояний в твердом теле можно получить, если измерять такую характеристику как коэффициенты отражения, в условиях, когда температура или электрическое поле являются периодической функцией времени.

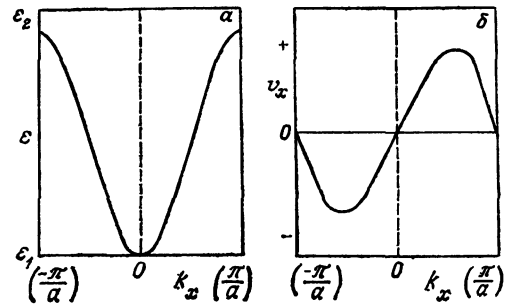


Рис. 3.54. Зависимость энергии электрона (*а*) и скорости электрона (*б*) от приведенного волнового вектора для зонных состояний в одномерном кристалле. Производная v_x по k_x определяется эффективной массой m_1 для области вблизи центра зоны и «дырочной» эффективной массой m_2 для областей вблизи краев зоны.

во внешнем поле **E** он может беспрепятственно перемещаться из одного состояния в другое с отличающейся энергией.

Предположим сначала, что решетка одномерна и периодический потенциал имеет такой вид, что энергия электрона минимальна (например, $\varepsilon=\varepsilon_1$) в центре зоны и максимальна ($\varepsilon=\varepsilon_2$) на границах зоны $k_x=\pm(\pi/a)$. Зависимость $\varepsilon(\mathbf{k})$ для промежуточных энергий изображена на рис. 3.54, *а*. На дне этой зоны

$$\varepsilon=\varepsilon_1+(\hbar^2k_x^2/2m_1), \tag{3.161}$$

а соответствующая скорость электрона равна

$$v_x=\frac{1}{\hbar}\left(\frac{d\varepsilon}{dk_x}\right)=\frac{\hbar k_x}{m_1}, \tag{3.162}$$

то есть она меняется с k_x линейно. Однако, v_x достигает максимального значения (положительного или отрицательного) при некоторой энергии вблизи середины зоны. Для более высоких энергий (значения k_x, расположенные вблизи левой и правой границ зоны) электроны движутся менее быстро. Поскольку вблизи потолка зоны

$$\varepsilon=\varepsilon_2-\frac{\hbar^2}{2m_2}\left[k_x\pm\frac{\pi}{a}\right]^2, \tag{3.163}$$

величина

$$v_x=\frac{-\hbar}{m_2}\left[k_x\pm\frac{\pi}{a}\right] \tag{3.164}$$

описывает поведение скорости вблизи зонной границы. Электрон в состоянии, расположенном *на* краю зоны, оказывается

неподвижным в согласии с предшествовавшим обсуждением дифракции Брэгга и стоячих волн для состояний на краю зоны. Вся кривая скорости для рассматриваемого одномерного кристалла имеет вид, изображенный на рис. 3.54, *б*.

Если рассеяние и внешнее электрическое поле отсутствуют, электрон описывается *одной точкой* в зоне Бриллюэна и может двигаться прямолинейно и с постоянной скоростью сколь угодно долго. При приложении конечного внешнего поля E_x (и все еще без рассеяния) из одномерной формы уравнения (3.144) следует, что k_x должно изменяться со временем согласно уравнению

$$dk_x/dt = -eE_x/\hbar. \tag{3.165}$$

Предположим теперь, что постоянное поле E_x приложено в течение длительного времени. Из уравнения (3.165) следует, что точка в зоне, описывающая наш электрон, будет двигаться с постоянной скоростью влево до тех пор, пока не будет выполняться условие $k_x=-\pi/a$ (а $v_x=0$). Что произойдет при дальнейшем действии поля? Это зависит от того, насколько поле велико.

Туннелирование Зинера и другие виды туннелирования

Если поле очень велико, существует конечная вероятность того, что электрон может осуществить переход через запрещенную щель в состояние с $k_x=-\pi/a$ в следующей, более высоколежащей зоне. Предположим, что запрещенная щель между потолком рассматриваемой зоны и следующей, более высокой зоной составляет $\Delta\varepsilon$. Тогда электрическое поле E_x переводит электрон из нижних состояний при $x=0$ на уровень верхних состояний при $x=(\Delta\varepsilon/eE_x)$ Зинер[59] указал на то, что электрон может *туннелировать* через запрещенную щель в следующую зону при условии, что $(\Delta\varepsilon/eE_x)$ не слишком велико по сравнению с длиной волны де Бройля для электрона и по сравнению с межатомным расстоянием a. Таким образом, вероятность *туннелирования Зинера* изменяется с электрическим полем примерно, как $\exp(-am\Delta\varepsilon^2/4\hbar^2eE_x)$. Для $\Delta\varepsilon=1$ эВ и $a=1$ Å эта вероятность пренебрежимо мала для поля, меньшего 10^7 В/м.

Туннелирование Зинера определяет напряжение пробоя обратносмещенного полупроводникового p—n-перехода, если это резкий переход между сильно легированными областями n-типа и p-типа[60]. (Для более плавного p—n-перехода максимальное

---

[59] *Zener C.*— Proc. Roy. Soc., **A145**, 523 (1934).

[60] См. любую из книг, перечисленных в общей библиографии в разделе «Электронные свойства твердых тел».

обратное напряжение определяется не механизмом Зинера, а лавинным пробоем.) Туннелирование также обусловливает ток при малых напряжениях, приложенных в прямом направлении для полупроводникового *туннельного диода*[61].

Очень большое электрическое поле, способное вызвать туннелирование Зинера, не может быть создано *внутри* металла. В разделе 3.3 мы обсуждали эмиссию во внешнем поле на поверхности металла, зависящую от туннелирования. Вариантами такого процесса является туннелирование через очень тонкий (100 Å или меньше) слой вакуума или диэлектрика из одного металла в другой, особого внимания заслуживают ситуации, в которых металлы с одной или с обеих сторон от диэлектрика находятся в сверхпроводящем состоянии. В разделе 3.6 мы коротко остановимся на теории сверхпроводимости БКШ и успешном предсказании ею малой энергетической щели при энергии Ферми. Малость этой энергетической щели означает, что напряжение, необходимое для того, чтобы вызвать туннелирование между сверхпроводниками, может составлять всего 100 мкВ.

Туннелирование, в котором участвуют состояния выше энергетической щели сверхпроводника в сандвич-структуре металл — диэлектрик — металл, было впервые открыто экспериментально Живером[62]. Позднее Джозефсон[63] указал на то, что если диэлектрический слой в таком сверхпроводящем сандвиче достаточно тонкий, то через поверхность раздела может течь *ток сверхпроводимости* в виде *электронных пар*. Такое туннелирование является основой для ряда применений[64], к которым мы вернемся в конце разд. 3.6.

Отражение Брэгга, вызванное электрическим полем

Вернемся к рассмотрению зоны, содержащей единственный электрон, в условиях, когда нет рассеяния и электрическое поле настолько мало, что туннелирование Зинера отсутствует. Что произойдет, если электрическое поле E_x действует достаточно долгое время, так что k_x (в соответствии с уравнением (3.165)) достигает зонной границы при $k_x = -\pi/a$? Модель, развитая

---

[61] *Esaki L.*— Phys. Rev., **109**, 603 (1957).

[62] *Giaever I.*— Phys. Rev. Lett., **5**, 147 (1960).

[63] *Josephson B. D.*— Phys. Lett., **1**, 251 (1960); Rev. Mod. Phys., **36**, D6 (1964). Нобелевская премия по физике в 1973 г. была разделена между Джозефсоном, Есаки и Живером за их значительный вклад в наше понимание явлений туннелирования.

[64] См., например, *Solymar L.* Superconductivity Tunneling and Applications, Chapman and Hall, 1972. [Имеется перевод: *Солимар Л.* Туннельный эффект в сверхпроводниках и его применение.— М.: Мир, 1974.]

Хоустоном[65], предсказывает, что электрон в этом случае должен испытать рассеяние Брэгга, а затем сразу же оказаться в состоянии внутри той же зоны с $k_x = +\pi/a$. Под действием постоянного электрического поля волновой вектор электрона k_x снова изменяет направление в сторону левой границы зоны.

Модель Хоустона предсказывает, что под действием постоянного электрического поля точка, в которой находится электрон в реальном пространстве, будет двигаться туда и обратно, проходя через нулевую скорость всякий раз, когда $k_x = 0$ или $k_x = \pm(\pi/a)$. Следовательно, единственный не рассеивающийся электрон в пустой зоне при усреднении по большому промежутку времени дает *нулевой* вклад в электропроводность. Вопрос о том, сколько времени продолжается полное колебание в **k**-пространстве и в реальном пространстве, ставится в задаче 3.29, а в следующей задаче рассматривается отклик электрона в такой пустой зоне на поле, меняющееся со временем.

Один из способов описания любопытного результата о нулевом вкладе в электропроводность электрона в пустой зоне за длительный период времени состоит в следующем: говорят, что отклик электрона способствует проводимости, пока его масса положительна, но препятствует ей, когда электрон расположен в той части зоны, где положительной величиной является *масса дырки*.

Рабинович и Зак[66] показали, что представление Хоустона о блоховском электроне, осциллирующем в пустой зоне, является некорректным. Эти авторы указывают, что зонные состояния возмущаются присутствием приложенного электрического поля (эффект Штарка); это возмущение сравнимо с внутризонными поправками к одноэлектронному блоховскому описанию электронных состояний. Конечно, модель Хоустона невозможно как-то имитировать практически, поскольку каждая зона содержит очень много электронов и каждый из них испытывает от 10^9 до 10^{14} соударений в 1 с.

Несмотря на туман сомнений, вызываемых последним абзацем, нам будет удобно продолжать пользоваться уравнением (3.144) для описания движения электрона в **k**-пространстве в многомерном твердом теле при наличии электрического поля. Мы будем также считать, что происходит брэгговское отражение, когда волновой вектор **k** электрона достигает границы зоны.

Как указывалось выше, скорость электрона в многомерном кристалле не обязательно обращается в нуль, если состояние, занимаемое электроном, находится на зонной границе, однако

---

[65] *Houston W. V.*— Phys. Rev., **57**, 184 (1940).
[66] *Rabinovitch A., Zak J.*— Phys. Lett., **40A**, 189 (1972).

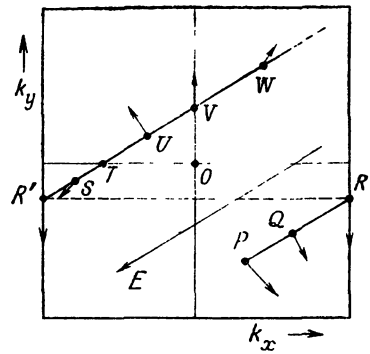


Рис. 3.55. Влияние электрического поля **E** на положение электрона в незаполненной зоне в последовательные моменты времени (в предположении, что рассеяние отсутствует). Для каждого электронного состояния, начиная с состояния P и далее, длина и направление стрелки показывают скорость, которую должен иметь электрон в реальном пространстве для соответствующего набора поверхностей постоянной энергии. Отражение Брэгга происходит, когда k_x достигает зонной границы; в этом случае $v_x=0$. В отсутствие рассеяния будут происходить последующие отражения Брэгга.

в нуль обращается перпендикулярная к границе компонента скорости. На рис. 3.55 показаны точки в **k**-пространстве, последовательно занимаемые электроном в двумерной решетке под действием электрического поля **E** в отсутствие рассеяния. Поле включается, когда электрон находится в состоянии, которому соответствует точка P. Скорость электрона в реальном пространстве в этот момент времени показана стрелкой, выходящей из P. Длина и направление стрелки определяются величиной $\nabla_k\varepsilon$ согласно формуле (3.141).

Под действием электрического поля **E** волновой вектор электрона перемещается сначала в Q, затем в R согласно уравнению (3.144). Если мы примем точку зрения Хоустона о брэгговском отражении на границе зоны, то в результате отражения электрона в точке R радиус-вектор электрона в **k**-пространстве последовательно пройдет положения от R' до W и т. д. Движение в реальном пространстве определяется изменяющимися значениями $\nabla_k\varepsilon$ в проходимых электроном областях **k**-пространства. Таким образом, в одни моменты времени электрон набирает энергию от внешнего поля, в другие электрон отдает свою энергию полю.

Движение электрона в реальном трехмерном твердом теле, происходящее под действием электрического поля, не соответствует картине, которую мы сейчас изобразили, в результате действия принципа Паули и в результате рассеяния. Принцип Паули запрещает электрону под действием поля перемещаться в состояния, лежащие ниже энергии Ферми, а рассея-

ние на дефектах или фононах как в металлах, так и в неметаллических твердых телах обычно прерывает вызванное полем движение **k** и изменение связанных с **k** характеристик, когда электрон пройдет расстояние, составляющее незначительную часть ширины зоны. Чем меньше плотность фононов и чем совершеннее кристалл, тем больше вероятность наблюдать перемещение электрона по конечной части зоны. В соответствии с этим исследования формы зон и формы поверхности Ферми обычно выполняются при низких температурах на тщательно ориентированных и чистых монокристаллах.

Циклотронный резонанс

Согласно уравнениям Максвелла, магнитное поле, действующее на электрон, стремится изменить направление движения электрона, не изменяя его энергии. Это следует из формулы (3.35) для силы Лоренца. Таким образом, магнитная индукция B_z оказывает влияние на движение в плоскости xy, не изменяя движения в направлении z. Если электрон не рассеивается, то он описывает в плоскости xy некоторую орбиту, движение по которой накладывается на любое движение в направлении z.

Квазисвободный электрон со скалярной массой m^* описывает круговую орбиту радиусом r, по которой электрон движется с угловой частотой ω_c. Связь между этими величинами определяется условием равенства центробежной силы ($m^*\omega_c^2 r$) и уравновешивающей ее силы Лоренца ($r\omega_c eB_z$). Таким образом, угловая *циклотронная частота* равна

$$\omega_c = eB_z/m^*; \tag{3.166}$$

она не зависит от кинетической энергии электрона. (От энергии зависит размер орбиты в реальном пространстве, поскольку $\varepsilon = m^*\omega_c^2 r^2/2$.) Циклотронная частота для обычно применяемых магнитных полей лежит в радио- и микроволновой области электромагнитного спектра, так как

$$\nu_c \equiv (\omega_c/2\pi) = 28{,}0\,(B_z m/m^*)\ \text{ГГц} \tag{3.167}$$

для магнитной индукции, выраженной в теслах.

Под действием магнитного поля движение электрона в реальном пространстве сопровождается прецессией в **k**-пространстве по траектории с *постоянной энергией* в зоне Бриллюэна. Конечно, для очень сильно вырожденного электронного газа в металле это движение наблюдается только для электронов с энергией Ферми, т. е. для электронов, которые описывают в **k**-пространстве орбиты вокруг поверхности Ферми. Поскольку *какое-то* рассеяние электронов на фононах и дефектах неизбежно даже в почти идеальном кристалле при низких темпе-

ратурах, отчетливо выраженное циклотронное движение может быть получено только при условии $(\omega_c\tau_m)>1$, т. е. когда электрон может пройти значительную часть своей магнитной орбиты до того, как он будет рассеян.

Бóльшая часть электронов с энергией Ферми имеет отличную от нуля компоненту импульса, параллельную B_z. Эти электроны описывают в k-пространстве круговую траекторию с радиусом, меньшим радиуса ферми-сферы. Их траектория в реальном пространстве складывается из движения по окружности в плоскости xy и прямолинейного движения в направлении z. Однако некоторые электроны с энергией Ферми обладают нулевой компонентой импульса в z-направлении. Под действием поля B_z эти электроны должны двигаться по экваториальной траектории (по «большому кругу») вокруг сферы Ферми, а их движение в реальном пространстве также является чисто круговым — на него не налагается никакое прямолинейное движение. Такая экваториальная орбита вокруг сферы Ферми представляет собой простейший вид *экстремальной орбиты* — того класса орбит, который очень важен в экспериментах по *циклотронному* резонансу. Даже когда форма поверхности Ферми далека от сферической, существуют определенные экстремальные траектории, которые могут быть определены и использованы для характеристики топологии поверхности.

Теперь должно быть очевидно, что сферическая поверхность Ферми может быть обнаружена в металле только в силу случайных обстоятельств. Гораздо более типична ситуация, когда магнитное поле B_z заставляет электроны с энергией Ферми двигаться в k-пространстве вокруг поверхности Ферми по траектории, вдоль которой эффективная масса непрерывно изменяется. Тогда скорость, с которой волновой вектор меняется со временем, непостоянна; это ясно уже из того, что магнитная сила, действующая на электрон, равна $\hbar(d\mathbf{k}/dt)$ и также равна $-e(\mathbf{v}\times\mathbf{B})$. В результате скорость движения электрона по орбите в реальном пространстве не постоянна, что исследуется для простого случая в задаче 3.32.

В экспериментах по *циклотронному резонансу* используется поглощение электромагнитной энергии на радиочастоте ω, когда магнитная индукция $\mathbf{B}$ подобрана таким образом, что $\omega=\omega_c$. Тогда использование различных комбинацией ω и $\mathbf{B}$ позволяет (в принципе) получить информацию относительно тензора эффективной массы для электрона с энергией Ферми. Фактическая теория циклотронного резонанса гораздо более сложна как для полупроводников, так и для металлов.

Для полупроводникового материала, в котором плотность свободных электронов мала, эксперименты по циклотронному резонансу могут быть выполнены с электромагнитными вол-

нами, проникающими в твердое тело. Трудности, которые при этом возникают, связаны с топологией поверхностей постоянной энергии и с гибридными плазменными резонансами[67] в том случае, когда концентрация свободных электронов не слишком мала.

Частоты, используемые для исследования циклотронного резонанса в металле, всегда гораздо меньше плазменной частоты (поскольку концентрация электронов в металле настолько велика, что и частота ω_p становится очень большой). Как было указано в разд. 3.3, для $\omega<\omega_p$ вещественная часть диэлектрической проницаемости отрицательна. В соответствии с этим металл для таких частот непрозрачен и глубина проникновения δ (толщина скин-слоя) гораздо меньше толщины образца. В этом случае от средней длины свободного пробега электрона λ зависит, чем будут определяться электрические характеристики поверхности для электромагнитных волн радиодиапазона: *нормальным скин-эффектом* или *аномальным скин-эффектом*. Первый случай осуществляется при $\lambda<\delta$, а второй при $\lambda>\delta$.

В последнем случае можно возбудить циклотронное движение, комбинируя действие постоянной магнитной индукции (например, B_z) и высокочастотного электромагнитного поля; при этом используется геометрия, предложенная Азбелем и Канером[68] (рис. 3.56). Названные авторы указали, что если постоянная магнитная индукция B_z лежит в плоскости поверхности, то циклотронное движение должно происходить в плоскости, пересекающей поверхность. Некоторые циклотронные орбиты при этом достигают области высокочастотного скин-слоя, как показано на рисунке. Электроны, движущиеся по таким орбитам, приближаясь к поверхности, могут испытывать действие высокочастотного поля с угловой частотой ω и приобретать энергию от поля или отдавать ее полю в зависимости от соотношения между ω и циклотронной частотой ω_c. Таким образом, поверхностный импеданс кристалла по отношению к высокочастотному излучению является функцией величины магнитной индукции.

---

[67] Согласно зонной теории выражение (3.107) для плазменной частоты заменяется выражением, содержащим эффективную массу: $\omega_p=(ne^2/m^*\varepsilon_0)$ Поэтому, если концентрация электронов в полупроводнике не настолько мала, чтобы выполнялось неравенство $\omega_p\ll\omega_c$, максимальное взаимодействие электромагнитного поля с модами электронного движения будет происходить не на самой частоте ω_c, а на верхней гибридной частоте $(\omega_c^2+\omega_p^2)^{1}$ Влияние плазменного резонанса на измерение циклотронного резонанса обсуждается в работе: *Dresselhaus G., Kip A. F., Kittel C.*— Phys. Rev., **100**, 618 (1955).

[68] *Азбель М. Я., Канер Е. А.*— ЖЭТФ, 1956, т. 3, с. 772; J. Phys. Chem. Sol., **6**, 113 (1958).

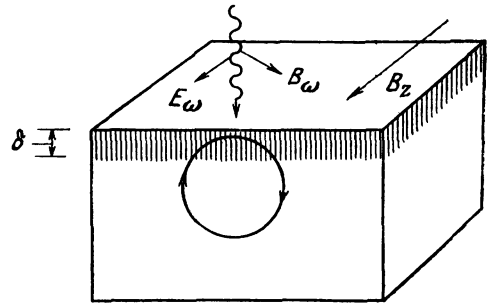


Рис. 3.56. Геометрия Азбеля — Канера для наблюдения циклотронного резонанса в металлическом кристалле. Заштрихован скин-слой, имеющий глубину δ для высокочастотного излучения с частотой ω. Показана одна из возможных орбит, проходящих через поверхностный слой. Такая орбита может соответствовать циклотронному движению, возникающему под действием магнитной индукции B_z, приложенной в плоскости поверхности. Наблюдение резонанса Азбеля — Канера должно проводиться на металлическом монокристалле высокой частоты и совершенства, одна грань которого [например, (100) или (111)] обработана с особой тщательностью, чтобы при низких температурах среднее время свободного пробега (а следовательно, средняя длина свободного пробега) было велико как в объеме кристалла, так и в скин-слое. Энергия высокочастотного поля может быть связана с энергией кругового движения электронов при условии $\lambda > \delta$. Если при этом также $\omega_c \tau_m \gg 1$, то может наблюдаться острый циклотронный резонанс, когда частота ω равна или кратна ω_c.

Для успешного наблюдения резонансных явлений следует работать с чистым совершенным монокристаллом при низких температурах, чтобы средняя длина свободного пробега была велика по сравнению с размером циклотронной орбиты. Поверхность, на которую падает высокочастотное излучение, должна быть хорошего качества, чтобы значение λ в приповерхностном слое было таким же, как в объеме. В этих условиях значение λ будет большим по сравнению с толщиной скин-слоя δ и движущийся по окружности электрон будет взаимодействовать с высокочастотным полем только в течение малой доли своего периода обращения. Азбель и Канер указали, что при $\lambda > \delta$ и $(\omega_c \tau_m) \gg 1$ взаимодействие между высокочастотным полем и циклотронным движением может быть обеспечено как при $\omega = \omega_c$, так и при значении ω, достаточно малом кратном ω_c. Пусть B_c — магнитная индукция, при которой $\omega = \omega_c$. Для магнитной индукции, составляющей целую долю от B_c, интервал между двумя последовательными попаданиями данного электрона в поверхностный слой равен нескольким периодам высокочастотного поля. Однако и в этом случае высокочастотное поле сможет повторить свое воздействие на электрон в тот момент, когда он снова окажется у поверхности. Азбель

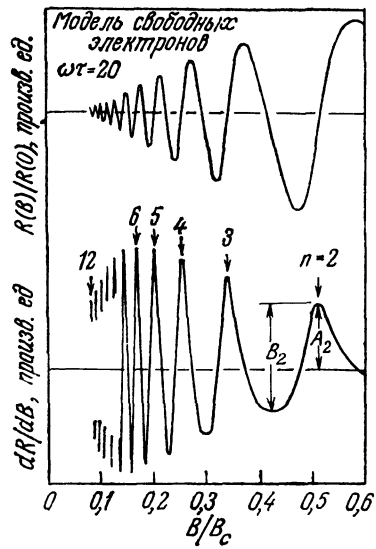


Рис. 3.57. Зависимость поверхностного сопротивления (вещественной части поверхностного импеданса) для свободного электронного газа в металле при частоте высокочастотного поля ω от индукции $\mathbf{B}$ (верхняя кривая). По оси абсцисс отложена нормированная величина B/B_c, где $B_c=\omega m/e$ — индукция, для которой ω и циклотронная частота совпадают. Эта кривая может быть рассчитана по формуле (3.168) модели Азбеля — Канера. Следует заметить, что производная dR/dB (нижняя кривая, имеющая осциллирующий ход) достигает максимума всякий раз, когда индукция B составляет целую долю от B_c. Это показывает, что эксперимент по циклотронному резонансу удобно строить таким образом, чтобы в нем измерялась и записывалась непосредственно величина dR/dB как функция магнитной индукции при фиксированной частоте высокочастотного поля [*Häussler P., Welles S. J.*— Phys. Rev., 152, 675 (1966)].

и Канер установили, что зависимость комплексного поверхностного импеданса от магнитной индукции определяется выражением

$$Z(B)=Z_0\,[1-\exp(-2\pi/\omega\tau_m)\exp(-2\pi i\omega_c/\omega)]^{1/3}, \qquad (3.168)$$

где магнитная индукция входит в величину ω_c. Осциллирующее поведение вещественной части этого импеданса (поверхностного сопротивления) показано на рис. 3.57. Там же показан ход производной (dR/dB) — величины, которую можно измерять непосредственно в эксперименте.

На рис. 3.58 приведены для примера результаты экспериментального наблюдения резонанса Азбеля — Канера на очень чистом образце меди при низких температурах. Различие двух кривых показывает, как важно, чтобы рассеяние электронов

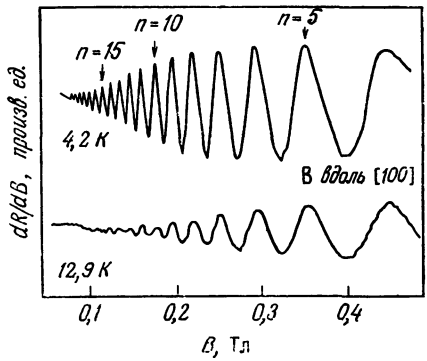


Рис. 3.58. Результаты экспериментального наблюдения резонанса Азбеля — Канера в кристалле чистой меди при двух температурах. Кривая для более высокой температуры сглажена из-за возросшего теплового рассеяния движущихся по циклотронной орбите электронов. Поверхность кристалла представляет собой плоскость (110), магнитное поле, направленное вдоль [100], лежит в этой плоскости. Наблюдается резонанс для электронов, движущихся по экстремальной «поясной орбите», охватывающей основной объем поверхности Ферми (см. рис. 3.51) [*Häussler P., Welles S. J.*— Phys. Rev. **152**, 675 (1966)].

было сведено к минимуму. Кривую, снятую при 4,2K, можно непосредственно сравнить с предсказаниями теории Азбеля — Канера и определить из нее размер орбиты для электронов с энергией Ферми в меди. Для такой ориентации полей, при которой были получены данные на рис. 3.58, важна электронная «поясная» орбита (belly orbit), когда электроны движутся в $\mathbf{k}$-пространстве почти по круговой траектории вокруг основного обхвата поверхности Ферми, показанной на рис. 3.51. Поясная орбита является *экстремальной*; она *максимизирует* циклотронный период; точно так же «шеечная орбита» вокруг шейки, показанной у границы зоны на рис. 3.51 и 3.52, *экстремальна* в том смысле, что она *минимизирует* циклотронный период по сравнению с соседними орбитами.

Особая важность экстремальных орбит связана с тем, что электроны, прецессирующие по орбитам, лежащим на несферической поверхности Ферми, обладают в данном магнитном поле множеством периодов. Однако вклады электронов с неэкстремальных орбит взаимно компенсируются из-за различия фаз. Основной вклад дает экстремальная область, в которой первая производная периода по компоненте $\mathbf{k}$, направленной вдоль магнитного поля, обращается в нуль. Эта область ответственна за значительный сигнал, находящийся в фазе.

Проблема электронных орбит в металле в случае, когда величина магнитной индукции достаточно велика, интенсивно раз-

рабатывается с середины 50-х годов и заслуживает гораздо более подробного рассмотрения. Заинтересованного читателя мы отсылаем к библиографии в конце главы; в перечисленных работах он найдет обсуждение таких вопросов, как «открытые орбиты» (экстремальные орбиты, охватывающие несколько повторяющихся зон), размерные резонансные эффекты (Гантмахер, 1964) и другие интересные результаты, позволившие так радикально улучшить наши представления о зонной структуре металлов.

Магнитные подзоны

Наше обсуждение парамагнитной восприимчивости газа свободных электронов (разд. 3.3) было основано на представлении о том, что заселенность состояний со спинами, антипараллельными магнитному полю, увеличивается за счет ее уменьшения для состояний с параллельными спинами (т. е. заполнение состояний с магнитным моментом, параллельным полю, увеличивается за счет уменьшения заполнения состояний с магнитным моментом, антипараллельным этому направлению). Было показано, что магнитная восприимчивость уменьшается за счет диамагнетизма [69]. Теперь вернемся к этому вопросу. При этом, обсуждая орбитальное циклотронное движение, мы (временно) пренебрежем существованием электронного спина.

Для модели свободных электронов (т. е. модели, для которой потенциальная энергия во всем пространстве постоянна) в отсутствие магнитного поля справедливо уравнение Шредингера

$$(-\hbar^2/2m)\nabla^2\psi = \varepsilon\psi. \tag{3.169}$$

Мы знаем, что решения этого уравнения представляют собой плоские волны, волновые векторы которых связаны с собственными значениями энергии соотношением

$$\varepsilon = \hbar^2k^2/2m. \tag{3.170}$$

Однако известно и то, что магнитная индукция B_z будет оказывать воздействие на движение каждого электрона в плоскости xy, оставляя невозмущенным его движение в z-направле-

[69] Диамагнитная организация электронных состояний была впервые описана Л. Ландау [*Landay L.*— Z. Phys., **64**, 629 (1930)]. Он рассматривал «свободный электронный газ» без всякого периодического потенциала. Пайерлс [*Peierls R. E.*— Z. Phys., **80**, 763 (1933)] показал, что электроны в зоне, полученной в модели сильной связи, обладают в сильном магнитном поле аналогичными свойствами. Затем Латтинджер [*Luttinger J. M.*— Phys. Rev., **84**, 814 (1951)] указал, что результат Ландау можно использовать для любой зоны, в которой энергия квадратично зависит от волнового вектора, просто подставляя вместо обычной массы электрона эффективную массу m^*.

нии. Сила Лоренца превращает движение электрона в сумму прямолинейного движения в направлении z и циклотронного движения в плоскости xy. Таким образом, каждый электрон движется по спиральной траектории.

Ландау описал это, показав, что оператор Гамильтона в уравнении Шредингера при наличии конечного магнитного поля содержит два дополнительных члена. С учетом магнитной индукции B_z уравнение (3.169) было записано Ландау в более общем виде:

$$\left\{(-\hbar^2/2m)\nabla^2+(m\omega_c^2/8)(x^2+y^2)-\right.$$
$$\left.-(i\hbar\omega_c)\left[x\left(\frac{\partial}{\partial y}\right)+y\left(\frac{\partial}{\partial x}\right)\right]\right\}\psi=\varepsilon\psi, \qquad (3.171)$$

где магнитная индукция выражена через циклотронную частоту свободных электронов $\omega_c=eB_z/m$. Уравнение (3.171) накладывает на движение электрона в плоскости xy ограничение, которое можно записать в следующем виде:

$$x=x_0+[(2p+1)\hbar/m\omega_c]^{1/2}\cos(\omega_c t),$$
$$y=y_0+[(2p+1)\hbar/m\omega_c]^{1/2}\sin(\omega_c t), \qquad (3.172)$$

где p может быть любым целым числом (включая нуль). Этот набор возможных видов движения в плоскости xy представляет собой набор состояний двумерного квантового гармонического осциллятора.

Для данного значения квантового числа p циклотронное движение (в реальном пространстве) происходит по траектории с радиусом

$$r_p=[(2p+1)\hbar/m\omega_c]^{1/2}. \qquad (3.173)$$

Отождествляя k_x с $(m/\hbar)(dx/dt)$ и k_y с $(m/\hbar)(dy/dt)$, можно видеть, что электронная прецессия в **k**-пространстве должна удовлетворять условию

$$(k_x^2+k_y^2)=(2p+1)m\omega_c/\hbar. \qquad (3.174)$$

Таким образом, все состояния, соответствующие данному значению квантового числа p, попадают в **k**-пространстве на цилиндрическую поверхность радиусом

$$k_p=[(2p+1)m\omega_c/\hbar]^{1/2}. \qquad (3.175)$$

Эта ситуация, представленная на рис. 3.59, коренным образом отличается от однородного распределения электронов в **k**-пространстве, с которым мы имели дело до сих пор. Совокупность электронных состояний, соответствующих данному значению квантового числа p, называется *магнитной подзоной*, или *уровнем Ландау*.

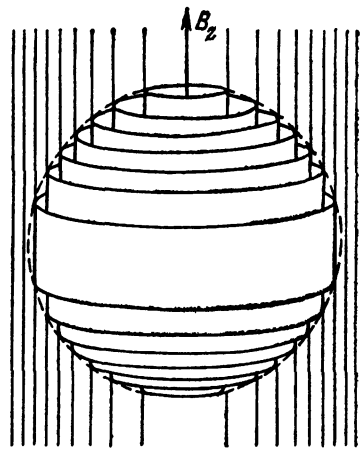


Рис. 3.59. Влияние магнитной индукции B_z на области k-пространства, которые в соответствии с моделью Ландау могут быть заняты электронами. Состояния, разрешенные условием (3.174), образуют ряд бесконечных коаксиальных цилиндров, параллельных B_z. Граница заполнения, устанавливаемая энергией Ферми, если пренебречь деталями, похожа на сферу Ферми в отсутствие поля. Практически края цилиндров смазаны из-за электронного рассеяния [*Chambers R. G.*— Canad. J. Phys., 34, 1395 (1956)].

Собственные энергии модифицированного Ландау уравнения Шредингера (3.171) равны

$$\varepsilon = (\hbar^2 k_z^2/2m) + \left(p + \frac{1}{2}\right)\hbar\omega_c. \qquad (3.176)$$

В этой полной энергии первый член справа представляет собой собственную энергию движения свободной частицы в направлении приложенного поля (не зависящую от величины этого поля), а второй член — это энергия циклотронного движения в плоскости xy, которое описывается формулами (2.172) — (3.175).

Теперь следует обсудить влияние магнитного поля на функцию плотности состояний $g(\varepsilon)$. Однако какой бы ни стала эта функция, энергия Ферми, как известно, может быть определена условием

$$n = \int_0^\infty g(\varepsilon)\left[1 + \exp\left(\frac{\varepsilon - \varepsilon_F}{k_0 T}\right)\right]^{-1} d\varepsilon. \qquad (3.177)$$

Известно также, что состояния ниже ε_F заполнены, а состояния выше ε_F пусты. На рис. 3.60 показана зависимость полной энергии электрона (3.176) от k_z для нескольких первых подзон, когда напряженность магнитного поля настолько велика, что все прочие уровни Ландау расположены полностью выше энергии Ферми. Вспомним, что для каждого уровня Ландау разрешены только те значения k_x и k_y, которые лежат на цилиндрической поверхности (3.174), коаксиальной с осью k_z; тогда становится ясно, что для зависимостей $\varepsilon(k_z)$, изображенных на

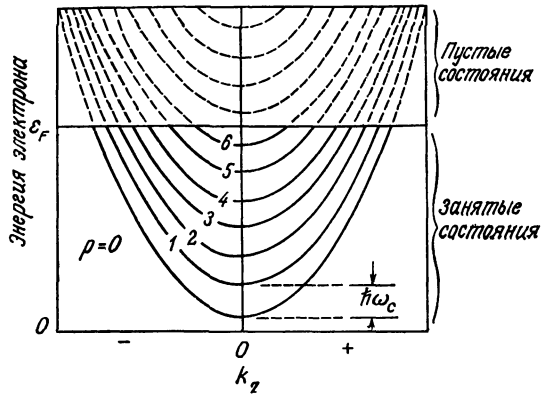


Рис. 3.60. Энергия электрона как функция k_z для первых семи магнитных подзон при приложении магнитной индукции B_z. Из формулы (3.176) видно, что для каждой подзоны минимальная энергия составляет $(p+1/2)\hbar\omega_c = (p+1/2)(e\hbar B_z/m)$. Предполагается, что энергия Ферми такова, что все подзоны выше седьмой совершенно пусты. Распределение заполненных состояний в **k**-пространстве для этого случая показано на рис. 3.59.

рис. 3.60, должны быть заполнены области **k**-пространства, показанные на рис. 3.59.

В слабом магнитном поле (или при малой подвижности электронов) магнитной перестройкой плотности состояний можно пренебречь, поскольку появление различимых уровней Ландау возможно только при $\omega_c\tau_m \gg 1$. Когда значение B_z достаточно велико, чтобы вызвать магнитное перераспределение плотности состояний $g(\varepsilon)$, не следует считать, что в присутствии магнитного поля зона содержит меньше электронных состояний, чем без него. Каждый уровень Ландау вмещает столько состояний, сколько их находится на площади, равной $2\pi m\omega_c/\hbar$, в плоскости k_xk_y в металле без поля. Роль магнитного поля состоит в том, что оно собирает группу из $m\omega_c/\hbar$ состояний (в единице объема), которые первоначально обладали в плоскости xy энергиями, распределенными в интервале от $p\hbar\omega_c$ до $(p+1)\hbar\omega_c$, причем энергия каждого состояния из этой группы становится равной $(p+1/2)\hbar\omega_c$. Процесс такой группировки схематически показан на рис. 3.61. Чем больше магнитная индукция, тем больше состояний находится в каждой группе и тем меньше групп (т. е. уровней Ландау) попадает внутрь заданного интервала энергии $(\varepsilon - \hbar^2k_z^2/2m)$.

Если рассматривать плотность состояний в достаточно большом интервале энергий или достаточно большой области **k**-пространства, то она оказывается слабо зависящей от магнитного поля, однако особенности ее распределения безусловно изменяются. Именно этим объясняется диамагнетизм электронного

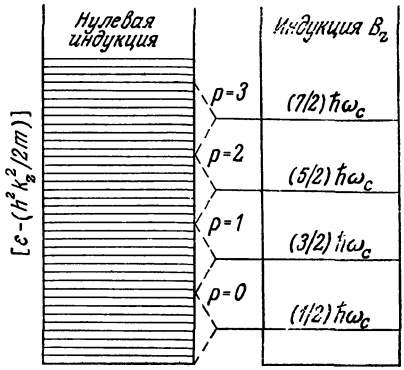


Рис. 3.61. Магнитная перестройка распределения электронных состояний по энергии. По оси ординат отложена энергия, связанная с движением в плоскости, перпендикулярной полю. Магнитная индукция B_z вынуждает группы из $m\omega_c/\hbar$ состояний (на единицу объема кристалла) собираться в сильно вырожденные состояния (подзоны, уровни Ландау) с энергиями $(p+1/2)\hbar\omega_c$. Каждое из этих состояний соответствует одному из цилиндров в **k**-пространстве, изображенных на рис. 3.59. Области **k**-пространства, расположенные между цилиндрами, перестают быть разрешенными.

газа и различные эффекты, которые обнаруживают зависимость от B_z. Для энергетического спектра (3.176) плотность состояний как функция энергии имеет вид

$$g(\varepsilon)=\frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sum_{p=0}^{p_{\text{макс}}}\frac{\hbar\omega_c}{\left[\varepsilon-\left(p+\frac{1}{2}\right)\hbar\omega_c\right]^{1/2}}, \qquad (3.178)$$

где каждый уровень Ландау соответствует одному из членов суммы. Значение $p_{\text{макс}}$ определяется тем условием, что знаменатель должен быть вещественной величиной. Чем больше рассматриваемая энергия, тем больше уровней Ландау возникает под этой энергией.

Как следует из выражений (3.177) и (3.178), полная плотность электронов в присутствии магнитного поля любой величины, как большого, так и малого, может быть выражена в довольно неудобном виде

$$n=\frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\times$$

$$\times\sum_{p=0}^{p_{\text{макс}}}\int_0^{\infty}\frac{\hbar\omega_c d\varepsilon}{\left[1+\exp\left(\frac{\varepsilon-\varepsilon_F}{k_0T}\right)\right]\left[\varepsilon-\left(p+\frac{1}{2}\right)\hbar\omega_c\right]^{1/2}}. \qquad (3.179)$$

Это выражение можно записать следующим образом:

$$n=\left[\frac{m^3k_0T\omega_c^2}{2\pi^4\hbar^4}\right]^{1/2}\sum_{p=0}^{p_{\text{макс}}}F_{-1/2}(x_p/k_0T), \tag{3.180}$$

где

$$x_p=\left[\varepsilon_F-\left(p+\frac{1}{2}\right)\hbar\omega_c\right], \tag{3.181}$$

а $F_{-1/2}(x_p/k_0T)$ — интеграл из семейства интегралов Ферми — Дирака, введенных в этой главе выше [12, 13].

Если магнитная индукция мала по сравнению с $(\varepsilon m/e\hbar)$, то суммирование в формуле (3.178) можно заменить интегрированием (как предлагается в задаче (3.33). Тогда соотношения между плотностью состояний, плотностью электронов и энергией Ферми принимают вид (3.44) и (3.46), обычный для квантовой теории свободных электронов. Конечно, если магнитная индукция достаточно мала, величина $(\omega_c\tau_m)$, по-видимому, становится меньше единицы и тепловое уширение энергетических уровней также будет способствовать смазыванию магнитных изменений в расположении энергетических уровней и их заполнении.

Если же значение B_z велико, то выражения (3.178)—(3.180) можно рассматривать как суммы с ограниченным числом членов: тогда $g(\varepsilon)$ обращается в бесконечность для каждой из энергий $(p+1/2)\hbar\omega_c$, представляющих собой минимальные значения *полной* энергии для каждого нового уровня Ландау. Разрывный характер функции $g(\varepsilon)$ продемонстрирован на рис. 3.62. В этом случае свойства электронного газа зависят от энергии Ферми ε_F, удовлетворяющей условию нормировки (3.179), от того, как расположена энергия Ферми по отношению к минимальным энергиям ближайших к ней уровней Ландау и от соотношения между энергиями ε_F и $\hbar\omega_c$ и шириной теплового распределения k_0T.

Чтобы оценить величину диамагнетизма электронного газа, заметим, что уровень Ландау объединяет состояния с энергиями движения в плоскости xy, находящимися в интервале от $p\hbar\omega_c$ до $(p+1)\hbar\omega_c$, в одну вырожденную группу с энергией $(p+1/2)\hbar\omega_c$. Средняя энергия, приходящаяся на электрон, такой перестройкой не изменяется, если все эти состояния заняты. Однако для уровня Ландау, квантовый номер которого не намного меньше, чем $(\varepsilon_F/\hbar\omega_c)$, магнитная перестройка, показанная на рис. 3.61, приводит к объединению сильно заполненных состояний из нижней части этого энергетического интервала с более слабо заполненными состояниями из верхних частей интервала. В результате средняя энергия, приходящаяся на

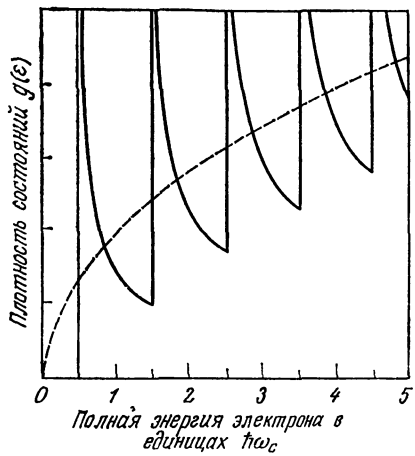


Рис. 3.62. Плотность состояний, определяемая формулой (3.178), как функция энергии при наличии большой магнитной индукции. Заметим, что ниже энергии $\hbar\omega_c/2$ плотность состояний $g(\varepsilon)$ равна нулю и что при этой энергии при всех последующих значениях энергии, равных нечетному кратному $\hbar\omega_c/2$, наблюдается неограниченный рост $g(\varepsilon)$. Бесконечное значение $g(\varepsilon)$, предсказываемое для нижней энергии каждой подзоны, практически становится конечным в результате теплового рассеяния. Штриховой кривой показана обычная плотность квантовых состояний (3.44) в отсутствие магнитной индукции.

одно *занятое* электронное состояние, для самых верхних подзон увеличивается в магнитном поле. Полная энергия электронов равна

$$n\langle\varepsilon\rangle = \frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} \times$$

$$\times \sum_{p=0}^{p_{\text{макс}}} \int_0^\infty \frac{\varepsilon\hbar\omega_c d\varepsilon}{\left[1+\exp\left(\dfrac{\varepsilon-\varepsilon_F}{k_0T}\right)\right]\left[\varepsilon-\left(p+\dfrac{1}{2}\right)\hbar\omega_c\right]^{1/2}}, \qquad (3.182)$$

где n дается формулой (3.179). Увеличение средней энергии $\langle\varepsilon\rangle$ по сравнению с ее значением $(3\varepsilon_F/5)$, полученным выше (см. формулу (3.55)) для поля, равного нулю, может быть описано с помощью диамагнитной восприимчивости. Как было показано в разд. 3.3, Ландау установил, что диамагнитная восприимчивость

$$\chi_{\text{диам}} = -(n\mu_0\hbar^2\omega_c^2/8\varepsilon_F B_z^2) = -(n\mu_0\mu_B^2/2\varepsilon_F) \qquad (3.183)$$

компенсирует одну треть парамагнитной восприимчивости, обусловленной электронным спином.

Магнитная восприимчивость, определяемая формулой (3.183), не зависит от B_z. Это следует из выражений (3.179) и (3.182) при условии, что B_z не слишком велико (или температура не слишком мала). Это замечание сделано нами потому, что не зависящая от поля диамагнитная восприимчивость Ландау (3.183) вычислена в предположении, что больцмановский хвост распределения Ферми простирается на интервал k_0T, захватывающий несколько магнитных подзон. Это соответствует случаю, когда величина $(\hbar\omega_c) \equiv (2\mu_B B_z)$ мала по сравнению с k_0T. Для очень сильного поля (или очень низкой температуры) распределение Ферми обрывается очень резко в пределах одной подзоны; средняя энергия электрона в этом случае зависит от точного соотношения между энергией Ферми и энергиями ближайших к ней верхнего и нижнего уровней Ландау. Изменение магнитного поля приводит к осциллирующему изменению магнитной восприимчивости — явлению, об экспериментальном наблюдении которого впервые сообщили де Гааз и ван Альфен (1931 г.).

Происхождение эффекта де Гааза — ван Альфена (дГвА) легко понять, рассмотрев разрывную плотность состояний, изображенную на рис. 3.62. По мере увеличения магнитной индукции B_z набор уровней Ландау, изображенных на рис. 3.59 в виде цилиндров в **k**-пространстве, расширяется наружу. Энергии этих подзон на рис. 3.60—3.62 движутся вверх. Для ряда критических значений B_z пороговые энергии магнитных подзон совпадают с энергией Ферми (как следует из формул (3.179) или (3.180)). Таким образом, вследствие разрывности плотности состояний энергия Ферми является осциллирующей функцией магнитного поля. На рис. 3.63 показаны осцилляции ε_F как функции B_z для температуры, по предположению равной $T = (\varepsilon_F/20k_0)$. В правой части этого рисунка может быть заполнен только нижний уровень Ландау.

Здесь следует еще раз вспомнить, что модель Ландау в ее первоначальном виде (и как она была изложена выше) — это модель квантового газа *свободных* электронов. Если тензор эффективной массы соответствует почти сферической поверхности Ферми в металле в отсутствие магнитного поля, то модель Ландау можно использовать без изменений; нужно только всюду заменить массу свободного электрона эффективной массой m^*. Для металла, в котором поверхность Ферми коренным образом отличается от сферической, дело обстоит сложнее, однако представление о магнитных подзонах остается чрезвычайно полезным для понимания того, что в нем происходит.

Для каждого из критических значений **B** на рис. 3.63 площадь самой удаленной от центра орбиты в **k**-пространстве (т. е. площадь сечения самого внешнего цилиндра на рис. 3.59)

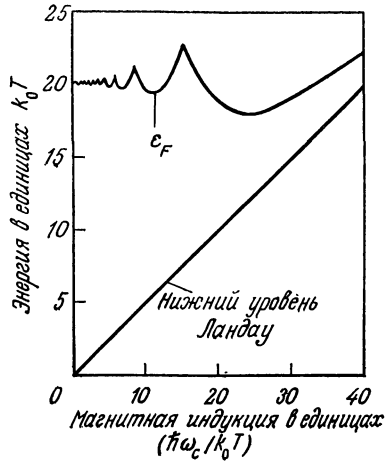


Рис. 3.63. Рассчитанный пример зависимости энергии Ферми от магнитной индукции для вырожденного электронного газа. Выбор температуры $T = (\varepsilon_F/20k_0)$ обеспечивает выполнение условия слабого поля. Величины, отложенные по оси ординат и абсцисс, выражены в безразмерных единицах. Следует отметить, что пикообразные максимумы энергии Ферми будут на самом деле слегка скруглены из-за электронного рассеяния (*Blakemore J. S.* Semiconductor Statistics.— Pergamon, 1962. [Имеется перевод: *Блейкмор Дж.* Статистика электронов в полупроводниках.— М.: Мир, 1964.])

равна площади экстремальной орбиты, соответствующей энергии Ферми. Необходимое условие состоит в том, что энергия Ферми должна совпадать с наименьшей возможной энергией для одного из уровней Ландау. Это условие не реализуется для фиксированных интервалов **B**, как ясно видно из рис. 3.63. Однако критическое условие оказывается периодическим по величине $1/B$. Осциллирующая часть магнитной восприимчивости в эффекте дГвА определяется осциллирующей частью средней энергии (3.182), которая имеет максимумы для той же линейной последовательности значений $1/B$. Установить эту последовательность значений можно следующим образом.

В присутствии магнитной индукции B_z площадь S орбиты в плоскости k_xk_y в k-пространстве может иметь одно из следующего ряда значений:

$$S = (2\pi m^*\omega_c/\hbar)\left(p + \frac{1}{2}\right) = (2\pi eB_z/\hbar)\left(p + \frac{1}{2}\right). \qquad (3.184)$$

Эта площадь равна умноженному на π квадрату радиуса, определяемого формулой (3.174). Максимальное возможное значение p (а следовательно, и S) должно соответствовать энергии Ферми и нулевой компоненте движения в z-направлении.

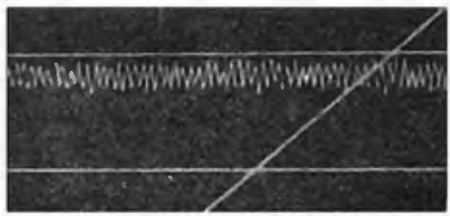


Рис. 3.64. Экспериментально полученная запись осциллирующей компоненты магнитной восприимчивости (эффект дГвА) в калии при температуре 1,07 К. Диагональная линия дает изменение магнитной индукции, которая в пределах рисунка меняется от 12,69 до 13,28 Тл. Расстояние между последовательными максимумами может быть использовано для определения площади поверхности Ферми в k-пространстве (см. задачу 3.34) [*Thorsen A. C., Berlincourt T. G.*— Phys. Rev. Lett, 6, 617, (1961)].

Из формул (3.40) и (3.49) следует, что радиус ферми-сферы в k-пространстве равен $k_F = (3\pi^2 n)^{1/3}$. Круг такого радиуса имеет экстремальную площадь

$$S_F = \pi (3\pi^2 n)^{2/3}. \tag{3.185}$$

Таким образом, самая большая орбита, допускаемая соотношением (3.184), будет точно соответствовать энергии Ферми для сильно вырожденного газа всякий раз, когда

$$1/B_z = (2\pi e/\hbar S_F)\left(p + \frac{1}{2}\right). \tag{3.186}$$

Это условие выполняется каждый раз, когда обратная магнитная индукция изменяется на величину

$$P \equiv \Delta (1/B_z) = 2\pi e/\hbar S_F = 9{,}55 \cdot 10^{15}\ S_F\ \text{Тл}^{-1}, \tag{3.187}$$

где сечение поверхности Ферми S_F выражено в $м^{-2}$. Таким образом, измеряя период P (в единицах $Тл^{-1}$) между последовательными максимумами дГвА, можно определить площадь сечения поверхности Ферми и с помощью (3.185) найти полную концентрацию электронов.

Простой пример экспериментальных данных по эффекту дГвА приведен на рис. 3.64. Эти данные для щелочного металла калия можно проанализировать в рамках изотропной модели свободных электронов, как предлагается в задаче 3.34. Для твердого тела, в котором эффективная масса анизотропна, орбиты в k-пространстве не будут круговыми, однако *площадь* орбиты все равно будет определяться соотношением (3.184) и условие (3.187) и в этом случае может быть использовано для определения S_F. Измерение периода дГвА для различных ориентаций можно использовать для определения размера и формы поверхности Ферми. Рассмотрим теперь результаты ис-

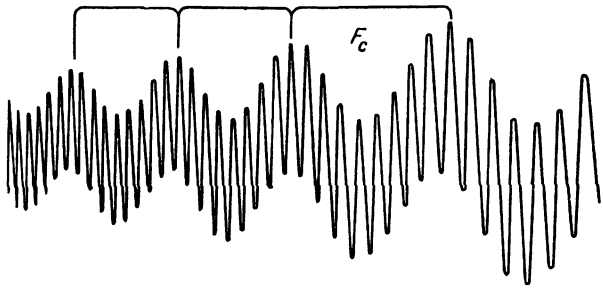


Рис. 3.65. Пример более сложной записи эффекта дГвА для металлического серебра [*Joseph A. S., Thorsen A. C.* — Phys. Rev., **138**, A1159(1965)]. Магнитное поле составляет угол $\theta=48°$ с направлением куба [001] в плоскости (110), угол $\sim 7°$ с направлением [111]. Поверхность Ферми серебра похожа на поверхность Ферми меди (см. рис. 3.51 и 3.52) с «шейками», касающимися границы зоны в направлениях [111] k-пространства.

следования эффекта дГвА в металлическом серебре, показанные на рис. 3.65. Поверхность Ферми серебра похожа на поверхность Ферми меди, показанную на рис. 3.51, 3.52. Наименьшая орбита на поверхности Ферми — это орбита вокруг ее «шейки» там, где поверхность Ферми достигает границы зоны Бриллюэна в одной из областей, эквивалентных [$^1/_2$ $^1/_2$ $^1/_2$]. Колебания магнитной восприимчивости с коротким периодом, показанные на рис. 3.65, соответствуют такой «шеечной» орбите. На эти колебания восприимчивости накладывается более медленная модуляция, которая приписывается биениям, возникающим из-за наличия двух различных типов орбит вокруг основного «пояса» квазисферы Ферми. Для всех металлов, кроме самых простейших, интерпретация данных по эффекту дГвА чрезвычайно сложна, и несколько случаев ее успешного осуществления, накопленных к настоящему времени, стоят в ряду самых значительных достижений в экспериментальной физике твердого тела за 1950—1960-е годы.

Кроме осциллирующей компоненты магнитной восприимчивости, в вырожденном электронном газе обнаружено много других эффектов[70], характеризующихся осциллирующей зависимостью от $(1/B)$ при $(\omega_c\tau_m)\gg 1$. Одна из таких зависимостей показана в очень простом виде на рис. 3.66. В качестве примера взят эффект Шубникова — де Гааза (ШдГ), впервые обнаруженный в 1930 г. Л. Шубниковым и В. Дж. де Гаазом в полуметалле висмуте. Эффект ШдГ заключается в осциллирующей

[70] Различные явления с осциллирующими зависимостями описаны в работе: *Kahn A. H., Frederikse H. P. R.*, Solid State Physics, v. 9, eds. F. Seitz, D. Turnbull, Academic Press, 1959.

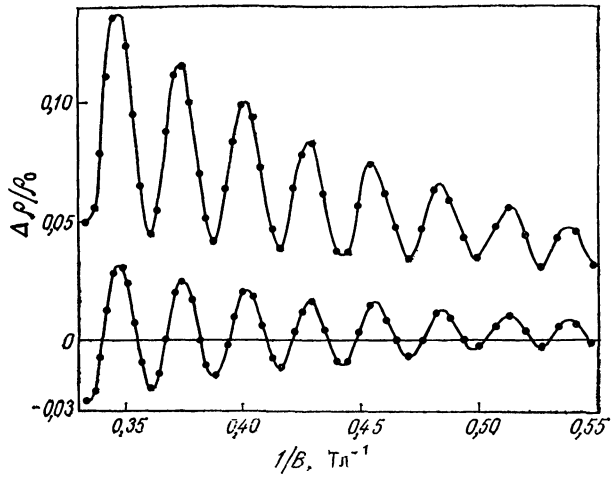


Рис. 3.66. Простой пример эффекта Шубникова — де Гааза [*Becker W. M., Fan H. Y.* Proceedings 7th, International Semiconductor Conference, Dunod, Paris, 1964]. Относительное магнетосопротивление ($\Delta\rho/\rho_0$) представлено как функция обратной магнитной индукции для полупроводникового образца GaSb n-типа при 4,2 К. Ток и магнитное поле перпендикулярны для верхней кривой и параллельны для нижней кривой. По величине постоянной Холла $R_H = -4{,}8 \cdot 10^{-6}$ м³/Кл можно вычислить полную плотность электронов (задача 3.35) и сравнить ее с величиной, рассчитанной по периоду ШдГ.

зависимости электрического сопротивления от магнитной индукции (осцилляции накладываются на монотонно изменяющееся магнетосопротивление, которое для некоторых твердых тел и некоторых ориентаций может быть гораздо больше осциллирующего члена). Как и в эффекте дГвА, осцилляции в ШдГ-эффекте обусловлены движением уровней Ландау относительно энергии Ферми.

Для многих металлов и полупроводников со сложной поверхностью Ферми интерпретация эффекта ШдГ так же сложна, как и эффекта дГвА или результатов других методов исследования. Данные, приведенные на рис. 3.66, отличаются такой простотой потому, что они получены для зонных электронов (в полупроводнике, а не в металле), для которых тензор эффективной массы сводится к простому скаляру. На изотропность эффективной массы указывает то обстоятельство, что период P верхней кривой (поперечное магнетосопротивление) и нижней кривой (продольное магнетосопротивление) одинаков. Расчеты, предлагаемые в задаче 3.35, подтверждают, что период эффекта ШдГ для кристалла, которому соответствуют кривые на рис. 3.66, согласуется со значением полной электронной концентрации, полученной по величине постоянной Холла.

В гл. 4 мы рассмотрим осциллирующую компоненту магнетосопротивления слабо легированного полупроводника с *невырожденным* электронным газом. Появление этой компоненты носит название *магнетофононного резонанса* и обусловлено взаимодействием структуры уровней Ландау, медленных электронов на первом уровне Ландау и оптических фононов с энергией, соответствующей забросу электрона на более высокий уровень Ландау. Для магнетофононного резонанса необходима высокая подвижность электронов, поэтому обычно он не может конкурировать с другими процессами в вырожденном электронном газе металла или сильно легированного полупроводника.

3.6. Сверхпроводимость *

Сверхпроводящее состояние

Можно сказать, что современная физика низких температур началась с успешного сжижения гелия, осуществленного в Лейдене Камерлинг-Оннесом (1908 г.). Оннес сразу же использовал жидкий гелий для многочисленных исследований металлов при низких температурах и в 1911 г. обнаружил, что электрическое сопротивление чистой ртути резко падает при охлаждении ниже 4,2 К. К его удивлению, добавление к ртути примесей не приводило при самых низких температурах к появлению остаточного сопротивления, и к 1913 г. Оннес пришел к выводу, что «ртуть перешла в новое состояние, которое в силу исключительности его электрических свойств можно назвать *сверхпроводящим состоянием*». Рис. 3.67 воспроизводит поведение электросопротивления в окрестности T_c — *температуры перехода в сверхпроводящее состояние* (или *критической* температуры), о котором сообщил Оннес.

Прежде чем переходить к более подробному описанию явлений, связанных со сверхпроводимостью, которое будет дано в настоящем разделе, укажем, что сверхпроводник обладает бесконечной электропроводностью на постоянном токе, однако выше некоторой частоты обнаруживает нормальное сопротивление. Аномалия в изменении электронной теплоемкости при переходе в сверхпроводящее состояние (при отсутствии скрытой теплоты перехода) показывает, что этот переход электронов в более упорядоченное состояние представляет собой переход второго рода (в термодинамическом смысле). Сверхпроводник

* Уже после выхода в свет английского издания этой книги была открыта высокотемпературная сверхпроводимость, о которой мы кратко упомянули в предисловии.— *Прим. ред.*

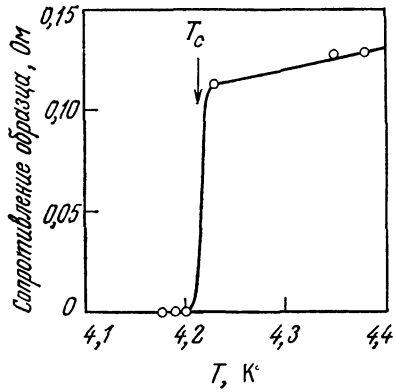


Рис. 3.67. Переход ртути из нормального состояния в сверхпроводящее, наблюдавшийся Оннесом [*Onnes H. K.* — Leiden Comm., 124c (1911)]. Согласно данным Оннеса, ниже температуры перехода T_c сопротивление меньше 10^{-5} Ом, что соответствует электропроводности не меньшей, чем 10^{15} $Ом^{-1} \times м^{-1}$. По современной температурной шкале T_c для ртути составляет 4,15 К.

является идеальным диамагнетиком и полностью выталкивает магнитный поток. Приложение магнитного поля, достаточно сильного, чтобы проникнуть в металл, разрушает сверхпроводимость. Все эти явления в течение полувека не находили теоретического объяснения, хотя некоторые свойства и могли быть систематизированы в рамках термодинамической и электродинамической теорий. Природа электронной перестройки была объяснена микроскопической теорией, развитой в 1957 г. Бардином, Купером и Шриффером [71]. Мы будем, как это принято, сокращенно называть ее теорией сверхпроводимости БКШ.

Согласно теории БКШ взаимодействие между электронами посредством виртуального обмена фононами может привести к тому, что их *полная* энергия (кинетическая энергия плюс потенциальная) станет меньше полной энергии независимых электронов в распределении Ферми при условии, что при нулевой плотности тока электроны занимают пары состояний с противоположными направлениями импульса и спина. Конечный ток, однажды возникнув, будет поддерживаться всем электронным распределением, однородно смещенным от точки $\mathbf{k}=0$. Чтобы изменить такое электронное распределение, системе необходимо сообщить конечное количество энергии (равное ширине энергетической щели), что весьма маловероятно.

[71] *Bardeen J., Cooper L. N., Schrieffer J. R.* — Phys. Rev., 108, 1175 (1957). Это классическое изложение очень сложных представлений теории БКШ. В книгах Линтона, Купера и Тинккама, приведенных в библиографии, эти представления излагаются более подробно.

Таким образом, теория БКШ не только объясняет, почему в сверхпроводнике «не затухает» ток, но позволяет также понять и другие свойства сверхпроводника и объясняет эксперименты, указывающие на существование энергетической щели в спектре электронов.

Возникновение сверхпроводимости

Сверхпроводимость представляет собой широко распространенное явление[72]. К моменту написания этих строк существование сверхпроводящей фазы было обнаружено у 25 химических элементов. Температуры перехода для некоторых из этих элементов приведены в табл. 3.4. Многие сплавы и интерметаллические соединения также оказались сверхпроводниками. Несколько примеров таких соединений также дано в табл. 3.4. Бинарное соединение может быть сверхпроводником даже в том случае, когда один из составляющих его элементов (или оба!) не является сверхпроводником. Как было предсказано Коэном[73], сверхпроводимость была обнаружена у полупроводников с большой концентрацией свободных электронов.

Возможно, что любой чистый элемент ниже некоторой достаточно низкой критической температуры становится сверхпроводником, хотя может понадобиться много времени, чтобы подтвердить или опровергнуть такое утверждение. У одновалентных щелочных и благородных металлов, исследованных до температур ниже 0,1 К, не было обнаружено каких-либо признаков сверхпроводимости. Вполне возможно, что эти металлы должны быть исключены из числа сверхпроводников, так как электрон-фонон-электронное взаимодействие, приводящее при низких температурах к сверхпроводимости, должно сочетаться с сильным электрон-фононным взаимодействием (т. е. с невысокой электропроводностью) при более высоких температурах.

Сверхпроводимость отсутствует также у ферромагнитных металлов. Хорошо известно, что магнитные примеси оказывают разрушающее действие на сверхпроводимость того твердого

---

[72] Обширная сводка экспериментальных данных, накопленных по сверхпроводимости к середине 50-х годов, содержится в обзорной статье: *Serin B.*— In: Handbuch der Physik, v. 15, Springer, 1956. Критерии возникновения сверхпроводимости даются в работе: *Matthias B. T., Geballe T. H., Compton V. B.*— Rev. Mod. Phys., 35, 1, (1963). См. также: *Matthias B. T.* Progress in Low Temperature Physics, v. 2, ed. C. J. Gorter, Wiley-Interscience, 1957.

[73] Теоретическое предсказание сделано в работе: *Cohen M. L.*— Phys. Rev., 134, 511 (1964). Имеются экспериментальные данные о сверхпроводимости (с очень малым значением T_c) ряда полупроводников, в которых может быть создана высокая электронная плотность и в которых имеет место сильное взаимодействие между электронами и колебаниями решетки.

Таблица 3.4. **Температуры перехода в сверхпроводящее состояние для ряда элементов, соединений и сплавов***

Элемент	T_c, К	Соединение	T_c, К
Технеций	11,2	Nb_3Ge	23,2
Ниобий	9,46	Nb_3Ga	20,3
Свинец	7,18	Nb_3Sn	18,05
Тантал	4,48	NbN	16,0
Ртуть	4,15	Mo_3Ir	8,8
Индий	3,41	NiBi	4,25
Алюминий	1,19	AuBe	2,64
Кадмий	0,56	$PdSb_2$	1,25
Титан	0,40	TiCo	0,71
Иридий	0,14	$AuSb_2$	0,58
Вольфрам	0,01	$ZrAl_2$	0,30

* Данные взяты из работы: *Matthias B. T., Geballe T. H., Compton V. B.* — Rev Mod. Phys., 35, 1 (1963). Данные для Nb_3Ga — из работы: *Webb G. W. et al.* — Solid. State Comm., 9, 1769 (1971), а для Nb_3Ge — из работы: Physics Today, October 1973.

тела, в которое они вводятся. Это легко согласовать с удивительными магнитными свойствами сверхпроводника. Маттиас[72] вывел эмпирические правила для возникновения сверхпроводимости, согласно которым предпочтительными оказываются такие твердые тела, в которых число валентных электронов на атом равно по крайней мере двум, но не превышает девяти. Маттиас связывает также высокую критическую температуру переходных элементов с нечетным числом валентных электронов и обсуждает причину отсутствия сверхпроводимости у всех редкоземельных элементов, следующих за лантаном. Критическая температура T_c коррелирует также с атомной массой и с межатомными расстояниями и конфигурациями атомов. Все это сложным образом связано с микроскопической теорией сверхпроводимости БКШ.

Проводимость на постоянном и переменном токе

Если проводимость сверхпроводника на постоянном токе действительно бесконечно велика, то ток, возникший в замкнутом витке сверхпроводящего материала, должен существовать бесконечно долго без всякой электродвижущей силы. Если же сверхпроводящий виток с индуктивностью L обладает *ко-*

нечным сопротивлением R, то текущий по нему ток должен медленно спадать в соответствии с формулой

$$I = I_0 \exp(-Rt/L). \tag{3.188}$$

Попытки определить скорость затухания таких токов начались с первых исследований сверхпроводимости. Оннес и Тайн (1924 г.) пришли к выводу, что электропроводность свинца в сверхпроводящем состоянии составляет не менее 10^{18} Ом$^{-1}\cdot$м$^{-1}$, в то время как для свинца высокой чистоты при температуре, немного превышающей T_c, она равна 10^{11} Ом$^{-1}\cdot$м$^{-1}$. Многочисленные эксперименты, проведенные после этого (один из них обсуждается в задаче 3.36), отодвинули нижний предел критической температуры. По данным Квина и Итнера[74] электропроводность на постоянном токе составляет не менее 10^{25} Ом$^{-1}\cdot$м$^{-1}$.

Важно, однако, подчеркнуть, различие между *сверхпроводником*, т. е. твердым телом с бесконечной электропроводностью на постоянном токе, и *идеальным проводником*, т. е. твердым телом, в котором электроны находятся в блоховских состояниях, не испытывая рассеяния. Бардин (1956 г.) указал, что целесообразнее рассматривать сверхпроводник как предельный случай диамагнетика, чем как предельный случай проводника с бесконечно большой электропроводностью. При $T=0$ электропроводность бесконечна только до некоторой предельной частоты, а при конечной температуре (лежащей ниже T_c) небольшие потери существуют на переменном токе на всех частотах.

Отсутствие сверхпроводимости на оптических частотах было обнаружено на самых ранних этапах исследования (визуально сверхпроводник непрозрачен!), а измерения, выполненные в 1930-х годах, показали, что сверхпроводимость исчезает при частоте, превышающей 10^9 Гц, но меньшей, чем 10^{14} Гц. Более современные исследования в далекой инфракрасной области позволили определить точную частотную зависимость в этом интересном диапазоне для большого числа сверхпроводников. Типичный результат по поглощению в тонкой пленке показан на рис. 3.68. Порог наступает при энергии фотона порядка $4k_0T_c$, что должно быть связано с величиной энергетической щели между сверхпроводящим и «нормальным» состояниями электрона. Существование такой щели было постулировано Ф. Лондоном (1935 г.) как неотъемлемая часть его феноменологической теории сверхпроводимости, и оно лежит в основе микроскопической теории БКШ.

---

[74] *Quinn D. J., Ittner W. B.*— J. Appl. Phys., **33**, 748 (1962).

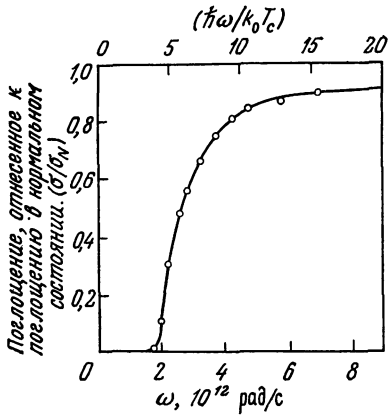


Рис. 3.68. Частотная зависимость поглощения электромагнитных волн в тонкой сверхпроводящей пленке индия. Ордината нормирована на величину поглощения в нормальном состоянии. Заметим, что выше угловой частоты, примерно равной $(4k_0T_c/\hbar)$, сверхпроводимость отсутствует [*Ginsberg D. M., Tinkham M.*— Phys. Rev., **118**, 990 (1960)].

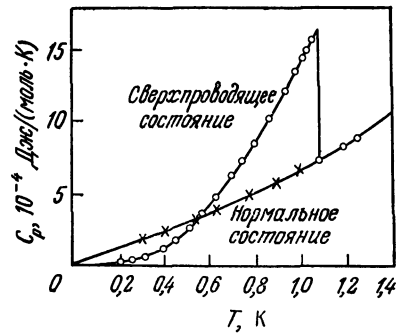


Рис. 3.69. Температурная зависимость удельной теплоемкости галлия в нормальном и сверхпроводящем состояниях [*Phillips N.* — Phys. Rev., 134, 385 (1964)]. Результаты для галлия в нормальном состоянии ниже температуры T_c были получены при приложении магнитного поля 0,02 Тл.

Теплоемкость

С переходом в сверхпроводящее состояние в металле не связано никакой теплоты перехода, однако имеется аномалия в ходе электронной составляющей теплоемкости. Пример такого аномального хода показан на рис. 3.69. Разрыв удельной теплоемкости указывает на переход второго рода из относительно разупорядоченного (нормального) состояния в более упорядоченное (сверхпроводящее) состояние, обладающее меньшей энтропией. Аналогичная особенность теплоемкости имеет место, когда жидкий ^{4}He становится сверхтекучим. Общепринято рассматривать сверхпроводящий металл как сверхтекучую среду, состоящую из бозонов, которые, согласно теории БКШ, представляют собой пары электронов с противоположно направленными импульсами и спинами. То, что сверхпроводимость представляет собой сугубо квантовое явление, в котором волновые функции электронов согласованы на макроскопических расстояниях, было очевидно уже из ранних феноменологических теорий.

Полная удельная теплоемкость металла в «нормальном» состоянии представляет собой сумму, состоящую из члена, про-

порционального T^3, обусловленного колебаниями решетки (дебаевского члена), и члена, прямо пропорционального абсолютной температуре, отражающего температурную зависимость энергии фермиевского электронного газа. Если использовать «нормальные» данные для галлия, приведенные на рис. 3.69, то эта сумма равна

$$C_p = 6\cdot 10^{-5}T^3 + 6\cdot 10^{-4}T \text{ Дж}\cdot\text{моль}^{-1}\text{К}^{-1}. \quad (3.189)$$

Вклад фононов в сверхпроводящем состоянии остается неизменным, однако зависимость от температуры электронной составляющей становится гораздо более сильной. В некоторых сверхпроводящих металлах электронная теплоемкость $C_{эл}$ изменяется при низких температурах примерно как T^3, однако у галлия, как и у некоторых других сверхпроводников, наблюдается экспоненциальный ход теплоемкости:

$$C_p = (6\cdot 10^{-5}T^3 + 9{,}5\cdot 10^{-3}\exp(-1{,}9/T)\text{ Дж}\cdot\text{моль}^{-1}\text{К}^{-1}. \quad (3.190)$$

Эта экспоненциальная зависимость означает, что при абсолютном нуле все электроны в сверхпроводнике находятся в полностью упорядоченном состоянии с нулевой энтропией и теплоемкостью, однако электроны могут быть термически переброшены через энергетическую щель (ширины порядка $4k_0T_c$) в «нормальные» состояния системы. В соответствии с такой двухжидкостной моделью доля электронов, находящихся в сверхпроводящем (или сверхтекучем) состоянии, будет падать, обращаясь в нуль при некоторой температуре — критической температуре T_c. Такое представление содержится и в теории БКШ, но в более сложной форме, так как в модели БКШ *сама* ширина щели уменьшается с ростом температуры.

Сверхпроводимость и магнитные поля

Во время своих ранних исследований Оннес обнаружил, что сверхпроводящее состояние можно разрушать обратимым образом, пропуская через проволоку сильный ток. Вскоре стало ясно, что причиной возвращения к нормальной проводимости служит магнитное поле, возникающее при протекании тока. Сверхпроводимость может существовать только, если внешнее магнитное поле не превышает некоторого критического значения H_c, величина которого изменяется от некоторого максимального значения H_0 при $T=0$ до нуля при критической температуре. Для многих сверхпроводников эта зависимость приближенно описывается формулой

$$H_c = H_0[1-(T/T_c)^2], \quad (3.191)$$

которую часто называют *законом Тайна* (1924 г., 1929 г.) (Tuyn's law). Зависимость (3.191) представлена сплошной

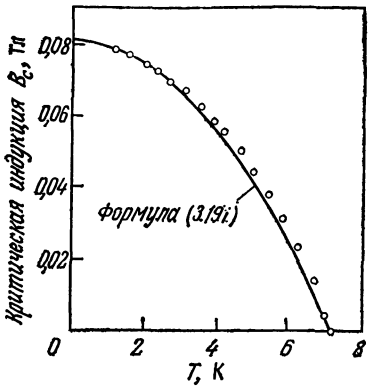


Рис. 3.70. Температурная зависимость магнитного поля, необходимого для разрушения сверхпроводимости в свинце (T_c=7,19 K, B_0=0,0803 Тл). Кружками обозначены экспериментальные точки [*Decker D. et al.*— Phys. Rev, **112**, 1888 (1958)]. Сплошная кривая соответствует закону Тайна (3.191). Отклонение экспериментальных точек от этой зависимости показано в увеличенном масштабе на рис. 3.71.

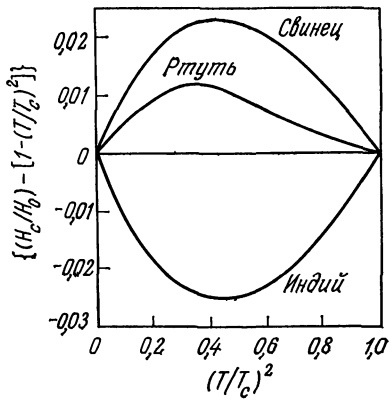


Рис. 3.71. Представленное в крупном масштабе расхождение между законом Тайна (3.191) и результатами экспериментального определения температурной зависимости критического магнитного поля. Для свинца изображено расхождение между сплошной кривой и экспериментальными точками на рис. 3.70. Заметим, что расхождение может иметь любой знак, однако отрицательное отклонение, которое наблюдается в случае индия, лучше согласуется с теорией БКШ [*Bardeen J., Schrieffer J. R.*— Progr. Low. Temp. Phys, 3, 170 (1961)].

кривой на рис. 3.70 вместе с экспериментальными данными для свинца, которые удовлетворительно согласуются с кривой. [Поскольку магнитные поля обычно указываются в литературе не в А/м, а в гауссах, эрстедах или теслах, представляется удобным на рис. 3.70 и нескольких следующих рисунках и формулах описывать критическое состояние через критическую *индукцию* $B_c=\mu_0 H_c$ (Тл), а не через напряженность приложенного поля H_c (А/м).]

Однако данные для свинца на рис. 3.70 не совсем точно следуют закону Тайна. Это расхождение наглядно видно из рис. 3.71 благодаря выбору координат. На этом рисунке приведены также для сравнения данные для ртути и индия. Кривые показывают, что для более точного описания зависимости H_c от T необходимо представить ее в виде полинома.

На рис. 3.71 использованы безразмерные переменные: магнитное поле в единицах H_0, а температура в единицах T_c. Эти и другие безразмерные переменные часто используются при срав-

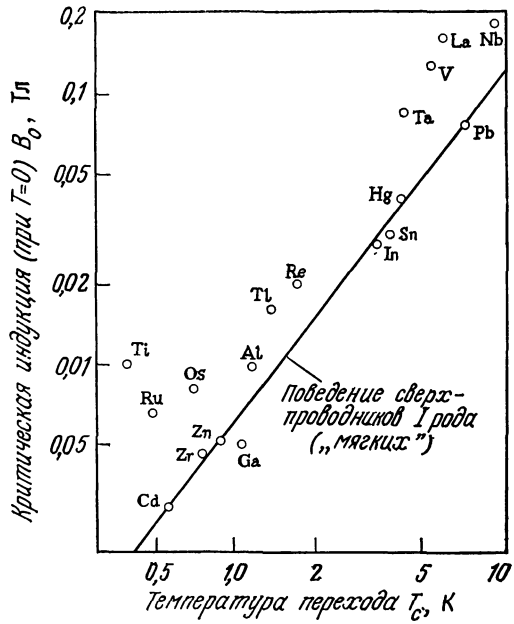


Рис. 3.72. Критическая магнитная индукция B_0, разрушающая сверхпроводящее состояние при $T=0$, представленная как функция T_c для элементарных сверхпроводников (Lynton).

нении данных для различных металлов. Многие свойства сверхпроводящего состояния были систематизированы с помощью безразмерных систем координат еще до появления микроскопической теории явления.

Критическое поле при нулевой температуре (H_0 или B_0) обычно больше для материалов с высокой критической температурой, однако, как можно видеть из рис. 3.72, результаты для различных элементов отличаются значительным разбросом. Этот разброс становится гораздо меньше, если разделить сверхпроводники на две группы в соответствии с кривыми их диамагнитного отклика и отнести линейную зависимость, изображенную на рис. 3.72, только к «мягким» сверхпроводникам, или *сверхпроводникам I рода*. Это те сверхпроводящие элементы, которые находятся в группе клеток, расположенной в правой части периодической таблицы элементов; именно эти сверхпроводники наиболее удовлетворительно описываются теорией. «Жесткие» сверхпроводники, или *сверхпроводники II рода*, к которым относятся группы переходных металлов и актинидов, а также бóльшая часть сверхпроводящих соединений, имеют более сложную структуру, что отражается на их

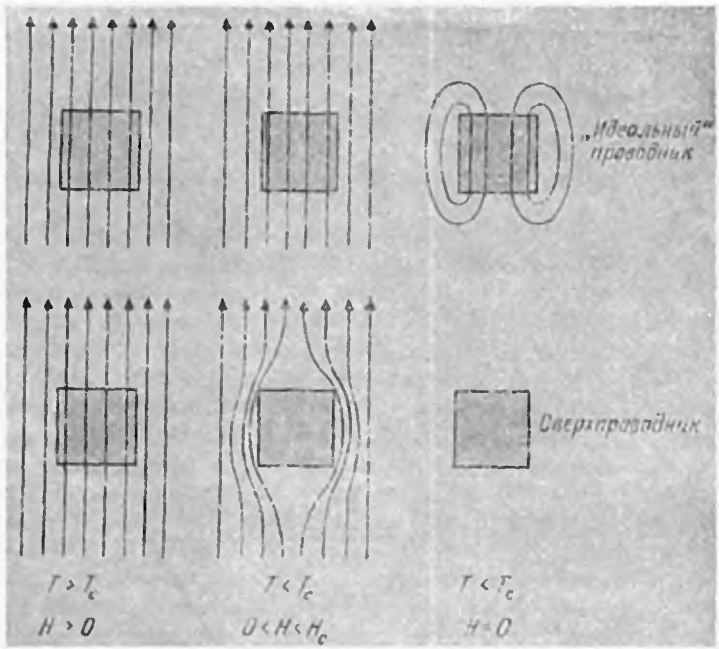


Рис. 3.73. Иллюстрация различия между воображаемым «идеальным проводником» (вверху) и реальным сверхпроводником (внизу). Как показано Мейсснером и Оксенфельдом (1933 г.), сверхпроводник выталкивает магнитный поток из своей внутренней области.

магнитных свойствах. Прежде чем обсуждать разницу в магнитном поведении сверхпроводников I и II рода, следует описать предполагаемое взаимодействие идеального проводника с магнитным полем.

Если бы можно было поместить идеальный проводник во внешнее магнитное поле, то силовые линии магнитного потока совсем не проникали бы в образец. Наведенные поверхностные токи противодействовали бы внешнему полю. Однако находящийся в поле нормальный проводник (верхняя левая схема на рис. 3.73), который станет «идеальным проводником», когда на него действует поле, должен вести себя, как показано в центральной верхней схеме, а после удаления внешнего поля (верхняя правая схема) должны возникнуть незатухающие токи, которые замкнут внутренний магнитный поток. Таким образом, конечное состояние охлажденного «идеального» проводника» будет зависеть от предшествовавших ему состояний и переходы не будут обратимыми.

Другая ситуация имеет место в сверхпроводнике, как было показано Мейснером и Оксенфельдом (1933 г.) путем измерений магнитного поля, окружающего металлическую сферу. Они установили, что состояние сверхпроводника не зависит от действовавшего на него до перехода магнитного поля; при охлаждении металла ниже температуры перехода в сверхпроводящее состояние он выталкивает наружу все магнитные силовые линии. (Если внешнее поле больше H_c, то магнитный поток не выталкивается, потому что тогда металл *не станет* сверхпроводящим.) Все силовые линии, вытолкнутые из внутренней части металла, должны быть скомпенсированы незатухающим наведенным поверхностным током. Выключение внешнего поля наводит поверхностный ток противоположного направления, который уничтожает первоначальный ток, так что в сверхпроводнике не остается ни поля, ни тока. Благодаря *эффекту Мейснера,* состоящему в выталкивании магнитного потока, переход между сверхпроводящим и нормальным состояниями является обратимым.

При обсуждении эффекта Мейснера важно обратить внимание на то, что магнитная индукция **B** не может существовать во *внутренней части* образца. В отличие от электрического поля, которое совсем не может проникать в сверхпроводник, магнитное поле может проникать в сверхпроводник на глубину λ, которая обычно имеет порядок 10^{-8}—10^{-7} м. В соответствии с феноменологической теорией электродинамики сверхпроводящего состояния, развитой Г. и Ф. Лондонами (1935 г.), глубина проникновения определяется формулой

$$\lambda = (m\varepsilon_0 c^2/e^2 n_s)^{1/2}, \tag{3.192}$$

где n_s — плотность электронов, находящихся в сверхтекучем состоянии. При нулевой температуре в таком состоянии находятся все электроны, однако по мере приближения температуры к T_c число сверхпроводящих электронов падает до нуля, делая возможным бесконечно глубокое проникновение поля. Измерение глубины проникновения поля показывает, что она изменяется с температурой приблизительно следующим образом [75]:

$$(\lambda/\lambda_0) = [1-(T/T_c)^4]^{-1/2}, \tag{3.193}$$

что соответствует предсказаниям термодинамической модели сверхпроводников, развитой Гортером и Казимиром (1934 г.).

---

[75] Некоторые из самых ранних измерений глубины проникновения поля в сверхпроводник были выполнены Понтиусом [*Pontius R. B.*— Phyl. Mag., **24**, 787 (1937)]. Исследованию температурной зависимости λ в олове посвящена работа: *Laurmann E., Shoenberg D.*— Proc. Roy. Soc., **A198**, 560 (1949).

[В связи с формулой (3.193) еще раз укажем на удобство использования безразмерных координат при обсуждении явлений, связанных со сверхпроводимостью.]

Квантование магнитного потока и незатухающие токи

Выталкивание магнитного потока из сверхпроводящего тела, конечно, имеет место в *односвязном* теле, не содержащем дырок или их топологических эквивалентов. Предположим теперь, что эксперимент, которому соответствует рис. 3.73, выполнен на полом металлическом цилиндре. Охлаждение ниже температуры T_c приведет к тому, что магнитный поток будет вытолкнут из металла, но останется во внутренней полости цилиндра. Эти силовые линии *окажутся захваченными,* когда внешнее поле будет выключено, и будут поддерживаться незатухающим электрическим током, циркулирующим в окружающем их цилиндре. Величина захваченного магнитного потока уменьшается по мере спада этого незатухающего тока.

Ф. Лондон[76] на основании того, что сверхпроводимость представляет собой квантовое явление, предположил, что захваченный магнитный поток должен быть квантован с размером кванта $2\pi\hbar/e = 4{,}14 \cdot 10^{-15}$ Вб. Чтобы установить квантовый характер магнитного потока, необходимо было бы проводить эксперименты на металлических цилиндрах с очень малой площадью поперечного сечения. Конечно, квантование магнитного потока требует квантования полного незатухающего тока (см. задачу 3.37).

Измерения того типа, который был предложен Лондоном, были затем выполнены[77]; оказалось возможным выделить поток, соответствующий очень малому квантовому числу. Величина кванта магнитного потока составляет половину от величины, постулированной Лондоном:

$$\varphi = \pi\hbar/e = 2{,}07 \cdot 10^{-15} \text{ Вб}, \qquad (3.194)$$

откуда следует, что величина фундаментального заряда носителей тока сверхпроводимости составляет $-2e$. Это хорошо согласуется с представлением о парах Купера, лежащим в основе теории БКШ и состоящим в том, что пара электронов с противоположно направленными спинами и волновыми векторами образует единичный элемент сверхпроводящей системы.

Свободная энергия сверхпроводящего цилиндра минимальна, когда по нему совсем не течет ток, и обладает рядом минимумов для токов, соответствующих разрешенным значениям квантованного магнитного потока. Таким образом, непре-

[76] *London F.* Superfluids, v. 1, Wiley, 1950.

[77] *Deaver B. S., Fairbank W. M.*— Phys. Rev. Lett., 7, 43 (1961); *Doll R., Näbauer M.*— Phys. Rev. Lett., 7, 51 (1961).

рывное, плавное уменьшение сверхпроводящего тока не разрешено. Уменьшение тока может происходить только макроскопическими ступеньками, что требует весьма маловероятной высокой согласованности между возможными источниками рассеяния электронов. В этом состоит одна из причин (в дополнение к представлению о «бесконечной» электропроводности») стабильности во времени, наведенных незатухающих токов.

Диамагнетизм сверхпроводников I и II рода

Выталкивание магнитного потока (эффект Мейснера), описанное нами в связи с нижним рядом схем на рис. 3.73, имеет место, как было сказано, только до некоторой предельной напряженности поля, зависящей от H_c, от формы образца и от того, относится ли данный сверхпроводник к I или II роду. Даже для сверхпроводника I рода, если мы хотим, чтобы выталкивание магнитного потока вплоть до поля H_c было полным, следует использовать образцы такой формы, для которой размагничивающий фактор практически равен нулю (например, длинный сплошной стержень). Выше этого поля проводимость опять становится нормальной. Для образца сферической формы размагничивающее поле имеет конечную величину и магнитный поток будет полностью выталкиваться, только пока внешнее приложенное поле остается меньше $(2H_c/3)$.

Предположим, что образец сверхпроводника I рода имеет форму сплошного длинного цилиндра. В присутствии внешнего поля, меньшего H_c, в этом образце установятся циркулирующие поверхностные токи, благодаря которым внутри образца $\mathbf{B}=0$. Массивный образец ведет себя как *идеальный диамагнетик* с отрицательной намагниченностью $\mathbf{M}=-(\mathbf{H}-\mathbf{B}/\mu_0)$, изменяющейся с полем, как показано кривой *I* на рис. 3.74. В поле выше критического поля H_c проводимость металла должна стать «нормальной» и величина $(\mathbf{B}/\mu_0)$ внутри твердого тела будет отличаться от H снаружи от него столь же мало, как в любом парамагнетике.

Вторая кривая на рис. 3.74 указывает на совершенно другое поведение сверхпроводника II рода в магнитном поле. В материал этого типа возможно частичное проникновение магнитного потока в полях, превышающих первое критическое поле H_{c1}. В полях, больших H_{c1}, но не превосходящих второе предельное значение, обозначенное H_{c2}, материал находится в так называемом *«промежуточном»* состоянии. «Жестким» называют такой сверхпроводник II рода, для которого поле H_{c2} очень велико. [К моменту написания этих строк стали известны такие сверхпроводники II рода, в которых H_{c2} составляет $2{,}8\cdot10^7$ А/м ($B_{c2}=35$ Тл) при абсолютном нуле.] Такие мате-

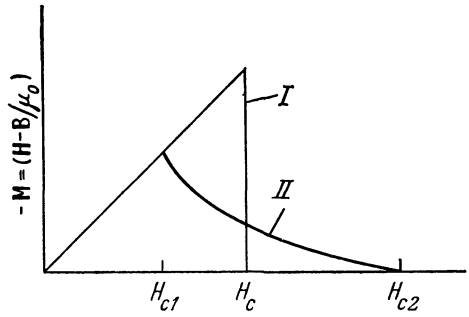


Рис. 3.74. Магнитное поведение сверхпроводников I и II рода. Полное выталкивание магнитного потока (эффект Мейснера) имеет место в *идеальном диамагнетике*, для которого зависимость намагниченности от магнитного поля соответствует кривой, изображенной здесь для сверхпроводника I рода (или «мягкого» сверхпроводника). В сверхпроводнике I рода сверхпроводящее состояние внезапно исчезает при критическом поле H_c. *Частичное* выталкивание магнитного потока, при котором все еще сохраняется сверхпроводимость на постоянном токе, имеет место в сверхпроводнике II рода для полей, превышающих H_{c1}, но меньших H_{c2}.

риалы сейчас практически используются для изготовления сверхпроводящих катушек для электромагнитов; предполагается, что их можно будет использовать для генерации и передачи электроэнергии (см. последний раздел настоящей главы), поскольку плотность тока, который может передаваться без разрушения сверхпроводящего состояния, составляет 10^{10} А/м^2.

Сверхпроводник II рода в промежуточном состоянии обладает бесконечно большой проводимостью на постоянном токе, хотя он и состоит из расположенных мозаичным образом нитевидных областей «нормального» и сверхпроводящего материалов. Существование такой смеси сверхпроводящего и несверхпроводящего материалов было постулировано Ландау (1937 г.). Можно показать, что для материалов, удовлетворяющих определенному ряду требований, такая структура обладает меньшей энергией, чем полностью нормальное или идеальное диамагнитное состояние. Грубо сформулированное условие существования промежуточного состояния заключается в том, чтобы *длина когерентности*, связанная с волновыми функциями сверхпроводящего состояния (Пиппард, 1953 г.), была меньше введенной Ландау глубины проникновения λ. Для сверхпроводника I рода волновые функции когерентны на расстояниях, гораздо больших, чем глубина проникновения, поэтому требуется меньшая энергия для создания единственного внешнего барьера между магнитным полем и сверхпроводником, чем для образования нитевидной структуры.

Микроскопическая теория сверхпроводимости

В этом разделе уже неоднократно упоминалась микроскопическая теория БКШ, которая впервые объяснила поведение электронов в сверхпроводнике. Эта теория, являющаяся одним из важнейших достижений теории сверхпроводимости, появилась примерно через 50 лет после открытия самого явления и примерно на 30 лет позже того, как Блох дал описание одноэлектронных квантовых состояний в нормальном металле. В течение всего этого времени проблема сверхпроводимости вызывала значительный теоретический интерес, и большая часть появившихся за этот период работ Гортера, Казимира, Лондона, Гинзбурга, Ландау и Пиппарда по феноменологической теории создала тот фундамент, на котором основывается описание сверхпроводимости как с макроскопической, так и с микроскопической точки зрения.

Заметив, что сверхпроводимость наблюдается преимущественно в тех твердых телах, обычная проводимость которых не слишком велика (в тех, в которых электроны часто рассеиваются фононами), Фрелих[78] в 1950 г. предложил модель, согласно которой электрон при движении по решетке непрерывно испускает и вновь поглощает виртуальные фононы. В этой модели Фрелих использовал представления, развитые им раньше и касающиеся поляронных состояний (электрон плюс поляризация) в ионных кристаллах. Фрелих предположил, что электрон в металле возмущает атомы в своей окрестности, вызывая их колебания, а возмущение решетки затем в свою очередь воздействует на электрон. Он показал, что при таком типе взаимодействия основное состояние может обладать меньшей энергией, чем система невзаимодействующих электронов, целиком заполняющих область Ферми, и что такое сверхпроводящее основное состояние должно быть отделено от «нормальных» состояний электронов проводимости электрической щелью. Распределение электронов в k-пространстве, которое дало бы это новое состояние с меньшей энергией, следовало вычислять с помощью теории возмущений (которая, увы, в этом случае не работает).

Непригодность теории возмущений для модели Фрелиха не так существенна. Гораздо более важна та его идея, что если главную роль играют электрон-фононные взаимодействия, то для изотопов данного элемента критическая температура T_c должна быть пропорциональна температуре Дебая θ_D. Это означает, что температура T_c должна изменяться, как $M^{-1/2}$, где M — масса изотопа.

---

[78] *Fröhlich H.*— Phys. Rev., **79**, 845 (1950).

Когда Фрелих высказал свое предположение, ему были неизвестны результаты двух экспериментальных исследований[79] зависимости T_c от M для изотопов ртути, которые показали, что T_c *действительно* изменяется как $M^{-1/2}$. Аналогичная зависимость T_c от массы изотопа была после этого получена для других элементов. Благодаря такому подтверждению *изотопического эффекта* модели Фрелиха и Бардина[80], связывающие сверхпроводимость с электрон-фононными взаимодействиями, были всесторонне исследованы.

Важное недостающее звено было внесено в это исследование Купером[81]. Его идея состояла в том, что электрон-фонон-электронное взаимодействие в сверхпроводящем состоянии может уменьшить (или даже свести к нулю) кулоновское отталкивание между двумя электронами по сравнению с его величиной в нормальном состоянии. Купер рассмотрел последовательность событий, которая начинается с того, что электрон деформирует решетку в своей окрестности. Это приводит к рождению фонона, который распространяется по кристаллу. Затем второй электрон включается во взаимодействие, поглощая фонон. Если обозначить начальный и конечный волновые векторы первого электрона через $\mathbf{k}_a$ и $\mathbf{k}_a^*$, а соответствующие волновые векторы для второго электрона — через $\mathbf{k}_b$ и $\mathbf{k}_b^*$, то

$$\mathbf{k}_a - \mathbf{k}_a^* = \mathbf{k}_b^* - \mathbf{k}_b = \mathbf{k}_{\text{фонон}}. \tag{3.195}$$

На первый взгляд мы могли бы представить себе аналогичный закон сохранения и для *энергии*, приобретаемой одним электроном за счет ее убыли у второго электрона. Однако поскольку процесс обмена фононом виртуальный, энергия сохраняется для *кристалла в целом*, а энергия пары электронов не обязана оставаться той же самой, что и до обмена. Таким образом, Купер показал, что для пары электронов, находящихся над самой поверхностью Ферми, фононный обмен, если он приводит к сколь угодно слабому притягивающему взаимодействию, будет создавать связанное состояние.

Бардин, Купер и Шрифер[71], развивая эту возможность, показали, что преимущественными условиями для конденсации в связанные пары при нулевой плотности тока обладают электроны с противоположно направленными спинами и противоположными величинами волновых векторов. В спаривании участвуют электроны, находящиеся в состояниях из интервала энергий шириной $k_0\theta_D$ около энергии Ферми. В лекциях, прочи-

---

[79] *Maxwell E.*— Phys. Rev., 78, 477 (1950); *Reynolds C. A. et al.*— Phys. Rev., 78, 167 (1950).

[80] *Bardeen J.*— Phys. Rev., 78, 167 (1950).

[81] *Cooper L. N.*— Phys. Rev., 104, 1189 (1956).

танных Шриффером, Купером и Бардином в Стокгольме в связи с присуждением им Нобелевской премии по физике 1972 г. (перепечатанных в июльском выпуске 1973 г. журнала Physics Today), содержится блестящее изложение того, как в результате работы трех авторов была создана теория БКШ.

Таким образом, предполагается, что при абсолютном нуле и нулевой плотности тока основное состояние сверхпроводника есть состояние с высокой степенью корреляции, в котором электронные состояния в k-пространстве, расположенные вблизи поверхности Ферми, в наибольшей возможной степени заняты парами с противоположными значениями спина и волнового вектора. Основное состояние всех куперовских пар может быть представлено одной когерентной волновой функцией. Куперовская пара ($\mathbf{k}_1\uparrow$, $-\mathbf{k}_1\downarrow$) путем обмена соответствующим виртуальным фононом может перейти в какую-либо другую незанятую пару состояний ($\mathbf{k}_2\uparrow$, $-\mathbf{k}_2\downarrow$). Квантование магнитного потока служит убедительным экспериментальным свидетельством в пользу существования таких пар.

Между энергией куперовской пары и энергией двух отдельных неспаренных электронов имеется значительная энергетическая щель. Эта энергия во много раз больше энергии, необходимой только для того, чтобы разрушить одну пару, так как при заполнении одного электронного состояния неспаренным электроном исчезает возможность фононного обмена для какой-либо из оставшихся куперовских пар, который мог бы перевести один из электронов пары в это заполненное состояние. Таким образом, энергетическая щель, столь очевидно проявляющаяся в инфракрасных свойствах и в поведении электронной теплоемкости, в приближении БКШ возникает как естественное свойство.

Как уже было сказано, каждая куперовская пара симметрична относительно $\mathbf{k}=0$ в сверхпроводнике *с нулевой плотностью тока.* При наличии отличного от нуля незатухающего тока вся сфера Ферми смещается как целое (вместе с окружающей ее энергетической щелью) и каждая куперовская пара имеет отличный от нуля импульс. В отличие от обычного электрического тока в «нормальном» проводнике этот сверхпроводящий ток не может релаксировать в результате рассеяния на фононах или локализованных дефектах. Энергетическая щель стабилизирует куперовские пары, препятствуя любому малому изменению полного импульса. Таким образом, бесконечно большая проводимость сверхпроводника на постоянном токе также возникает как естественное (и существенное) следствие теории.

Согласно теории БКШ, энергетическая щель уменьшается до нуля по мере того, как температура приближается к T_c,

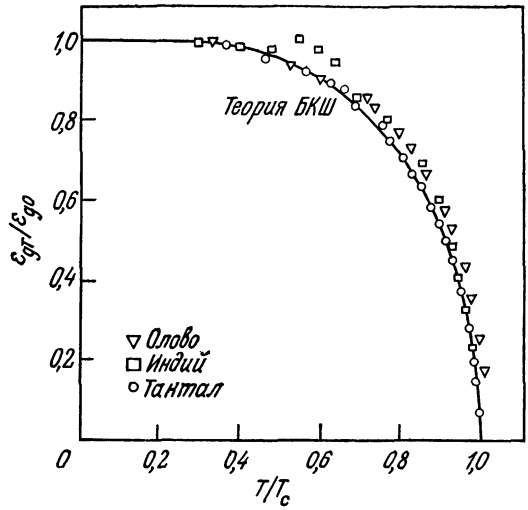


Рис. 3.75. Температурная зависимость ширины энергетической щели для сверхпроводимости. По оси ординат отложена ширина щели, отнесенная к ее ширине при $T=0$, а по оси абсцисс — температура, отнесенная к температуре перехода. Кривая соответствует теории БКШ. Данные для олова получены из измерений акустического затухания [*Morse R. W., Bohm H. V.*— Phys. Rev., **108**, 1094 (1957)]; данные для индия — из экспериментов по электронному туннелированию [*Giaever I., Megerle K.*— Phys. Rev., **122**, 1101 (1961)]; данные для тантала — из экспериментов по туннелированию [*Townsend P., Sutton J.*— Phys. Rev., **128**, 591 (1962)].

поскольку доля электронов, находящихся в спаренных состояниях, падает при повышении температуры. Полная величина зазора между состояниями, расположенными над и под щелью, условно характеризуются величиной $2\varepsilon_{g0}$ при нулевой температуре и $2\varepsilon_{gT}$ при конечной температуре. Связь между ε_{gT} и T в теории БКШ описывается интегральным уравнением, которое, как и большинство формул для сверхпроводящих свойств, удобно записывать для безразмерных величин $(\varepsilon_{gT}/\varepsilon_{g0})$ и (T/T_c). Сплошная кривая на рис. 3.75 показывает взаимную зависимость этих двух безразмерных переменных; здесь же для сравнения приведены экспериментальные данные для ширины щели в трех элементах. Данные для In и Ta на этом рисунке были получены с помощью уже упоминавшегося туннельного метода Живера[62], в котором измеряются вольтамперные характеристики для сандвич-структуры металл — диэлектрик — сверхпроводник. Диэлектрический слой в этом сандвиче настолько тонкий (<100 Å), что электроны могут легко туннелировать из состояний в металле с одной стороны в состояния сверхпро-

водника с другой стороны. Таким способом электроны с энергией Ферми в «нормальном» металле с одной стороны сандвича могут сканировать энергетическую зависимость плотности состояний в сверхпроводнике. При изменении приложенного напряжения смещения в пределах нескольких милливольт наблюдаются разрывы тока, соответствующие верхнему и нижнему краям энергетической щели сверхпроводимости. Такие измерения вместе с другими экспериментальными данными (поглощение в далекой инфракрасной области, затухание ультразвука и т. д.) дают убедительное подтверждение того, что основное состояние сверхпроводника отделено от состояний «нормальной» проводимости, как это предсказывается теорией Бардина, Купера и Шриффера.

На рис. 3.76 в сильно преувеличенном виде представлено различие спектра и заполнения состояний в нормальном металле и в сверхпроводнике. На рис. 3.76, *а* показана плотность состояний при $T=0$ в отсутствие сверхпроводимости (такое состояние можно получить, приложив магнитное поле, превышающее H_{c0}). Эта плотность состояний описывается формулами (3.44) и (3.160). Функция заполнения Ферми—Дирака (3.45) равна единице до энергии ε_{F0} и нулю для любой более высокой энергии.

Основное состояние сверхпроводника в теории БКШ при нулевой температуре изображено на рис. 3.76, *б*. Плотность состояний равна нулю для энергий, расположенных внутри интервала $\pm\varepsilon_{g0}$ с каждой стороны от энергии Ферми, по обе из сторон щели скапливаются смещенные состояния. При $T=0$ все электроны входят в куперовские пары и ни один из них не возбужден в более высокие состояния.

На рис. 3.76, *в* представлена плотность состояний при конечной температуре, меньшей T_c. Энергетическая щель ε_{gT} теперь меньше, чем ε_{g0}. Более того, не все электроны остаются в сверхпроводящих парах, часть из них находится в состояниях выше $(\varepsilon_{F0}+\varepsilon_{gT})$, оставляя при этом незаполненными состояния («дырки для пар») в области ниже $(\varepsilon_{F0}—\varepsilon_{gT})$. В конце концов щель уменьшается до нуля и доля спаренных электронов обращается в нуль, когда температура достигает T_c.

Такая картина соотношения между заполненными и незаполненными состояниями в сверхпроводнике полезна тем, что позволяет наглядно себе представить, как с помощью экспериментов, подобных экспериментам Живера, можно сканировать зависимость плотности состояний от энергии, измеряя вольтамперные характеристики туннельного перехода, содержащего сверхпроводник с обеих или с одной стороны от перехода. Изображенная картина не вполне точна в том смысле, что в теории БКШ основное состояние состоит не из невзаимодействую-

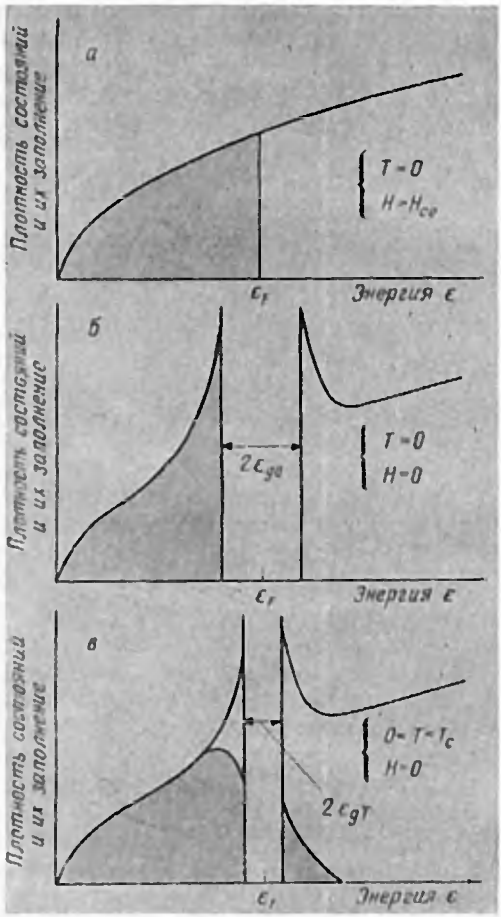


Рис. 3.76. Наглядная иллюстрация различия в спектре состояний нормального металла и сверхпроводника при очень низких температурах. Ширина энергетической щели для сверхпроводимости в частях *б* и *в* рисунка сильно преувеличена.

щих одночастичных состояний, подчиняющихся статистике Ферми—Дирака, а из взаимодействующих электронных пар, что приводит к отклонению от статистики Ферми.

Применение сверхпроводимости

Практические приложения сверхпроводимости подробно обсуждаются в ряде книг, к которым относятся монографии Фишлока, Ньюхауза, Солимэра и Уильямса, приведенные в библио-

графии в конце главы. В данном разделе мы хотим обрисовать диапазон ее приложений, как реализующихся, так и потенциально возможных.

Из всех сверхпроводящих устройств наибольшее распространение получил на сегодняшний день *сверхпроводящий магнит*, в котором большой постоянный ток в спиральной обмотке из сверхпроводящей проволоки используется для создания сильного магнитного поля для лабораторных и промышленных целей. Неудобство, связанное с необходимостью иметь жидкий хладагент, в который должна быть погружена катушка электромагнита, часто компенсируется теми преимуществами, которые дает возможность получения с ее помощью магнитной индукции, достигающей 15 Тл при пренебрежимо малых омических потерях (I^2R). Создание таких сверхпроводящих магнитов стало возможно только после того, как были найдены очень «жесткие» сверхпроводящие материалы II рода, для которых поле H_{c2} при рабочей температуре имеет большую величину[82]. Некоторые из этих материалов приведены в верхней части правого столбца табл. 3.4. Они характеризуются большими значениями H_{c2} при $T=0$ (для Va_3Ga при абсолютном нуле, согласно сообщениям, B_{c2} достигает 35 Тл) и высокими значениями T_c. Получение многих из этих материалов сопряжено со значительными трудностями прежде всего из-за их исключительной хрупкости.

От хрупкости, свойственной поликристаллическим материалам, можно в какой-то мере избавиться устранением примесей и отжигом (т. е. термообработкой). К несчастью, эти попытки облегчить задачу конструктора магнита сопряжены с уменьшением способности к переносу большого тока. Так происходит потому, что сверхпроводник II рода в промежуточном состоянии (выше H_{c1}, но ниже H_{c2}) представляет собой мозаичную структуру из нитей сверхпроводящего и «нормального» материала. Прохождение тока, создающего поле H, превышающее H_{c1}, приводит к миграции под действием силы Лоренца вихрей магнитного потока, связанных с «нормальными» включениями. Поэтому практически важно, чтобы в объеме материала II рода были распределены нерегулярности решетки, которые должны закреплять * эти вихри и таким образом предотвращать потери энергии, связанные с их движением. Хрупкость проволоки — это та цена, которую приходится платить за то, чтобы сверхпроводящий магнит был способен создавать очень большие поля.

---

[82] *Kunzer J. E.*— Rev. Mod. Phys., 33, 501 (1961).

* Английское слово pin буквально означает «пришпилить». Отсюда происходит название явления — «пиннинг». См.: *Абрикосов А. А.* Основы теории металлов.— М.: Наука, 1987.— *Прим. перев.*

Протекание переменного тока по сверхпроводнику II рода при поле, превосходящем нижний предел H_{c1}, неизбежно сопровождается гистерезисными потерями, поскольку небольшое периодическое движение магнитных вихрей, связанных с нитями «нормального» материала, невозможно исключить полностью. Таким образом, при больших полях материал обладает конечным сопротивлением на переменном токе. Чем больше частота, тем больше ожидаемые омические потери. Из-за этого при использовании жестких сверхпроводников на частотах, составляющих несколько сотен герц, токи должны иметь такую величину, чтобы поле все время оставалось меньше H_{c1}. С учетом указанного ограничения частоты была выполнена работа (см. книги Ньюхауза и Уильямса) по использованию жестких сверхпроводников для обмоток электрических генераторов переменного тока и электродвигателей. Здесь затраты энергии, связанные с охлаждением, должны компенсироваться тем выигрышем, который дает уменьшение массы (предположительно в 10 раз) при данной рабочей мощности. Рассматривалась также возможность использования сверхпроводящих материалов для подземных кабелей с применением жестких сверхпроводников II рода для передачи электроэнергии на постоянном токе и I рода (вероятно, в виде тонкой пленки) для системы электропередачи на переменном токе.

С тех пор как впервые стало известно о «жестких» сверхпроводниках с большими значениями T_c и H_{c2} и высокой допустимой силой тока[82], исследователей заинтересовало, насколько высокой можно сделать температуру T_c. У двух материалов, приведенных вверху правого столбца табл. 3.4, T_c близка к точке кипения жидкого водорода. Однако отсюда не следует, что именно водород, а не гелий может быть использован в качестве хладагента во всех будущих приложениях сверхпроводников. (Щель сверхпроводимости при $T=T_c$ обращается в нуль, как и способность выталкивать магнитный поток и переносить без потерь сверхпроводящий ток. Серьезно говорить о сверхпроводящих приборах, охлаждаемых жидким водородом, станет возможно, когда температура T_c будет лежать существенно выше 23 К. Маттиас[83] предвидит возможность получения таких сплавов, в которых T_c может достигать 30 К, однако он сомневается, что эту температуру можно поднять выше 25—30 К *. Некоторые авторы[84] возлагают большие надежды на то, что

---

[83] *Matthias B. T.*— Physics Today, 24, 8, 23 (1971).

* О высокотемпературной сверхпроводимости см. предисловие редакторов перевода.— *Прим. ред.*

[84] Например, *Little W. A.*— Phys. Rev., 134, A1416 (1964).

органические комплексы с переносом заряда, обладающие высокой электропроводностью в одном направлении, могут обладать одномерной сверхпроводимостью с высокой температурой T_c. Однако большинство усилий в этой области не увенчалось успехом из-за фазового перехода материала в диэлектрик при охлаждении. Важно различать те органические соли с переносом заряда, которые представляют собой металлы или полуметаллы, и те, которые из-за изменения внутримолекулярных расстояний обладают зонной структурой, где отсутствуют частично заполненные зоны. Пайерлс в своей классической книге (см. библиографию) предупреждал, что энергия частично заполненной 1D-зоны может понизиться в результате возникновения последовательности межатомных расстояний, образующих (непроводящую) сверхрешетку. Поэтому, хотя перхлорат с органическим катионом (HMTSF) и *является* сверхпроводником [85], значение T_c для этой соли лишь ненамного превышает 1 К. Научный интерес к 1D-проводимости в свое время несомненно приведет к получению несколько бóльших значений T_c.

Ряд приложений сверхпроводимости основан на использовании сверхпроводников I рода; при этом главный интерес представляют не столько высокие температура перехода или магнитные поля, сколько возможность применения высоких частот. Примером такого прибора является сверхпроводящий линейный ускоритель, в котором высокочастотные объемные резонаторы окружены тонким слоем свинца или ниобия, охлаждаемым сверхтекучим гелием. Сверхпроводящий линейный ускоритель в Станфордском университете [86] работает на частоте 950 МГц, и добротность Q внутри резонатора достигает 10^{10}. Поскольку в сверхпроводник I рода магнитное поле, меньшее, чем объемное критическое поле, практически не проникает, при движении по такому сверхпроводнику магнитного потока потерь не возникает. Даже магнитный поток, проникающий на глубину Ландау λ, определяемую формулой (3.192), входит в проводник и выходит из него за счет обратимого механизма, не сопровождающегося потерями. Это свойство лежит в основе создания *сверхпроводящих опор* без трения, которые способны выдерживать нагрузку 10^4 Н/м^2 и более. Такие опоры предполагается использовать для высокоскоростных транспортных средств. Отсутствие трения между сверхпроводником I рода и магнитным полем привело также к созданию *сверхпроводящего гироскопа,* в котором вращающаяся сфера из сверхпроводника I рода «подвешена» в магнитном поле.

---

[85] *Bechgaard K. et al.*— Phys. Rev. Lett., **46**, 852 (1981).
[86] *Fairbank W. M., Schwettman H. A.*— Cryog. Eng. News, **2**, 8, 46 (1967).

На рис. 3.67 показано открытое Оннесом резкое изменение сопротивления с температурой при переходе ртути в сверхпроводящее состояние. Исследования, проводимые с 1911 г., многократно подтвердили резкий характер этого перехода для различных материалов. Сверхпроводящий болометр [87] представляет собой тонкую полоску материала, поддерживаемую с помощью соответствующего термостатирующего устройства при температуре T_c. Под действием небольшого количества поглощенного теплового излучения сопротивление элемента по крайней мере частично возвращается к «нормальному» значению. Болометр такого типа может успешно конкурировать с другими тепловыми детекторами в отношении его чувствительности к излучению и постоянной времени, однако он требует чрезвычайно точного контроля температуры, что ограничивает возможности применения этого прибора.

Сильное различие между сопротивлением в «нормальном» и в сверхпроводящем состоянии было положено Баком [88] в основу двухпозиционного переключательного элемента, называемого *криотроном*. В нем имеются две обмотки из сверхпроводящей проволоки (или две тонкие пленки), ток в одной из которых определяет наличие или отсутствие сверхпроводимости в другой. Поэтому такое устройство может служить двоичным логическим элементом или элементом памяти в цифровом компьютере. Были также предприняты попытки использовать переходные токи для создания магнитным путем «островков» нормального материала, расположенных в определенных местах в целом сверхпроводящей тонкой пленки. Такие островки должны быть окружены циркулирующим сверхпроводящим током, который затем может быть детектирован и использован в качестве основы памяти компьютера. Важность оптимизации технологии производства компьютеров в наши дни не нуждается в обосновании.

Ряд применений сверхпроводимости опирается непосредственно на квантовую природу этого состояния вещества. Так, квантование магнитного потока в сверхпроводящем кольце, описываемое формулой (3.194), было использовано [89] в чувствительном магнитометре, который мог также служить и абсолютным амперметром переменного тока.

Туннельные переходы на основе сандвич-структур типа металл—диэлектрик—сверхпроводник (МДС) или сверхпроводник—диэлектрик—сверхпроводник (СДС) уже упоминались

[87] *Andrews D. H., Milton R. M., de Sorbo W. J.*— J. Opt. Soc. Am., 36, 518 (1946). См. также: *Nartin D. H., Bloor D.*— Cryogenics, 1, 1 (1961).

[88] *Buck D. A.*— Proc. IRE, 44, 481 (1956).

[89] *Deaver B. S., Goree W. S.*— Rev. Sci. Instr., 38, 311 (1967).

выше в связи с экспериментами Живера[62]. Эксперименты, использующие такие сандвич-переходы, позволяют получить информацию о плотности состояний в сверхпроводнике или о других свойствах твердых тел, из которых состоит сандвич. Более того, на основе туннельных СДС-переходов были созданы усилители, генераторы и детекторы высокочастотного электромагнитного излучения, а также фононов высокой частоты. В книге Солимера[64] обсуждается применение туннельных переходов, в том числе переходы Джозефсона.

В переходах Джозефсона[63], которые ранее были названы СДС-переходами, изолирующий слой настолько тонок (~ 10 Å), что сверхпроводящие куперовские пары могут переходить из одного металла в другой. Первое поразительное свойство такой структуры состоит в том, что конечный сверхпроводящий ток может протекать по цепи, содержащей переход Джозефсона, при нулевом напряжении на контакте! Этот сверхпроводящий ток очень чувствителен к магнитному полю и обращается в нуль всякий раз, когда полный магнитный поток в диэлектрическом зазоре равен целому кратному $(\pi\hbar/e)$. Именно поэтому переход Джозефсона может быть использован в качестве основного элемента чувствительного магнитометра или гальванометра. Когда к переходу Джозефсона приложено малое напряжение V_0, ток через него осциллирует с угловой частотой

$$\omega_0 = (eV_0/\hbar). \tag{3.196}$$

Напряжение в 1 мкВ приводит к генерации частоты $(\omega_0/2\pi) = 484$ МГц. Путем измерения частоты, соответствующей известному напряжению[90], величина отношения $(\hbar/e)$ была получена с очень большой точностью. Эффект Джозефсона было предложено использовать при сравнении международных эталонов напряжения.

Задачи

3.1. Все щелочные металлы от лития до цезия кристаллизуются в о. ц. к.-структуре с расстоянием между ближайшими соседями, равным 3,03; 3,71; 4,62; 4,87; 5,24 Å соответственно. Радиусы одновалентных положительных ионов этих металлов соответственно равны: 0,68; 0,98; 1,33; 1,48; 1,67 Å. Покажите, что доля объема решетки, не занятого атомными остатками, монотонно уменьшается от 94 % для лития до 82 % для цезия. Рассчитайте

[90] *Parker W. H., Taylor B. N., Langenberg D. N.*— Phys. Rev. Lett., **18**, 287 (1967). Численные значения, содержащиеся в таблице в конце книги, взяты из работы: *Taylor B. N., Parker W. H., Langenberg D. N.*— Rev. Mod. Phys., **41**, 375 (1969); эти значения получены с использованием работы тех же авторов по измерению $\hbar/e$ с помощью эффекта Джозефсона.

число свободных электронов в 1 м³ и среднюю плотность заряда электронов в металлическом натрии. Решив уравнение Пуассона, покажите, что в центре расположенной внутри кристалла натрия сферической области радиусом 10 Å, для которой концентрация электронов на 15 % меньше, чем во всем остальном кристалле, значение потенциала увеличено примерно на 12 В.

3.2. Покажите, что электрон радиусом R пройдет расстояние $(\pi R^2 N)^{-1}$ прежде, чем его траектория пересечет одну из точек, распределенных в пространстве с плотностью N в единице объема, если R мало по сравнению с $N^{-1/3}$. Объясните смысл этого результата в связи со справедливостью формулы (3.7).

3.3. Выведите соотношение между электропроводностью, температурой и средней длиной свободного пробега для меди на основе модели Друде. Используйте кривую для электропроводности (рис. 3.1), чтобы получить соответствующую кривую для зависимости λ от температуры. Известно, что расстояние между ближайшими соседями для меди составляет 2,55 Å. Следует предположить также, что электронный газ содержит один электрон на атом.

3.4. Представьте электропроводность в формуле (3.33) в виде интеграла по скорости электрона и в виде интеграла по энергии электрона для больцмановского распределения скоростей и энергий. Покажите, что оба эти выражения приводят к формуле (3.34) (интегрируйте по частям). Получите выражение для среднего времени свободного пробега τ_m и для подвижности электрона μ [м²/(В·с)] в модели Лоренца, если длина свободного пробега составляет 100 Å, и постройте кривые, показывающие, как эти величины меняются с температурой.

3.5. Сферическое тело массой M и электрическим зарядом Q движется в вязкой среде с постоянной скоростью (AQE_x) под действием электрического поля E_x. Затем в направлении z прикладывается магнитное поле с величиной индукции B_z. Составьте уравнение движения в плоскости xy и найдите радиус кривизны траектории для случая, когда магнитное поле мало. (*Указание*: компонента скорости в x-направлении фактически не изменяется.)

3.6. Представьте свободный электронный газ, в котором две трети электронов движутся со скоростью s, а одна треть — со скоростью $2s$. Каждая из этих групп может быть в свою очередь разделена на шесть равных частей, движущихся в положительном и отрицательном направлениях осей x, y и z. Рассмотрите электропроводность, эффект Холла и магнетосопротивление такого свободного электронного газа, если ток идет в направлении x, а магнитное поле приложено в направлении z. Если нужно инжектировать еще один электрон, для которого поле Холла и магнитное поле компенсируют друг друга, какой скоростью должен обладать этот электрон в x-направлении?

3.7. Покажите, что спектр разрешенных состояний для электрона, заключенного в кубический ящик с ребром L, имеет такой же вид, как и в случае циклических граничных условий. (Эта задача предназначена для читателей, которым уже приходилось применять волновую механику.)

3.8. Идеальный газ Максвелла — Больцмана подчиняется закону $pV = RT = N_A k_0 T$. Найдите соответствующее соотношение между давлением, удельным объемом и температурой для газа Ферми — Дирака. Вам нужно показать, что давление электронного газа равно двум третям его кинетической энергии, приходящейся на единицу объема.

3.9. По постоянным решетки и валентностям соответствующих элементов установлено, что концентрация свободных электронов в натрии, меди и алюминии составляет соответственно 2,5; 8,5 и $18 \cdot 10^{28}$ м$^{-3}$. Докажите, что зна-

чение энергии Ферми для натрия действительно 3,1 эВ, как на рис. 3.9, и покажите, что электроны, имеющие энергию Ферми, движутся в этих трех металлах со скоростями 1,04; 1,57 и $2{,}01 \cdot 10^6$ м/с соответственно. Насколько увеличивается масса высокоэнергичного электрона в алюминии за счет релятивистского эффекта?

3.10. Используя формулы (3.46), (3.49), (3.50) и (3.52), покажите, что небольшое понижение уровня Ферми в металле при нагревании его от абсолютного нуля описывается формулой (3.53). Чему равно это понижение для металлического натрия при 300 К?

3.11. Вычислите электронную теплоемкость в расчете на 1 м^3 при 750 К для металлов натрия, меди и алюминия (необходимые численные значения содержатся в задаче 3.9). Какую долю она составляет для каждого из металлов от его полной *молярной* теплоемкости, если известно, что температура 750 К во всех случаях гораздо выше температуры Дебая?

3.12. Покажите, что выражение (3.62) согласуется с формулой для удельной теплоемкости (3.59). Для определения плотности состояний при энергии Ферми воспользуйтесь формулой (3.44).

3.13. Показанное на рис. 3.12, *б* расщепление между начальными энергиями для двух составляющих плотности электронных состояний фактически возможно при чрезвычайно большой магнитной индукции. Какой должна быть магнитная индукция (в теслах) в металлическом натрии, чтобы энергетическое расщепление $2\mu_B\mu_0 H = 2\mu_B B$ было равно $k_0 T$ при комнатной температуре (300 К)? Какую величину должна иметь магнитная индукция, чтобы расщепление составляло 10 % от энергии Ферми?

3.14. Сравните выражение (3.4) для классической функции распределения с выражением (3.73) для равновесной функции распределения Ферми — Дирака. Первое из этих выражений пропорционально n (объемной плотности частиц), а второе нет. Почему? Покажите, что f_0 в выражении (3.73) есть действительно *нормированная* функция распределения.

3.15. Исследуйте свойства производной $(-\partial f/\partial \varepsilon)$ для распределения Ферми — Дирака. Покажите, что она симметрична, а интеграл от нее равен единице. Как ведет себя эта функция по сравнению с функциями, обычно используемыми в качестве аналитического представления дельта-функции Дирака?

3.16. Ответьте на вопросы, поставленные в задаче 3.3, решая их теперь для меди в рамках модели проводимости Зоммерфельда.

3.17. Будем считать, что кристалл золота состоит из соприкасающихся сфер радиусом 1,44 А, расположенных с симметрией г. ц. к.-решетки. Каждый атом дает один свободный электрон. Электропроводность чистого золота составляет $5 \cdot 10^7$ Ом$^{-1}\cdot$м$^{-1}$ при $T = 273$ К. Теперь предположим, что в кристалле золота имеется 0,1 % беспорядочно расположенных вакансий. Рассчитайте среднюю длину свободного пробега $\lambda_{\text{деф}}$, определяемую исключительно рассеянием на вакансиях. По известному значению электропроводности оцените среднюю длину свободного пробега для рассеяния на фононах при 273 К и используйте правило Маттиссена для определения электропроводности дефектного кристалла.

3.18. Рассмотрите соотношение (формула (3.88)) между x-компонентой скорости электрона внутри металла и его гораздо меньшей скоростью после термоэмиссии. Покажите, что распределение x-компонент скоростей эмиттированных электронов имеет вид

$$n(u_x)\, du_x \sim u_x \exp\left(-mu_x^2/2k_0T\right) du_x,$$

если ее среднеквадратичное значение составляет $(u_x)_{\text{ср. кв}} = (2k_0T/m)^{1/2}$. Покажите, что компоненты скоростей в y- и z-направлении соответствуют средней кинетической энергии эмиттированных электронов, составляющей $2k_0T$ на электрон.

3.19. В экспериментальной установке в качестве источника термоэлектронной эмиссии используется чистая вольфрамовая нить. Найдено, что рабочая температура 2000 К обеспечивает требуемый ток эмиссии. Какая температура нити потребуется, если нить изготовлена из покрытого торием вольфрама? (Используйте параметры, приведенные в табл. 3.3.) Пусть температура нити ограничена испусканием излучения черного тела. Покажите, что замена вольфрама торированным вольфрамом позволяет уменьшить ток накала нити на 44 %, если считать, что электрическое сопротивление прямо пропорционально абсолютной температуре. Во сколько раз должен быть снова увеличен ток накала, если торированная нить содержит примеси, из-за которых эффективная работа выхода увеличивается на 25 %, тогда как константа Ричардсона остается неизменной?

3.20. Заметьте, что выражение для потенциальной энергии (3.100) для эмиссии во внешнем поле основано на формуле (3.98), а не на более точных соотношениях (3.99), несмотря на то, что последние более пригодны для малых значений x. Получите выражение для потенциальной энергии во внешнем поле, основанное на соотношениях (3.99), и покажите, что положение и высота ее максимума для разумных значений электрического поля не отличаются заметным образом от получающихся по формуле (3.101). Убедитесь, что электрическое поле, равное 10^6 В/м, понижает работу выхода только на 0,04 эВ на электрон и максимум потенциальной энергии для такого поля расположен при $x_{\text{макс}} = 190$ Å.

3.21. Покажите, что интегрирование силы (3.103) по полному времени, в течение которого первичный электрон взаимодействует с электроном проводимости в твердом теле, приводит к формуле (3.104), описывающей изменение величины и направления импульса. Каким должен быть параметр соударения b, чтобы электрон с начальной кинетической энергией 1 эВ приобрел в результате столкновения с первичным электроном, обладающим энергией 500 эВ, энергию 10 эВ?

3.22. Натрий становится прозрачным в ультрафиолетовой области спектра при длинах волн, меньших $\lambda = 210$ нм. Насколько хорошо это согласуется с длиной волны $\lambda_p = (2\pi c/\omega_p)$, соответствующей плазменной частоте (3.108)? Для численного сопоставления рассчитайте минимальную среднюю плотность электронов в ионосфере, если известно, что сигналы радиовещания с частотой, не превышающей 3 МГц, обычно отражаются от слоя ионосферы обратно на Землю.

3.23. На рис. 3.31 показано, что четвертая зона квадратной решетки (рис. 3.30) имеет такой же размер, как и первая, и что ее можно привести к той же форме. Нарисуйте единую кривую линию, пересекающую основную часть первой зоны этой решетки и покажите, как эта линия разбивается по многочисленным фрагментам разорванной четвертой зоны. Это не бессмысленное упражнение, так как ниже в этой главе вы столкнетесь с движением электрона в k-пространстве под действием электрического и/или магнитного полей.

3.24. Постарайтесь решить головоломку по собиранию правильного шестиугольника приведенной зоны из фрагментов зон гексагональной решетки (рис. 3.33), начиная со второй по пятую.

3.25. Изобразите первые четыре зоны двумерной прямоугольной решетки с отношением межатомных расстояний в примитивной ячейке, равным 2 : 1.

3.26. Для многоатомной решетки, которая кристаллизуется в г. п. у.-структуре, зона Бриллюэна — шестиугольная призма. Определите размеры этой призмы в **k**-пространстве, если межатомное расстояние равно a. Покажите, что допустимая плотность состояний в **k**-пространстве соответствует для каждой зоны Бриллюэна двум электронным состояниям на атом.

3.27. Начертите две первые зоны Бриллюэна для двумерной решетки с отношением межатомных расстояний, равным 2 : 1, как в задаче 3.25. Впишите в нее поверхность Ферми для почти свободных электронов таким образом, чтобы в представлении приведенных зон она достигала центра второй зоны. Изобразите приведенные зоны для первой и второй зон, показав, что в этом случае степень заполнения зон электронами как раз достаточна для того, чтобы разделить пустые области первой зоны на четыре фрагмента.

3.28. Покажите, что для любого состояния совершенно свободного электронного газа групповая скорость вдвое превышает фазовую. Определите групповую и фазовую скорость в системе, в которой скорости сферически-симметричны, а соотношение между ε и **k** имеет вид

$$\varepsilon = Ak^2 - Bk^4,$$

и покажите, что групповая и фазовая скорости совпадают, если $\varepsilon = (2A^2/9B)$.

3.29. Рассмотрите одномерный кристалл, в котором зависимость от волнового вектора имеет вид

$$\varepsilon = \varepsilon_1 + (\varepsilon_2 - \varepsilon_1)\sin^2(ak_x/2).$$

Поместите в эту зону один электрон, считая, что он не рассеивается. Рассмотрите поведение эффективной массы электрона, его скорости и положения электрона в реальном пространстве под действием не зависящего от времени электрического поля, полагая, что отражение Брэгга происходит на зонных границах, как предполагает Хоустон. Если $a = 1$ Å, то в течение какого времени должно быть приложено поле 100 В/м, чтобы электрон мог совершить одно полное колебание в пространстве? Если ширина зоны составляет 1 эВ, то каков диапазон расстояний, покрываемый в этом колебании?

3.30. Рассмотрите опять зону и электрон из задачи 3.29. Опишите поведение электрона в синусоидально модулированном электрическом поле. Сначала поместите электрон у дна зоны, а затем рассмотрите действие синусоидального поля на электрон, расположенный в средней по энергии точке зоны.

3.31. Рассмотрите электрон в натрии. Электрон обладает энергией Ферми, а его начальная скорость лежит в плоскости xy. Покажите, что магнитная индукция $B_z = 1$ Тл вызовет циклотронное движение с радиусом орбиты 6 мкм. В каком соотношении находится площадь орбиты в реальном пространстве (в м^2) с площадью орбиты, описываемой в **k**-пространстве (в м$^{-2}$)?

3.32. Рассмотрите задачу об экстремальной циклотронной орбите в плоскости k_xk_y поверхности Ферми для анизотропного твердого тела. Выражение для энергии Ферми имеет вид $\varepsilon = (\alpha k_x^2 + \beta k_y^2)$. Пусть на тело действует индукция B_z. Найдите скорость, с которой изменяется **k** при движении по орбите, если α и β не равны между собой. Найдите орбиту, описываемую электроном в реальном пространстве, для этой ситуации.

3.33. Покажите, что в случае малой магнитной индукции суммирование в формуле (3.178) можно заменить интегрированием по бесконечно большому интервалу, что дает обычную плотность квантовых состояний (3.44).

Таким же образом покажите, что в пределе слабого поля выражение (3.179) переходит в (3.46).

3.34. Используя выражение для разрешенной плотности состояний в **k**-пространстве, покажите, что сечение ферми-сферы S_F для холодного вырожденного электронного газа дается выражением (3.185). Определите расстояние P [в $Тл^{-1}$ или в $м^2/(В \cdot с)$] между последовательными пиками в эффекте дГвА на рис. 3.64 и оцените соответствующее сечение S_F орбиты Ферми в $м^{-2}$. Вычислите концентрацию электронов n ($м^{-3}$), составляющих этот электронный газ (предполагая, что тензор массы электрона для калия сводится к скаляру). Покажите также, что электроны в калии обладают эффективной массой, очень близкой к массе свободного электрона, поскольку в этом металле $\varepsilon_F = 2{,}0$ эВ.

3.35. В связи с решением задачи 3.34 рассмотрите экспериментальные данные для полупроводника GaSb на рис. 3.66. Заметьте, что приведенная для него постоянная Холла равна $R_н = -4{,}8 \cdot 10^{-6}$ $м^3/Кл$, и рассчитайте по нему полную концентрацию электронов и площадь экстремального сечения поверхности Ферми в **k**-пространстве. Покажите, что такая же площадь орбиты получается, если определить ее по периоду P обратного магнитного поля в эффекте ШдГ.

3.36. В работе *File J., Mills R. G.*— Phys. Rev. Lett., **10**, 93 (1963) использовался соленоид из сверхпроводящего сплава Nb_3Zr. Он представляет собой трубку диаметром 10 см, на которую в два слоя были намотаны 984 витков проволоки диаметром 0,5 мм (длина соленоида 25 см). Предположив, что эта длина достаточно велика, чтобы считать соленоид бесконечно длинным, вычислите индуктивность соленоида. Известно, что уменьшение тока сверхпроводимости в этом соленоиде в сверхпроводящем состоянии составляет 10^{-9} от его величины в час. Описывая это затухание формулой

$$I = I_0 \exp(-Rt/L),$$

определите верхний предел сопротивления обмотки, а из него — нижний предел электропроводности в сверхпроводящем состоянии. По данным авторов, электропроводность составляла не менее 10^{23} $Ом^{-1} \cdot м^{-1}$.

3.37. Рассмотрите первое квантовое состояние сверхпроводящего кольца, представляющего собой один виток алюминиевой проволоки. Проволока имеет диаметр $d = 2$ мкм, а кольцо — радиус $R = 20$ мкм. Предполагая, что магнитный поток через кольцо определяется формулой (3.194), вычислите, какой ток должен циркулировать по кольцу, а также величину **H** или **B** в центре кольца. Теперь воспользуйтесь более точным выражением

$$\phi = (\pi\hbar/e)\,[1 + (8\pi\lambda^2/Rd)]$$

для магнитного потока через кольцо, размеры которого не бесконечно велики по сравнению с глубиной проникновения Ландау λ. Считайте, что $n_s = 1{,}8 \cdot 10^{29}$ $м^{-3}$ электронов находятся в сверхпроводящем состоянии при температуре эксперимента, что справедливо для алюминия вблизи абсолютного нуля. Насколько изменяет эта поправка вычисленные ранее значения тока в кольце и магнитного поля в его центре?

3.38. Эта задача дает некоторое представление о том, насколько трудно сконструировать сверхпроводящий магнит, способный создавать большое поле, которое не разрушает сверхпроводящего состояния. На диэлектрическую цилиндрическую болванку радиусом $R = 50$ мм намотайте катушку Гельмгольца (два витка свинцовой проволоки, разделенные зазором R). Проволока имеет диаметр $d = 1$ мм. Какой ток можно пропускать через каждый

виток, если устройство погружено в жидкий гелий ($T=4{,}2$ К), и какая величина **H** или **B** будет при этом на оси в точке посередине между витками? (Величину критического поля для свинца при рабочей температуре определите по рис. 3.70. Поле на поверхности металлической проволоки не должно превышать эту величину.)

Литература

Экспериментальные исследования электропроводимости металлов

Blatt F. J., Schroeder P. A., Foiles C. L., Greig D. Thermoelectric Power of Metals, Plenum Press, 1976. [Имеется перевод: *Блатт Ф. Дж., Шредер П. А., Фойлз К. Л., Грейг Д.* Термоэлектродвижущая сила металлов.— М.: Металлургия. 1980.]

Gerritson A. N.— In: Handbuch der Physik, v. 19, Springer, 1956, p. 137.

Hurd C. M. The Hall Effect in Metals and Alloys, Plenum Press, 1972.

McDonald D. K. C.— In: Handbuch der Physik, v. 14, Springer, 1956, p. 137.

Meaden G. T. Electrical Resistance of Metals, Plenum Press, 1965.

Теория свободных электронов

Jones H.— In: Handbuch der Physik, v. 19, Springer, 1956, p. 227.

Lorentz H. A. The Theory of Electrons, Dover, 1952. [Имеется перевод: *Лоренц Г. А.* Теория электронов.— М.: Гостехиздат, 1956.]

Mott N. F., Jones H. Theory of Properties of Metals and Alloys, Dover, 1958.

Peierls R. Quantum Theory of Solids, Oxford Univ. Press, 1965. [Имеется перевод 1-го издания: *Пайерлс Р.* Квантовая теория твердых тел.— М.: ИЛ, 1956.]

Seitz F. Modern Theory of Solids, McGraw-Hill, 1940. [Имеется перевод: *Зейтц Ф.* Современная теория твердого тела.— М.: Гостехиздат, 1949.]

Wilson A. H. Theory of Metals, Cambridge University Press, 2nd ed.; 1965. [Имеется перевод 1-го издания: *Вильсон А.* Квантовая теория металлов.— М.: ОНТИ, 1941.]

Ziman J. M. Electrons and Phonons, Oxford Univ. Press, 1960. [Имеется перевод: *Займан Дж.* Электроны и фононы.— М.: ИЛ, 1962.]

Ziman J. M. Principles of the Theory of Solids, Cambridge Univ. Press, 2nd ed., 1972. [Имеется перевод: *Займан Дж.* Принципы теории твердого тела.— М.: Мир, 1974.]

* *Френкель Я. И.* Введение в теорию металлов.— М.: Гостехиздат, 1958.

Коллективные электронные явления: плазма в твердых телах

Haken H. Quantum Field Theory of Solids, North-Holland, 1976.

Hoyaux M. F. Solid State Plasmas, Pion, 1970.

Inkson J. C. Many-Body Theory of Solids, Plenum Press, 1984.

Platzman P. M., Wolf P. A. Waves and Interactions in Solid State Plasmas, Academic Press, 1973.

Rickayzen G. Greens Functions and Condensed Matter, Academic Press, 1984.

* *Пожела Ю. К.* Плазма и токовые неустойчивости в полупроводниках. Наука, 1977.

Зонная теория

Callaway J. Energy Band Theory, Academic Press, 1964.

Callaway J. Quantum Theory of Solid State, Academic Press, 1974, 2 vols.

Fletcher G. C. Electron Band Theory of Solids, North-Holland, 1971.

Harrison W. A. Electronic Structure and the Properties of Solids, Freeman, 1980. [Имеется перевод: *Харрисон У.* Электронная структура и свойства твердых тел.— М.: Мир, 1983.]
Harrison W. A. Solid State Theory, Dover, 1980.
Slater F. C.— In: Handbuch der Physik, v. 19, Springer, 1956, p. 1.

Поверхность Ферми и динамика электронов проводимости

Абрикосов А. А. Введение в теорию нормальных металлов.— М : Наука, 1972.
* *Абрикосов А. А.* Основы теории металов.— М.: Наука, 1987.
Cracknell A. P., Wong K. C. The Fermi Surface, Oxford Univ. Press, 1973.
Faber T. E. Introduction to the Theory of Liquid Metals, Cambridge Univ. Press, 1972.
The Fermi Surface, eds. W. A. Harrison, M. B. Webb, Wiley, 1960.
Pippard A. B. Dynamics of Conduction Electrons, Gordon and Breach, 1968.

Сверхпроводимость и ее применения

Blatt J. M. Theory of Superconductivity, Academic Press, 1964.
Kuper C. G. An Introduction to the Theory of Superconductivity, Oxford Univ. Press, 1968.
Lynton E. A. Superconductivity, Halsted, 3rd ed., 1971. [Имеется перевод 2-го издания: *Линтон Э.* Сверхпроводимость.— М.: Мир, 1971.]
Applied Superconductivity, ed. V. L. Newhouse, Academic Press, 1975, 2 vols.
Superconductivity, ed. R. D. Parks, Dekker, 1969, 2 vols.
Solymar L. Superconductivity Tunneling and Applications, Halsted, 1972.
Tinkham M. Introduction to Superconductivity, Krieger, 1980.
Williams J. E. C. Superconductivity and its Applications, Pion, 1970.

4

Полупроводники

При обсуждении в гл. 3 свойств металлов в некоторых случаях мы не могли не затронуть смежного вопроса электронной проводимости диэлектриков и полупроводников. В данной главе мы рассмотрим этот тип твердых тел детально.

В металле можно изменять *подвижность* носителей заряда (изменяя концентрацию и спектр фононов путем изменения температуры или изменяя концентрацию дефектов в кристалле), однако *число* носителей заряда при этом остается постоянным. Это число свободных электронов может быть охарактеризовано определенным значением электрохимического потенциала (энергией Ферми) или каким-либо другим удобным инвариантным параметром, и самая совершенная экспериментальная техника в случае металлов может дать нам информацию только о свойствах электронов, обладающих энергией Ферми [1].

В полупроводнике дело обстоит совершенно иначе. Число носителей заряда (рáвно как и их подвижности) зависит от температуры и от наличия дефектов и примесей. В термодинамическом равновесии вероятность заполнения всех электронных состояний с любыми энергиями можно выразить с помощью единственного параметра нормировки — уровня Ферми, однако положение уровня Ферми определяется условием электростатической нейтральности полупроводникового кристалла и не обязательно совпадает с энергией какого-либо из свободных электронов. Поэтому в полупроводнике необходимо определить *зависимость* уровня Ферми от температуры и концентрации дефектов [2]. При нарушении термодинамического равновесия

---

[1] В примечании [58] гл. 3 обсуждаются некоторые экспериментальные методики, которые дают информацию об электронных состояниях металла ниже энергии Ферми, в первую очередь фотоэмиссия и оптическое отражение.

[2] Термин «дефект» (flaw) был введен Шокли и Ридом (1952 г.) в качестве общего названия любого центра, который создает *локализованное* энергетическое состояние, способное отдавать или принимать по крайней мере один электрон. Так, посторонний (примесный) атом, очевидно, ведет себя как такой дефект и в том случае, когда он занимает регулярный узел решетки, и будучи внедренным в междоузлие. Собственные дефекты всех типов, обсуждавшиеся в разд. 1.6, также обычно оказывают некоторое воздействие на спектр электронов и могут вносить один или несколько «дефектных»

электроны в полупроводнике уже не могут быть охарактеризованы единственным уровнем Ферми и (опять в отличие от металла) наличие даже небольшого числа избыточных свободных электронов в полупроводнике может свидетельствовать о значительном отклонении от теплового равновесия.

В разд. 4.1 рассматривается равновесное распределение электронов по различным группам электронных состояний для «собственных» и «примесных» полупроводников. Материал этого раздела отражает научные интересы автора и поэтому изложен более подробно по сравнению с другими вопросами, обсуждаемыми в книге [3]. (Последний подраздел может быть опущен без ущерба для связи изложения с последующими разделами). Конечно, концентрация электронов не может быть измерена непосредственно. Мы измеряем какой-то параметр или комбинацию параметров твердого тела и на основании этих результатов *делаем вывод* о присутствии в кристалле определенного числа электронов и (или) дырок. Тем не менее даже в этом случае полезно рассмотреть влияние дефектов и теплового возбуждения на распределение электронов, поскольку от числа электронов и дырок радикальным образом зависит практически любой параметр, который мы можем измерить.

В разд. 4.2 рассмотрены явления переноса для свободных электронов в пространственно-однородном полупроводнике. При обсуждении явлений переноса почти всегда предполагается, что внешнее поле, приложенное к полупроводнику для наблюдения переноса, является достаточно слабым, чтобы вызвать заметное отклонение от термодинамического равновесия.

Наблюдаемые свойства большинства полупроводников усложняются из-за сложности поверхностей постоянной энергии в k-пространстве вблизи потолка валентной зоны и у дна зоны

энергетических уровней. От числа и спектра этих уровней зависит число свободных электронов и дырок в полупроводнике в равновесии. Дефекты участвуют также в генерации и аннигиляции свободных электронов и дырок как в равновесии, так и в неравновесном состоянии.

(По установившейся в советской литературе терминологии в дальнейшем при переводе для обозначения такого «универсального» дефекта мы будем пользоваться понятием «примесного центра» или просто «центра».— *Прим. перев.*)

[3] Автор проанализировал распределение электронов по зонным состояниям и состояниям примесных центров для равновесной и неравновесной ситуаций в однородном полупроводнике в монографии «Semiconductor Statistics» (Pergamon, 1962). [Имеется перевод: *Блейкмор Дж.* Статистика электронов в полупроводниках — М.: Мир, 1965.] К этой книге читатель будет отсылаться в ряде мест для получения более глубокого объяснения. Так, в ней механизмы и динамика генерации и рекомбинации носителей заряда для неравновесных ситуаций обсуждаются гораздо более подробно, чем в разд 4.4.

проводимости. Мы не имеем здесь возможности подробно обсудить эту непрерывно развивающуюся проблему, однако в разд. 4.3 ряд характерных особенностей сложной зонной структуры показан на примере хорошо изученных полупроводников. Правильное представление о реальной форме энергетических зон в полупроводнике чрезвычайно полезно, когда мы хотим выяснить, какой из механизмов релаксации доминирует в случае отклонения заселенности состояний электронами от равновесной. Разд. 4.4 объясняет, каким образом механизмы генерации и рекомбинации связаны с динамикой полупроводника в неравновесном состоянии. Эти вопросы очень важны, поскольку большинство применений полупроводников требует отклонения от равновесия.

4.1. Равновесная статистика электронов

Детальное равновесие

Если вещество находится в состоянии термодинамического равновесия, то его термодинамическое состояние можно описать с помощью температуры T. Действительно, понятие «температуры» имеет строгий смысл только, если вещество действительно *находится* в равновесии с окружающими телами, хотя часто мы не вполне точно говорим о «температуре» объекта, испытывающего слабое внешнее воздействие.

Эти замечания справедливы не только для полупроводящей среды, но и для любой другой, однако необходимость помнить о них оказывается гораздо острее для полупроводников, чем для металлов. Основная форма поверхности Ферми в металле сохраняется независимо от температуры и от того, существует или нет тепловое равновесие. Для малых концентраций подвижных электронов и дырок в полупроводнике важно знать, существует или нет равновесное распределение, и знать величину параметра «температуры», если равновесие имеет место.

Во всяком случае, когда термодинамическое равновесие в полупроводнике *существует,* температура T определяет количество и спектры электронов, дырок, фононов и фотонов внутри твердого тела. Фотоны оказываются распределенными по энергиям в соответствии с законом излучения Планка, поскольку эти квазичастицы подчиняются статистике Бозе—Эйнштейна. Фононы также подчиняются закону распределения Бозе—Эйнштейна (2.54), который управляет плотностью фононных мод, определяемых межатомными силами релаксации в твердом теле. Распределение электронов с различными потенциальными и кинетическими энергиями определяется законом

распределения Ферми—Дирака (3.45), причем параметр нормировки ε_F действует для *всех* энергий электронов.

В любой момент времени происходит поглощение фотонов излучения черного тела, приводящее к возбуждению электронов в состояния с более высокой энергией. В то же время каждый электрон, для которого находятся незанятые состояния с более низкой энергией, может перейти в одно из них, испуская при этом фотон соответствующей энергии. В состоянии термодинамического равновесия эти противоположные процессы должны совпадать по скорости как в целом, так и в каждой области спектра. Такое *детальное равновесие* существует при превращениях энергии фононов в энергию электронов и обратно и в любом другом процессе превращения энергии, который может происходить в твердом теле. Такого детального равновесия требуют законы термодинамики.

Из принципа детального равновесия следует, что существует единственное распределение электронов по энергиям, характеризуемое единственным уровнем Ферми ε_F для материала данного состава при данной температуре T. Величина этой энергии Ферми и распределение всех имеющихся электронов по полному спектру разрешенных состояний не зависят от скоростей различных возможных процессов перехода, поскольку речь идет о *не зависящем от времени* (стационарном) распределении.

Когда нарушается термодинамическое равновесие, временно нарушается детальное равновесие. В этом случае интересно выяснить, какие физические процессы будут доминировать в восстановлении равновесия. Для одного полупроводника релаксация к равновесию может происходить со скоростью, определяемой вре́менным преобладанием генерации фотонов над их поглощением. В другом твердом теле неравновесные электроны с высокой энергией наиболее легко могут быть термализованы за счет вре́менного превышения поглощения фононов над их генерацией. Все эти случаи важны, и о них вновь пойдет речь в разд. 4.4. Однако в остальной части данного раздела мы будем считать, что полупроводник можно исследовать и в состоянии термодинамического равновесия. Основываясь на этом предположении, мы обсудим равновесные распределения электронов без учета скоростей, с которыми могут происходить различные процессы превращения энергии и обратные им процессы.

Собственные и примесные полупроводники

В разд. 3.4 было указано, что свойства собственного полупроводника наблюдаются в любом кристалле-диэлектрике, когда величина k_0T достаточно велика для теплового возбуждения некоторого количества электронов из верхней части заполненной зоны в нижнюю часть следующей по высоте (обычно пустой) зоны[4]. Тогда мы ввели наименование *валентной зоны* для нормально заполненной зоны (названной так потому, что именно те электроны, которые занимают ее энергетические состояния, участвуют в образовании химических валентных связей между атомами) и *зоны проводимости* для следующей пустой зоны (поскольку любой электрон в этой зоне действительно свободен для участия в проводимости). В данной главе мы будем обозначать концентрации электронов в зоне проводимости и дырок в валентной зоне следующим образом:

n_0 — концентрация электронов в зоне проводимости в состоянии термодинамического равновесия;

n — общая концентрация электронов в зоне проводимости (не обязательно при равновесии);

$n_e = (n - n_0)$ — избыточная концентрация электронов в зоне проводимости, обусловленная отклонением от равновесия;

p_0 — концентрация дырок в валентной зоне в термодинамическом равновесии;

p — общая концентрация дырок в валентной зоне;

$p_e = (p - p_0)$ — избыточная концентрация свободных дырок, обусловленная отклонением от равновесия.

Если полупроводник является идеально чистым и кристаллографически совершенным, концентрации n_0 и p_0 должны быть равны, поскольку электроны в зоне проводимости могут появиться только путем возбуждения состояний валентной зоны. Собственную концентрацию электронов и дырок будем обозначать n_i. Распределение по энергиям как свободных дырок, так и свободных электронов характеризуется единственным уровнем Ферми. Этот собственный уровень Ферми расположен близко к середине запрещенной зоны, как показано на рис. 4.1, *а*. Далее в этом разделе мы увидим, почему положение уровня Ферми столь разительно отличается от случая металлов, но

---

[4] Из правой части рис. 3.40 можно получить очень грубое представление о самой верхней заполненной (валентной) зоне и самой нижней пустой (зоне проводимости) для кремния как типичного полупроводника. Примечания 47 и 48 гл. 3 содержат ссылки на статьи, в которых А. Вильсон ввел представление о собственном и примесном поведении полупроводника.

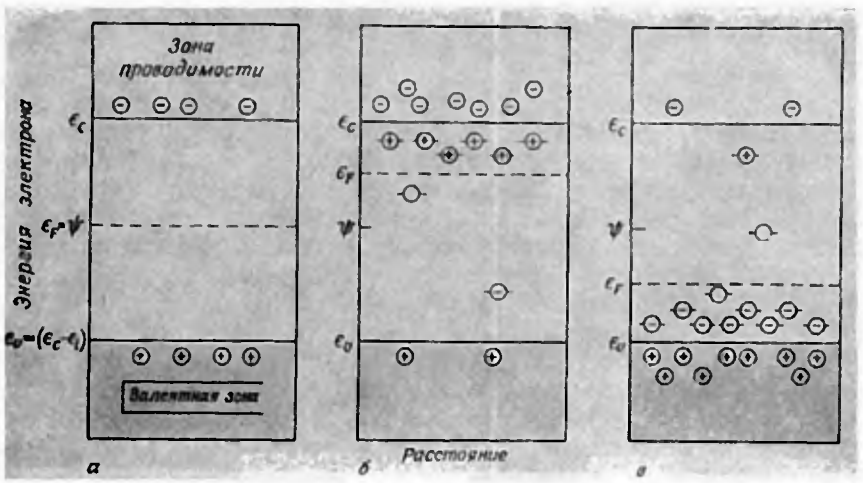


Рис. 4.1. *а* — равные концентрации свободных электронов в зоне проводимости и свободных дырок в валентной зоне для собственного полупроводника. Положение уровня Ферми, вытекающее из требования $n_0 = p_0 = n_i$, соответствует энергии ψ, которая обычно находится возле середины запрещенной зоны. *б* — тот же полупроводник, но с добавлением смеси локализованных примесных состояний, приводящих к тому, что $n_0 > n_i > p_0$. Это полупроводник *n*-типа, в котором уровень Ферми расположен выше ψ. *в* — *p*-тип проводимости возникает, когда добавленные примесные центры таковы, что уровень Ферми оказывается ниже ψ, так что $p_0 > n_i > n_0$.

сначала обсудим связь между концентрацией электронов и энергией Ферми для одной зоны.

Только в очень чистых полупроводниковых кристаллах при сравнительно высоких температурах преобладает собственное возбуждение. В частности, при низких температурах чаще всего оказывается, что собственная концентрация электронно-дырочных пар пренебрежимо мала по сравнению с концентрацией свободных носителей заряда какого-то *одного* типа, появившихся в результате теплового возбуждения уровней дефектов (примесных уровней). Таким образом, собственный полупроводник при охлаждении неизбежно должен стать примесным — *n*- или *p*-*типа*. Эти две возможности показаны на рис. 4.1, *б* и *в*.

В примесном полупроводнике *n*-*типа* свободных электронов больше, чем дырок. Название «*n*-тип» заставляет помнить о том, что процессы проводимости в таком кристалле определяются отрицательно (negative) заряженными носителями, так что эффект Холла и термоэлектрические эффекты (которые зависят от первой степени заряда подвижных носителей) *отрицательны*. Аналогичным образом эти эффекты *положительны* (positive)

в примесном полупроводнике *p-типа,* в котором преобладающими свободными носителями заряда являются положительно заряженные свободные дырки. *Донорная* примесь — это такой дефект, который может стать положительно заряженным, теряя при этом или отдавая системе один или более электронов. Полученные таким образом электроны могут стать подвижными, оказавшись в зоне проводимости (но могут быть сразу захвачены другим дефектом). Аналогично *акцепторная* примесь — это дефект, который может находиться в состояниях с одним или несколькими отрицательными зарядами, принимая, таким образом, на себя заряды электронов (и, возможно, освобождая подвижные дырки). В полупроводнике *n*-типа доноров больше, чем акцепторов, и их гораздо меньше, чем акцепторов, когда проявляется *p*-тип проводимости.

Величины n_0 и p_0 в примесном полупроводнике зависят как от температуры, так и от полного спектра локализованных примесных состояний, из которых одни могут быть ионизованы, а другие нет. Ионизация примесных центров, приводящая к увеличению концентрации свободных электронов, приводит также к уменьшению концентрации свободных дырок (и наоборот). Как показано на рис. 4.1, *б* и *в*, величина произведения $n_0 p_0$ в области, далекой от собственной проводимости, по существу точно равна n_i^2, и ниже будет показано, почему это так.

Концентрация электронов и уровень Ферми в полупроводнике с простой зоной

Для количественного обоснования предыдущих рассуждений о концентрации свободных носителей рассмотрим сначала, как конкретная энергия Ферми ε_F связана с заданной полной концентрацией электронов n_0 для простейшей зоны. Предположим, что эта зона имеет *единственный* энергетический минимум в центре зоны Бриллюэна ($\varepsilon = \varepsilon_c$) и что сферические поверхности постоянной энергии в k-пространстве распределены вокруг этого минимума по параболическому закону, как того требует скалярная, не зависящая от энергии эффективная масса m_c. Тогда плотность состояний для энергий, не слишком превышающих ε_c, составляет

$$g(\varepsilon) = \frac{1}{2\pi^2}\left(\frac{2m_c}{\hbar^2}\right)^{3/2}(\varepsilon - \varepsilon_c)^{1/2}, \qquad (4.1)$$

что можно сравнить непосредственно с формулами (3.44) или (3.160).

Предположим теперь, что некоторые состояния в такой зоне проводимости равновесно заполнены электронами, термически

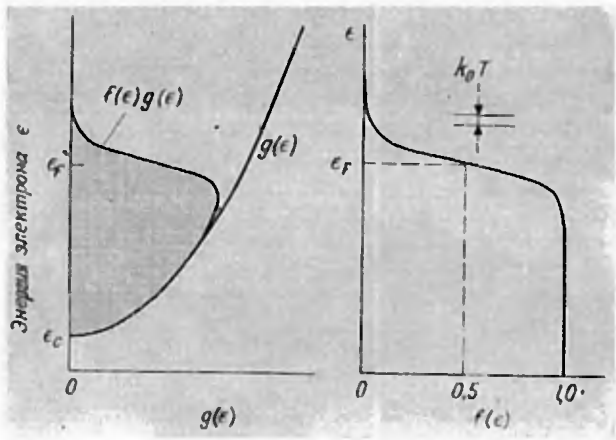


Рис. 4.2. Вырожденное распределение электронов проводимости в полупроводнике при очень большой концентрации электронов и (или) низкой температуре. Площадь затененной области в левой части рисунка представляет собой n_0 из уравнения (4.2) для ε_F. Это уравнение требует решения $\varepsilon_F > \varepsilon_c$. Следовательно, энергия Ферми и концентрация электронов связаны соотношениями (4.6).

возбужденными из нижних состояний [примесных и (или) состояний валентной зоны]. При температуре T и для полной концентрации n_0 м$^{-3}$ в равновесии существует единственное распределение по энергиям для этих электронов и единственная энергия Ферми. Вероятность того, что состояние с энергией ε заполнено, описывается функцией $f(\varepsilon)$, определяемой равенством (3.45), и искомое соотношение между концентрациями заполненных состояний с различными энергиями и полной концентрацией электронов имеет вид

$$n_0 = \int_{\varepsilon_c}^{\infty} f(\varepsilon)\, g(\varepsilon)\, d(\varepsilon) =$$

$$= \frac{1}{2\pi^2}\left(\frac{2m_c k_0 T}{\hbar^2}\right)^{3/2} \int_{(\varepsilon_c/k_0T)}^{\infty} \frac{[(\varepsilon-\varepsilon_c)/k_0T]^{1/2} d\,(\varepsilon/k_0T)}{1+\exp\,[(\varepsilon-\varepsilon_F)/k_0T]}\,. \qquad (4.2)$$

Это совпадает с выражением (3.46) модели Зоммерфельда для свободного электронного газа в металле (если не учитывать перенормировки эффективной массы).

Из рис. 4.2 и 4.3 видно, как плотность заполненных состояний меняется с энергией для двух важных предельных случаев.

1. Электронный газ в зоне проводимости *вырожден* (рис. 4.2), когда электронов проводимости много, а температура достаточно низка. В этих условиях при вычислении ин-

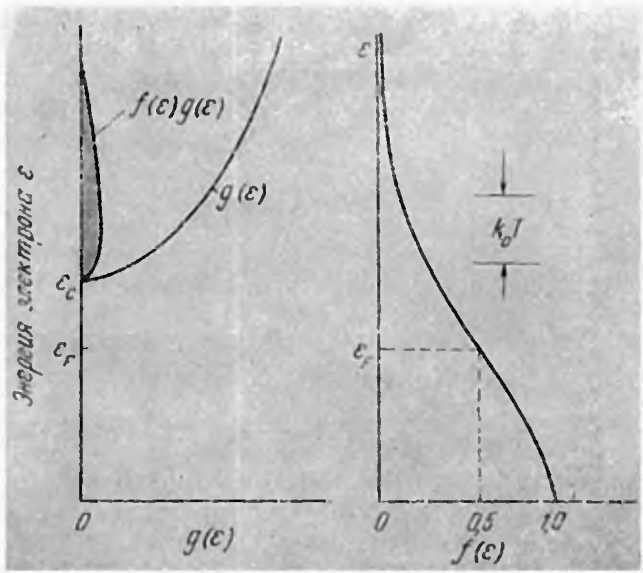


Рис. 4.3. Невырожденное, или «классическое», решение для уравнения (4.2), когда полное значение n_0 мало и (или) велика температура. Вероятность заполнения $f(\varepsilon)$ принимает для диапазона энергий, где отлична от нуля плотность состояний, предельный больцмановский вид, так что электроны, заключенные в заштрихованной области, образуют классический газ. Величина n_0 теперь связана с энергией Ферми соотношениями (4.9).

теграла (4.2) следует положить $\varepsilon_F - \varepsilon_c \gg k_0 T$ и все электропроводящие свойства будут определяться изменяющимся числом заполненных состояний ниже энергии Ферми.

2. Электронный газ является *невырожденным*, или *классическим*, когда полная концентрация электронов настолько мала (или температура так велика), что левая и правая части уравнения (4.2) уравновешивают друг друга при энергии Ферми, лежащей на несколько k_0T ниже, чем ε_c. Как видно из рис. 4.3, только малая доля зонных состояний оказывается при этом заполненной даже для самых низких состояний. Поскольку

$$f_i(\varepsilon) = \{1 + \exp[(\varepsilon - \varepsilon_F)/k_0T]\}^{-1} \to$$
$$\to \exp(\varepsilon_F/k_0T)\exp(-\varepsilon/k_0T) \quad \text{для} \quad \varepsilon \gg \varepsilon_F, \tag{4.3}$$

распределение электронов по энергиям, когда величина n_0 достаточно мала, действительно имеет классический больцмановский вид.

Выражение (4.2) может быть записано в виде

$$n_0 = \frac{1}{2\pi^2}\left(\frac{2m_c k_0 T}{\hbar^2}\right)^{3/2} F_{1/2}\left(\frac{\varepsilon_F - \varepsilon_c}{k_0 T}\right), \tag{4.4}$$

где использован один из семейства интегралов Ферми—Дирака $F_j(y_0)$, определенных выражением (3.50). При обсуждении поведения электронного газа в металлах в гл. 3 мы в основном имели дело с предельным случаем сильного вырождения, которое может иметь место и в полупроводнике при том условии, что

$$n_0 \gg \frac{1}{2\pi^2}\left(\frac{2m_c k_0 T}{\hbar^2}\right)^{3/2}. \tag{4.5}$$

Когда это неравенство действительно хорошо выполняется, имеют место соотношения

$$\begin{aligned} \varepsilon_F &\approx \varepsilon_c + (\hbar^2/2m_c)(3\pi^2 n_0)^{2/3}, \\ n_0 &\approx (1/3\pi^2)[2m_c(\varepsilon_F - \varepsilon_c)\hbar^2]^{3/2}, \end{aligned} \tag{4.6}$$

как это следует из асимптотического вида (3.52) для $F_{1/2}[(\varepsilon_F - \varepsilon_c)/k_0T)]$ в случае сильного вырождения.

Однако электронная заселенность зоны полупроводника часто оказывается слишком малой, чтобы удовлетворить условию (4.5). Чтобы рассматривать такие ситуации, более близкие к «классическим», оказывается удобным переписать уравнение (4.2) в виде

$$\begin{aligned} n_0 &= 2(m_c k_0 T/2\pi\hbar^2)^{3/2} \int_{\varepsilon_c/k_0T}^{\infty} \frac{(2/\pi^{1/2})\,[(\varepsilon - \varepsilon_c)/k_0T]^{1/2} d\,(\varepsilon/k_0T)}{1 + \exp\,[(\varepsilon - \varepsilon_F)/k_0T]} = \\ &= 2(m_c k_0 T/2\pi\hbar^2)^{3/2} \mathscr{F}_{1/2}\left(\frac{\varepsilon_F - \varepsilon_c}{k_0T}\right) = \\ &= N_c \mathscr{F}_{1/2}\left(\frac{\varepsilon_F - \varepsilon_c}{k_0T}\right). \end{aligned} \tag{4.7}$$

Здесь n_0 выражается произведением двух величин. Одна из них, $N_c = 2(m_c\, k_0 T/2\pi\hbar^2)^{3/2}$, имеет размерность (длина)$^{-3}$. Другая, безразмерная, представляет собой один из семейства интегралов $F_j(y_0)$, которые связаны с семейством $F_j(y_0)$ соотношением [5]

$$F_j(y_0) = \Gamma(j+1)\,\mathscr{F}_j(y_0). \tag{4.8}$$

Удобная особенность семейства интегралов $\mathscr{F}_j(y_0)$ состоит в том, что все они асимптотически стремятся к $\exp(y_0)$, когда y_0 принимает большие отрицательные значения. В частном случае интеграла $\mathscr{F}_{1/2}(y_0)$ этим асимптотическим выражением можно

[5] Эти две формы интегралов Ферми — Дирака и ссылки на их таблицы приведены в примечаниях 12 и 13 гл. 3. Выражая интеграл в первой строке соотношений (4.7) через $\mathscr{F}_{1/2}(\varepsilon_F - \varepsilon_c)/k_0T)$, мы воспользовались равенством $\Gamma(3/2) = (\pi^{1/2}/2)$.

пользоваться для $y_0 < -2$ с ошибкой, не превышающей 5 %. Следовательно,

$$n_0 \approx N_c \exp \frac{\varepsilon_F - \varepsilon_c}{k_0 T},$$

$$\text{если } (n_0/N_c) < 0{,}15 \qquad (4.9)$$

$$\varepsilon_F \approx \varepsilon_c - k_0 T \ln (N_c/n_0).$$

Простая взаимосвязь концентрации электронов с уровнем Ферми, выражаемая соотношениями (4.9), чрезвычайно полезна при обсуждении многих ситуаций в полупроводниках, когда нет вырождения электронного газа. Можно использовать соотношение, несколько более сложное, чем (4.9) (однако все еще содержащее в качестве главного члена экспоненту), для случаев с несколько большей концентрацией, т. е. с «полувырождением». С комбинацией поправочных членов величина $\mathscr{F}_{1/2}(y_0)$ может быть выражена в аналитическом виде для любого значения y_0, как положительного, так и отрицательного [6].

Величину N_c, определенную выражением (4.7) и использованную в соотношениях (4.9), часто называют *эффективной плотностью состояний зоны проводимости*. Происхождение этого названия ясно из соотношений (4.9), из которых видно, что в случае классического (невырожденного) распределения величина n_0 равна значению N_c, умноженному на больцмановскую вероятность заполнения для $\varepsilon = \varepsilon_c$. Таким образом, *полная* концентрация электронов (фактически распределенных по энергетическому интервалу, составляющему несколько k_0T, как показано на рис. 4.3) такая же, как если бы вместо зоны существовал набор N_c уровней с одной и той же энергией ε_c. В численном выражении получим

$$N_c = 2 (m_c k_0 T / 2\pi\hbar^2)^{3/2} = 4{,}83 \cdot 10^{21} (T m_c/m_0)^{3/2} \text{ м}^{-3}, \qquad (4.10)$$

где температура выражена в кельвинах. В задаче 4.1 исследуется простая проверка условий, при которых n_0 может быть либо гораздо больше N_c (вырождение), либо гораздо меньше N_c (невырожденный случай) в зависимости от температуры.

---

[6] Выражение $\mathscr{F}_{1/2}(y_0) \approx [0{,}27 + \exp(-y_0)]^{-1}$ хорошо согласуется с классической асимптотикой для больших отрицательных y_0 и дает ошибку, не превышающую 3 %, для $y_0 \leqslant +1$. Для бо́льших положительных значений y_0 выражение $\mathscr{F}_{1/2}(y_0) \approx (4/3\sqrt{\pi})(y_0^2 + \pi^2/6)^{3/4}$ дает столь же малую ошибку и имеет правильную асимптотику для больших положительных y_0. Обширная литература, начиная с работы Зоммерфельда (1928), посвященная попыткам выразить интегралы Ферми — Дирака в простом аналитическом виде, содержится в обзоре автора данной монографии: Solid State Electronics, 25, 1067 (1982).

Статистика свободных дырок

Для почти заполненной валентной зоны (рис. 4.4), верхняя область которой может быть охарактеризована эффективной массой m_v, число дырок (пустых состояний) в состоянии равновесия для температуры T и энергии Ферми ε_F составляет

$$p_0 = \int_{-\infty}^{\varepsilon_v} g(\varepsilon)\,[1 - f(\varepsilon)]\,d\varepsilon = \\ = \frac{1}{2\pi^2}\left(\frac{2m_v k_0 T}{\hbar^2}\right)^{3/2} \int_{-\infty}^{\varepsilon_v/k_0T} \frac{[(\varepsilon_v - \varepsilon)/k_0T]^{1/2} d\,(\varepsilon/k_0T)}{1 + \exp\,[(\varepsilon_F - \varepsilon)/k_0T}. \qquad (4.11)$$

Следовательно, можно записать p_0 как

$$p_0 = \frac{1}{2\pi^2}\left(\frac{2m_v k_0 T}{\hbar^2}\right)^{3/2} F_{1/2}\left(\frac{\varepsilon_v - \varepsilon_F}{k_0 T}\right) \qquad (4.12)$$

или

$$p_0 = N_v \mathscr{F}_{1/2}\left(\frac{\varepsilon_v - \varepsilon_F}{k_0 T}\right), \qquad (4.13)$$

где $N_v = 2(m_v k_0 T/2\pi\hbar^2)^{3/2}$ для валентной зоны — величина, аналогичная N_c для зоны проводимости. Уравнения (4.11)—(4.13) имеют следующие асимптотические выражения:

$$p_0 \approx (1/3\pi^2)\,[2m_v(\varepsilon_v - \varepsilon_F)\,\hbar^2]^{3/2} \text{ (вырождение)}, \\ p_0 \approx N_v \exp\left(\frac{\varepsilon_v - \varepsilon_F}{k_0 T}\right) \text{ (нет вырождения)} \qquad (4.14)$$

в полном соответствии с аналогичными выражениями для электронов у дна зоны проводимости. Статистика дырок полностью аналогична статистике электронов. Единственное различие состоит в том, что *уменьшению* энергии электрона соответствует *увеличение* энергии дырки. Аналогичность обеих статистик возникла так удачно потому, что функция $f(\varepsilon)$ симметрична относительно энергии ε_F, соответствующей 50 %-ному заполнению. Поэтому для любого значения энергии ζ величина $f(\varepsilon_F + \zeta)$ точно совпадает с величиной $[1 - f(\varepsilon_F - \zeta)]$.

О реальности существования дырки (отсутствия электронов в почти заполненной зоне) как квазичастицы, подчиняющейся собственной статистике и обладающей своими свойствами переноса, уже говорилось в разд. 3.5, и в данной главе мы будем считать это доказанным. Тщательный квантовомеханический анализ этого вопроса был выполнен, в частности, Шокли[7]. Шо-

[7] *Shockley W.* Electrons and Holes in Semiconductors, Van Nostrand, 1950, ch. 15. [Имеется перевод: *Шокли В.* Теория электронных полупроводников.— М.: ИЛ, 1953.]

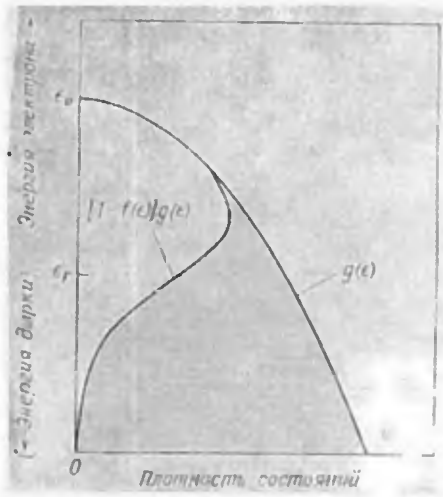


Рис. 4.4. Область расположения свободных дырок у потолка заполненной валентной зоны. Поскольку вероятность того, что состояние с энергией ε заполнено, есть $f(\varepsilon)=\{1+\exp[(\varepsilon-\varepsilon_F)/k_0T]\}^{-1}$, вероятность иметь пустое состояние (т. е. дырку) есть $[1-f(\varepsilon)]$, что равно $\{1+\exp[\varepsilon_F-\varepsilon)/k_0T]\}^{-1}$. Для конкретного случая, изображенного на рисунке, полная концентрация свободных дырок достаточно велика, чтобы создать вырождение, означающее для дырочной статистики, что уровень Ферми расположен *ниже* потолка валентной зоны.

кли пришел к выводу, что если справедлива *любая* одноэлектронная модель проводимости в твердом теле, то волновой пакет, описывающий одиночную дырку в заполненной зоне, обладает теми же свойствами, что и волновой пакет, описывающий одиночный электрон в пустой зоне.

Собственная концентрация электронно-дырочных пар

Теперь мы можем использовать формулы двух последних разделов, чтобы определить величину собственной концентрации электронно-дырочных пар n_i для полупроводника, содержащего пренебрежимо малое число примесных центров. Собственный полупроводник (рис. 4.5) имеет в двух зонах различные массы, m_c и m_v, так что N_c и N_v численно различны. Как и на рис. 4.1, мы используем символ ψ для обозначения величины уровня Ферми в области собственной проводимости, когда выполнено условие $n_0=p_0=n_i$

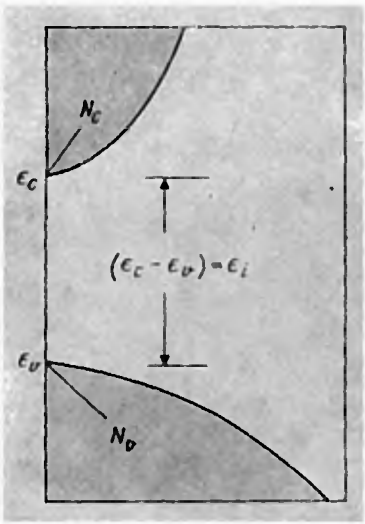


Рис. 4.5. Собственный полупроводник, для которого масса носителей в зоне проводимости m_c значительно меньше массы носителей в валентной зоне m_v. Это несовпадение усилено в степени $^3/_2$ для отношения N_v к N_c.

В предположении, что собственная щель ε_i велика по сравнению с k_0T, величина n_i будет мала по сравнению как с N_c, так и с N_v. Следовательно, в условии нейтральности

$$n_i = N_c \mathcal{F}_{1/2}\left(\frac{\psi - \varepsilon_c}{k_0 T}\right) = N_v \mathcal{F}_{1/2}\left(\frac{\varepsilon_v - \psi}{k_0 T}\right), \tag{4.15}$$

записанном с помощью равенств (4.7) и (4.13), можно воспользоваться предельными выражениями обоих интегралов Ферми для невырожденного случая. Тогда

$$n_i = N_c \exp\left(\frac{\psi - \varepsilon_c}{k_0 T}\right) = N_v \exp\left(\frac{\varepsilon_v - \psi}{k_0 T}\right). \tag{4.16}$$

Два уравнения (4.16) позволяют нам для невырожденного собственного полупроводника выразить каждую из величин n_i и ψ порознь как функцию температуры. Концентрация пар равна

$$n_i = (N_c N_v)^{1/2} \exp(-\varepsilon_i / 2k_0 T), \quad n_i \ll N_c, \; N_v, \tag{4.17}$$

где $\varepsilon_i = (\varepsilon_c - \varepsilon_v)$ — ширина энергетической щели.

Соответствующее изменение положения уровня Ферми с температурой имеет вид

$$\psi = \frac{1}{2}(\varepsilon_c + \varepsilon_v) + \frac{1}{2} k_0 T \ln(N_v/N_c), \quad n_i \ll N_c, \; N_v \tag{4.18}$$

или

$$\psi = \frac{1}{2}(\varepsilon_c + \varepsilon_v) + \frac{3}{4} k_0 T \ln(m_v/m_c).$$

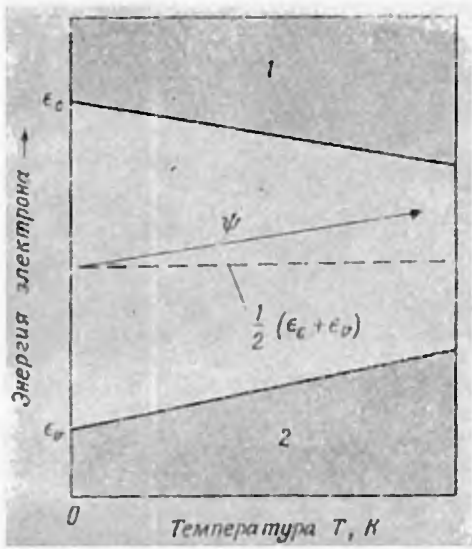


Рис. 4.6. Подъем собственного уровня Ферми ψ с ростом температуры по отношению к середине запрещенной зоны для собственного полупроводника, в котором $m_c < m_v$. Из-за несовпадения плотностей состояний в зоне проводимости (*1*) и валентной зоне (*2*) концентрации свободных электронов и свободных дырок могут оказаться равными, только если ψ лежит выше средней точки. Полупроводник, представленный на этом рисунке, обладает запрещенной зоной, уменьшающейся при нагревании, что смещает ψ еще ближе к зоне с малой массой. В предельном случае возможна ситуация, когда собственный полупроводник в зоне с легкой массой станет *частично вырожденным* при высоких температурах. [Так происходит в антимониде индия выше 400 К, согласно анализу *Austin, McClymont.*— Physica, **20**, 1077 (1954) См. также на рис. 4.9 результаты измерений собственной концентрации носителей заряда в InSb.] *1* — зона проводимости; *2* — валентная зона.

Таким образом, если только m_c и m_v не оказываются одинаковыми, собственный уровень Ферми ψ смещен от середины запрещенной зоны при любой температуре, отличной от абсолютного нуля (когда величина n_i бесконечно мала). Смещение уровня Ферми собственного полупроводника для случая несовпадения масс, соответствующего рис. 4.5, до некоторой степени проиллюстрировано рис. 4.6.

Экспоненциальный множитель в выражении (4.17) в основном (но не полностью) определяет всю температурную зависимость n_i. График зависимости n_i от $(1/T)$, построенный в полулогарифмическом масштабе, как на рис. 4.7, *а*, имеет средний наклон, в первом приближении удовлетворительно совпадающий с величиной $(-\varepsilon_i/2k_0)$, однако легко видеть, что график имеет небольшую кривизну. График же зависимости $\ln (n_i/T^{3/2})$

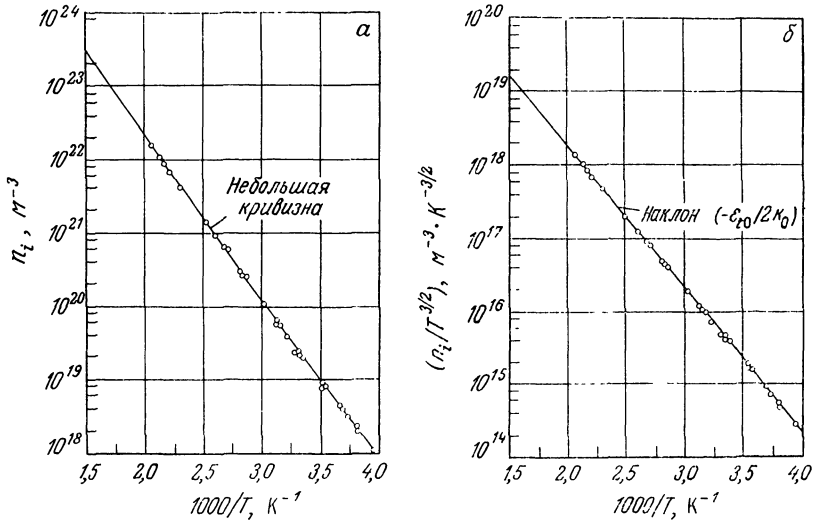


Рис. 4.7. *а* — полулогарифмическое представление собственной концентрации как функциии обратной температуры для германия. Экспериментальные точки представляют собой данные, взятые из работы: *Morin, Maita.* — Phys. Rev. **96**, 28 (1954). Обратите внимание на небольшую кривизну, обусловленную температурной зависимостью $T^{3/2}$ для множителя $(N_cN_v)^{1/2}$; *б* — те же результаты для $(n_i/T^{3/2})$ в качестве ординаты. Этот график может быть интерпретирован с помощью уравнения (4.20). Точки на графике получены с помощью экспериментальных измерений эффекта Холла.

от $1/T$ на том же рисунке представляет собой абсолютно прямую линию, поскольку неэкспоненциальная температурная зависимость множителя $(N_cN_v)^{1/2}$ при этом исключается.

Выражение (4.17) предполагает, что прямая линия на рис. 4.7, *б* имеет наклон $(-\varepsilon_i/2k_0)$. Это было бы действительно так, если бы не небольшое усложнение, состоящее в том, что собственная щель между валентной зоной и зоной проводимости германия слегка уменьшается с ростом температуры, что характерно для большинства полупроводников [8]. (Такое уменьшение ширины запрещенной зоны с нагревом схематически показано на рис. 4.6.) Изменения энергий для различных групп блоховских состояний в зависимости от температуры следует

[8] Соединения свинца PbS, PbSe, PbTe относятся к тем немногим полупроводникам, для которых собственная щель растет с нагревом (во всяком случае, до температур несколько выше 600 К). Для большинства полупроводников производная $(d\varepsilon_i/dT)$ очень мала при низких температурах и становится все более отрицательной с ростом температуры. В таблице величин $(d\varepsilon_i/dT)$ для 48 различных полупроводников, приведенной в книге Бьюба (*Bube R. H.* Photoconductivity of Solids, Wiley, 1960. [Имеется перевод: *Бьюб Р.* Фотопроводимость твердых тел.— М.: ИЛ, 1962]), только эти три соли свинца имеют положительные значения $d\varepsilon_i/dT$.

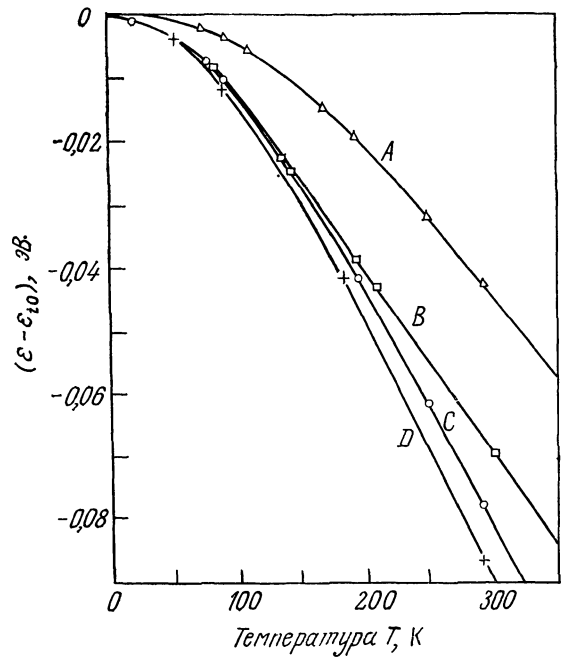


Рис. 4.8. Изменение ширины запрещенной зоны с температурой для некоторых хорошо известных полупроводников. Щель «прямая» для GaAs и InP и «непрямая» для германия и кремния (см. разд. 4.3). *A* — кремний, ε_{i0} = =1,156 эВ; *B* — InP, ε_{i0}=1,421 эВ; *C* —германий, ε_{i0}=0,741 эВ; *D* — GaAs, ε_{i0}=1,521 эВ. Данные для Ge и Si взяты из работы: *McLean T. P.* Progress in Semiconductors, v. 5, Wiley, 1960. Данные для GaAs: *Sturge M. D.*— Phys. Rev., **127**, 768 (1962). Данные для InP: *Turner W. J., Reese W. E., Pettit G. D.*— Phys. Rev., **136**, A1467, (1964).

ожидать как в результате изменения с T межатомных расстояний, так и от температурной зависимости амплитуды колебаний атомов кристаллической решетки. На рис. 4.8 показано изменение с температурой ширины запрещенной зоны для нескольких хорошо известных полупроводников, эти зависимости затрудняют интерпретацию результатов измерений собственной концентрации электронно-дырочных пар.

Если ширина запрещенной зоны может быть аппроксимирована линейной зависимостью

$$\varepsilon_i = \varepsilon_{i0} - \alpha T \tag{4.19}$$

в представляющем интерес интервале температур, то соотношение (4.17) может быть переписано в виде

$$(n_i/T^{3/2}) = \left[2\left(m_c^{1/2} m_v^{1/2} k_0/2\pi\hbar^2\right)^{3/2} \exp\left(\alpha/2k_0\right)\right] \exp\left(-\varepsilon_{i0}/2k_0 T\right), \tag{4.20}$$

где температурная зависимость правой части заключена исключительно в последнем экспоненциальном множителе. Следовательно, график $\ln(n_i/T^{3/2})$ от $(1/T)$ в полупроводнике может быть линейным, однако его наклон равен не $(-\varepsilon_i/2k_0)$, а $(-\varepsilon_{i0}/2k_0)$ и пересечение графика с осью ординат при $(1/T)\to 0$ определяется не только величинами m_c и m_v, но и параметром α.

Анализ экспериметальных данных по германию на рис. 4.7 предложен в задаче 4.2. Даже до выполнения этого анализа из сопоставления кривой для германия на рис. 4.8 с формулой (4.19) ясно, что ε_{i0} не является при *любой* температуре зазором между зоной проводимости и валентной зоной в этом полупроводнике. Трудности такого рода и даже еще бóльшие возникают при определении ширины запрещенной зоны по результатам измерений n_i в любом полупроводнике, в котором ε_i представляет собой полиномиальную функцию температуры [9]. (Это увеличивает ценность данных по оптическому поглощению в дополнение к анализу данных по измерению n_i). Переход от n_i к ε_i еще более труден, если эффективная масса плотности состояний для одной или обеих зон зависит от энергии или температуры. Среди прочих полупроводников подобной зависимостью особенно отличается такое соединение $A_{III}B_V$, как InSb, результаты измерений собственной концентрации которого представлены на рис. 4.9.

Затруднения в интерпретации измерений собственной концентрации InSb на рис. 4.9 усугубляются еще тем, что в этом полупроводнике с относительно малой шириной запрещенной

---

[9] Относительно простое нелинейное выражение для изменения ширины запрещенной зоны с температурой было предложено в работе: *Varshni Y. P.*— Phisica, **34**, 149 (1967). Варшни исходил из рассмотрения двух различных физических механизмов, приводящих к температурному изменению ширины запрещенной зоны. Он указал, что температурное сужение щели, обусловленное увеличением константы связи между электронами и колебаниями решетки, должно изменяться как T^2, если $T \ll \theta_D$, сменяясь линейной зависимостью при температурах, существенно превышающих температуру Дебая, когда возбужден весь спектр колебаний. На это сжатие щели налагается независимый вклад, обусловленный расширением решетки. Поскольку коэффициент теплового расширения для многих твердых тел при низких температурах очень мал и становится заметно больше при высоких температурах, соответствующий вклад в $\Delta\varepsilon_i$ для более высоких температур также должен быть линейным по T. Соответственно Варшни удалось удовлетворительно описать экспериментальные зависимости ширины запрещенной зоны от температуры для ряда полупроводников выражением вида

$$\varepsilon_i = \varepsilon_{i0} - AT^2/(T+B)^{-1}.$$

Однако значения B, дающие наилучшее согласие с экспериментальными зависимостями для различных полупроводников, плохо коррелируют с величиной температуры Дебая, определенной тепловыми методами.

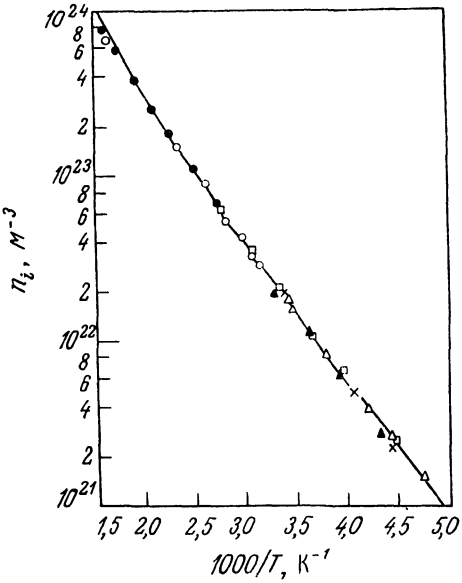


Рис. 4.9. Изменение с температурой собственной концентрации свободных носителей заряда в антимониде индия, полученное из измерений эффекта Холла и электропроводности [*Hrostowski H. J. et al.* — Phys. Rev, **100**, 1672 (1955)]. В корректном анализе должна быть учтена температурная зависимость ширины запрещенной зоны и эффективных масс, а также вырождение зоны проводимости при высоких температурах. Процедура, предложенная в задаче 4.3, не учитывает этих осложнений, и поэтому она будет давать для ε_i значения, отличающиеся от правильных величин.

зоны отношение (m_c/m_v) сильно отличается от единицы. Такое несовпадение плотностей состояний в двух зонах приводит к тому, что величина ψ при высоких температурах оказывается гораздо выше середины запрещенной зоны [как можно видеть из выражения (4.18)], и если ширина щели составляет всего несколько k_0T, то решение уравнений (4.15) может потребовать, чтобы величина ψ оказалась внутри зоны с малой массой. Намек на такое осложнение уже был сделан в подписи к рис. 4.6, но о нем умышленно ничего не было сказано при расчете (как мы теперь видим, ошибочном) зависимости ширины запрещенной зоны от температуры для InSb в задаче 4.3.

Предупреждения об осложнениях, изложенные выше, должны призывать читателя к бдительности по отношению к многочисленным неправильно рассчитанным значениям «собственной щели» в различных полупроводниках, приведенным в литературе. Как уже отмечалось выше в связи с рис. 1.54, гораздо

легче измерить величину с экспоненциальной зависимостью от $(1/T)$, чем убедиться, что измеренный наклон действительно есть та величина, которую *предполагаешь* определить.

Отклонение от модели собственного полупроводника

Для электрически нейтрального полупроводника в состоянии термодинамического равновесия любое различие между концентрациями подвижных электронов и дырок должно быть скомпенсировано зарядом $e(n_0-p_0)$, локализованным на примесных уровнях в запрещенной зоне. Статистику заполнения этих уровней мы рассмотрим в следующем разделе. Однако, ничего не зная относительно концентраций и природы примесных состояний, мы можем из соотношений (4.7), (4.13) и (4.15) заключить, что

$$n_0 p_0 = n_i^2 \left[\frac{\mathcal{F}_{1/2}\left(\frac{\varepsilon_F - \varepsilon_c}{k_0 T}\right) \mathcal{F}_{1/2}\left(\frac{\varepsilon_v - \varepsilon_F}{k_0 T}\right)}{\mathcal{F}_{1/2}\left(\frac{\psi - \varepsilon_c}{k_0 T}\right) \mathcal{F}_{1/2}\left(\frac{\varepsilon_v - \psi}{k_0 T}\right)} \right] \leqslant n_i^2. \qquad (4.21)$$

Произведение $n_0 p_0$ становится *меньше,* чем n_i^2, когда имеет место вырождение либо свободного электронного газа ($\varepsilon_F > \varepsilon_c$), либо свободного дырочного газа ($\varepsilon_F < \varepsilon_v$), поскольку интеграл Ферми $\mathcal{F}_{1/2}(y_0)$ для большого положительного аргумента растет не так быстро, как экспонента. Однако для многих примесных полупроводников, для которых уровень Ферми остается в пределах границ запрещенной зоны (т. е. когда $n_0 < N_c$ и $p_0 < N_v$), произведение $n_0 p_0$ *совпадает* с n_i^2. Следовательно, для невырожденного полупроводника (примесного или собственного) можно утверждать, что

$$n_0 = N_c \exp\left(\frac{\varepsilon_F - \varepsilon_c}{k_0 T}\right) = n_i \exp\left(\frac{\varepsilon_F - \psi}{k_0 T}\right),$$
$$p_0 = N_v \exp\left(\frac{\varepsilon_v - \varepsilon_F}{k_0 T}\right) = n_i \exp\left(\frac{\psi - \varepsilon_F}{k_0 T}\right) = \frac{n_i^2}{n_0}. \qquad (4.22)$$

Таким образом, если примесные уровни приводят к увеличению n_0, они должны одновременно *во столько же раз* (а не на ту же величину) уменьшить p_0 и наоборот. При рассмотрении свойств полупроводниковых материалов (а также приборов с p—n-переходом) очень полезно знать, что сдвиг электрохимического потенциала ε_F на величину $k_0 T$ сопровождается увеличением концентрации одного типа свободных носителей заряда в $e = 2{,}718$ раза и уменьшением концентрации носителей другого типа во столько же раз. Это не что иное, как «закон действующих масс», известный из термодинамики.

По прежнему отвлекаясь от конкретизации вида примесных состояний, предположим, что в невырожденном примесном полупроводнике n-типа $N_r = (n_0 - p_0)$. Тогда N_r обозначает *разность* концентраций тех дефектов, которые теряют электрон, и тех, которые *приобретают* электрон. Условие

$$p_0 = (n_i^2/n_0) = n_0 - N_r \tag{4.23}$$

позволяет выразить концентрацию *основных носителей* n_0 и концентрацию *неосновных носителей* p_0 через величину N_r и параметр собственного состояния n_i в следующем виде:

$$n_0 = (N_r/2)\left[(1 + 4n_i^2/N_r^2)^{1/2} + 1\right],$$
$$p_0 = (N_r/2)\left[(1 + 4n_i^2/N_r^2)^{1/2} - 1\right]. \tag{4.24}$$

На рис. 4.10 показано изменение этих двух величин от случая почти собственной проводимости [$N_r \ll n_i$, $n_0 \approx (n_i + {}^1/_2 N_r)$, $p_0 \approx n_i - {}^1/_2 N_r$)] до строго примесной [$N_r \gg n_i$, $n_0 \approx N_r$, $p_0 \approx (n_i^2/N_r)$]. Конечно, кривые электронной и дырочной концентрации должны поменяться местами для полупроводника p-типа, для которого основными носителями заряда являются дырки, а неосновными — электроны.

Переход от почти собственного состояния, соответствующего левой части графика на рис. 4.10, к полностью примесному состоянию происходит довольно резко при охлаждении собственного полупроводника. Это обусловлено тем, что величина n_i обычно быстро изменяется с температурой, в то время как величина N_r гораздо менее чувствительна к незначительному охлаждению. Эти замечания иллюстрируются кривыми на рис. 4.11, представляющими концентрации основных и неосновных носителей как функции температуры для двух образцов полупроводника InAs, который становится собственным при достижении верхнего края температурного интервала и примесным (p-типа) при охлаждении [10]. Температура перехода ниже для образца с меньшей концентрацией ионизованных примесных центров.

Для каждого образца на рис. 4.11 произведение $n_0 p_0$ есть зависящая от температуры, но не связанная со степенью чистоты образца величина n_i^2, в то время как разность $(p_0 - n_0)$ представляет собой избыток ионизованных центров акцепторного типа над ионизованными донорными центрами, причем эта величина очень мало меняется в температурном интервале,

---

[10] В задаче 4.4 предлагается построить график зависимости уровня Ферми от температуры для случаев собственного и примесного InAs (см. рис. 4.11). Отклонение отношения (p_0/n_0) от единицы сопровождается движением ε_F вниз от ψ.

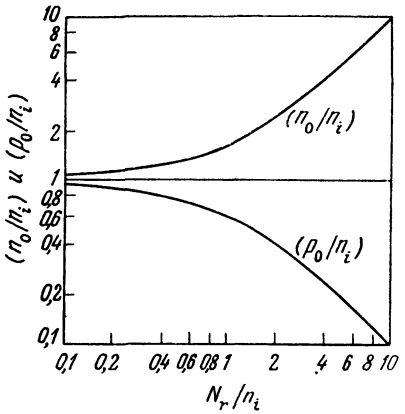


Рис. 4.10. Взаимосвязь концентраций свободных электронов и свободных дырок (в шкале, нормированной к собственной концентрации), когда избыток ионизованных донорных центров приводит к преобладанию электронов над дырками: $N_r = n_0 - p_0$. Эти кривые соответствуют соотношению (4.24). Для изображенной ситуации электроны являются основными носителями заряда, дырки — неосновными.

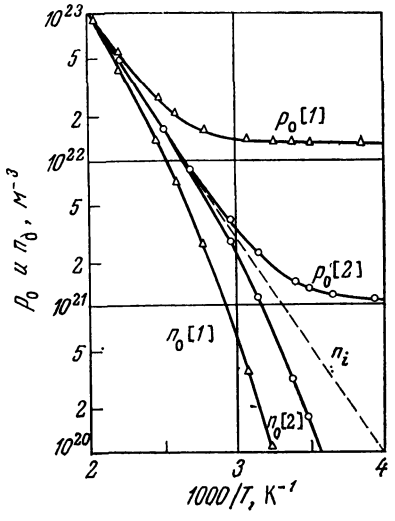


Рис. 4.11. Изменение с обратной температурой концентрации основных носителей заряда (дырок) и неосновных (электронов) для двух образцов InAs, которые приобретают проводимость *p*-типа при охлаждении. Ход собственной концентрации показан штриховой линией.

показанном на рисунке. Почти наверное концентрация основных носителей для каждого из образцов *уменьшится* при достаточном охлаждении (поскольку термическая ионизация даже довольно мелких примесных состояний не может осуществляться при самых низких температурах), однако такое снижение примесной концентрации основных носителей при 250 К еще не наблюдается. Соответственно концентрация неосновных носителей $n_0 = (n_i^2/p_0)$ должна следовать температурной зависимости n_i^2, и кривые для n_0 в нижней части рисунка имеют наклон, приблизительно равный $(-\varepsilon_i/k_0)$.

Донорные и акцепторные состояния

Конечно, можно определить число N_r как разность концентраций примесных центров, потерявших один электрон и принявших на себя один электрон. Однако чтобы понять, как и по-

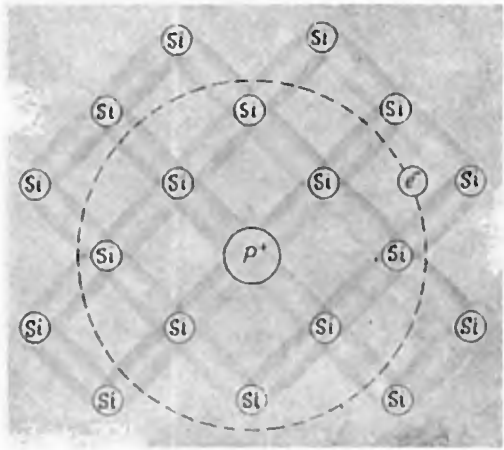


Рис. 4.12. Условное двумерное изображение части кристалла кремния. Кремний кристаллизуется в решетке алмаза (рис. 1.31, *а*), и каждый атом участвует в образовании ковалентных связей с четырьмя его ближайшими соседями. Атом фосфора показан замещающим один атом кремния, что приводит к образованию донорного центра замещения. Атом фосфора имеет один лишний электрон, не нужный для образования связей, и «донорный электрон» движется по большой орбите с малой энергией связи.

чему возникает разность N_r между концентрациями n_0 и p_0, мы должны иметь представление о природе и энергетическом спектре примесных состояний.

В качестве простого примера *донорного центра* в полупроводнике рассмотрим кристалл кремния, в котором один узел решетки занят атомом фосфора, как это схематически изображено на рис. 4.12. Видно, что каждый атом в кристалле образует четыре ковалентные связи со своими четырьмя ближайшими соседями. Это означает, что атом фосфора имеет один лишний электрон, не нужный для образования связей. Этот электрон не вполне свободен, потому что ядро атома фосфора обладает положительным зарядом на единицу бо́льшим, чем ядро атома кремния; избыточный электрон оказывается хотя и слабо, но связанным с окрестностью примесного дефекта. Следовательно, мы можем представить нейтральное состояние примесного центра (электрон на центре) *локализованным электронным состоянием* с энергией ниже зоны проводимости. На рис. 4.13 это состояние показано лежащим ниже дна зоны проводимости на величину энергии ε_d, это означает, что электрон из такого локализованного состояния может быть переведен в (нелокализованное) блоховское состояние, если ему сообщить дополнительную энергию, не меньшую чем ε_d. При этом донор

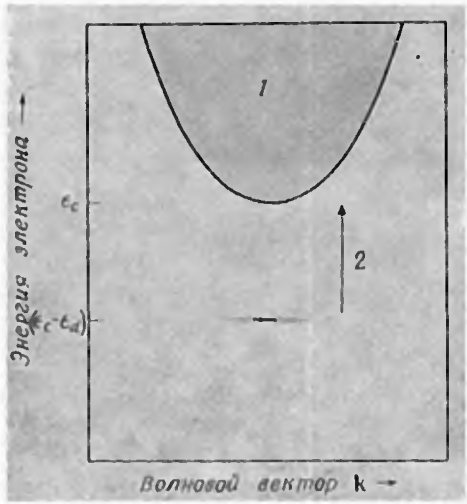


Рис. 4.13. Изображение состояния электрона на донорном центре с помощью уровня энергии, расположенного внутри запрещенной зоны. В соответствии с принципом неопределенности степень локализации, необходимая для образования связанного состояния, определяет ту область **k**-пространства, для которой его волновая функция отлична от нуля. *1* — зона проводимости; *2* — энергия связи электрона на доноре.

становится ионизованным, и на нем оказывается локализован положительный заряд.

В случае когда донором в кремнии является атом фосфора, энергия связи электрона на доноре мала, а орбита связанного электрона велика по сравнению с межатомным расстоянием. Тем не менее электрон, связанный на доноре, оказывается очень сильно ограниченным в реальном пространстве по сравнению с размером кристалла, поэтому на рис. 4.13 показано, что связанное состояние занимает конечную область **k**-пространства. Часто утверждается, что когда орбита связанного состояния велика по сравнению с межатомным расстоянием, величина кулоновского притяжения между примесным центром и связанным электроном ослаблена макроскопической диэлектрической проницаемостью полупроводника $\varkappa$. Используя аргументы такого рода, Бете (1942 г.) постулировал, что волновая функция донорного состояния может быть водородоподобной, видоизмененной только диэлектрической проницаемостью и перенормированной эффективной массой. В рамках этой модели (ссылка на которую была дана в примечании [57] гл. 3) не требуется знания природы (химической или металлургической)

данного центра, и уравнение Шредингера для такого водородоподобного центра может быть записано в виде

$$[(-\hbar^2/2m_c)\nabla^2-(e^2/4\pi\varkappa\varepsilon_0 r)]\psi=\varepsilon\psi. \tag{4.25}$$

Квантовая механика и полуклассическая модель Бора дают одни и те же значения для энергии связи основного состояния и «боровского радиуса» для такой пересчитанной водородной модели примеси:

$$\varepsilon_d=(m_c e^4/32\pi^2\hbar^2\varkappa^2\varepsilon_0^2),$$
$$a_d=(4\pi\varkappa\varepsilon_0\hbar^2/m_c e^2). \tag{4.26}$$

Имеется также водородоподобная последовательность возбужденных (но не ионизованных) состояний.

Мы увидим, что осложнения двух видов, одно — зависящее от характера зонной структуры, а другое — от природы примеси, часто приводят к пространственному перераспределению связанного заряда и делают энергию связи более сложной, чем дают формулы (4.25) и (4.26). Однако прежде чем переходить к проблемам и трудностям реальных донорных центров, удостоверимся в том, что и *акцепторные центры* также могут существовать.

На рис. 4.14 в условной форме изображен простейший вид акцепторной примеси замещения в кремнии — результат того, что атом бора занимает одно из регулярных мест в кремнии. Представляется естественным, что одна из четырех валентных связей, направленных к бору, должна остаться ненасыщенной, однако фактически положение отсутствующей связи может перемещаться из одного межатомного положения в другое, находясь только под действием экранированного кулоновского притяжения центрального отрицательного заряда. Наиболее просто такую ситуацию можно описать с помощью *связанной дырки,* движущейся в состоянии, зависящем от диэлектрической проницаемости, и с помощью тензора эффективной массы для свободных дырок. Дырка может стать полностью нелокализованной, если сообщить ей энергию ε_a; таким образом, нейтральное состояние акцептора мы представляем с помощью незанятого электронного состояния с энергией, лежащей на ε_a *выше* потолка валентной зоны. Такое условное представление изображено на рис. 4.15. Энергия связи ε_a и радиус a_a *водородоподобного* акцептора определяются теми же формулами (4.26), если заменить в них эффективную массу m_c на m_v.

Задача 4.5 рассматривает случай *мелких* донорных состояний (т. е. состояний с малой энергией связи и большим радиусом Бора) в условиях, при которых водородоподобное приближение уравнений (4.25) и (4.26) *может* быть справедливым.

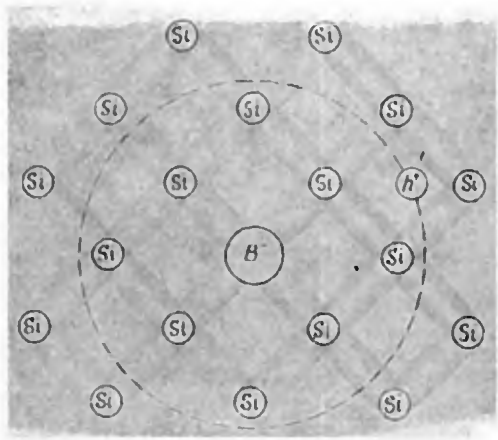


Рис. 4.14. Решетка кремния, как и на рис. 4.12, но кремниевый узел замещен теперь атомом бора, который имеет на один валентный электрон меньше, чем кремний. Одна валентная связь в окрестности атома бора должна быть незаполнена, в то время как вся область кристалла электрически нейтральна. Изменение положения этой незавершенной связи можно представить как орбитальное движение «связанной дырки» на акцепторном центре бора.

Однако, несомненно, такая модель является слишком упрощенной для большинства примесных центров как химической, так и металлургической природы в большинстве полупроводников.

Указывая, как может влиять структура зон на усложнение картины примесных состояний по сравнению с водородоподобной, Кон [11] замечает, что «донорное состояние можно рассматривать как волновой пакет, составленный в основном из блоховских волн, взятых вблизи дна зоны проводимости». Поэтому водородоподобные состояния, вытекающие из уравнения (4.25), могут иметь место, если эффективная масса m^* есть скаляр, не зависящий от энергии. Однако, когда эффективная масса представляет собой тензорную величину, требуется решить уравнение Шредингера гораздо большей сложности. Кон описал метод, с помощью которого может быть построена более сложная *теория эффективной массы* для мелких доноров и мел-

[11] *Kohn W.* Solid State Physics, eds. F. Seitz, D. Turnbull, Academic Press, 1957, v. 5. Теория эффективной массы для доноров в германии или кремнии осложнена тем, что в каждом из этих полупроводников любая поверхность постоянной энергии, расположенная чуть выше ε_c, состоит из нескольких вытянутых эллипсоидальных поверхностей, центры которых смещены от центра зоны Бриллюэна (ЗБ). Максимум валентной зоны попадает как раз в центр ЗБ, однако акцепторные состояния и в Si и в Ge осложняются тем, что *две* валентные зоны имеют общий максимум.

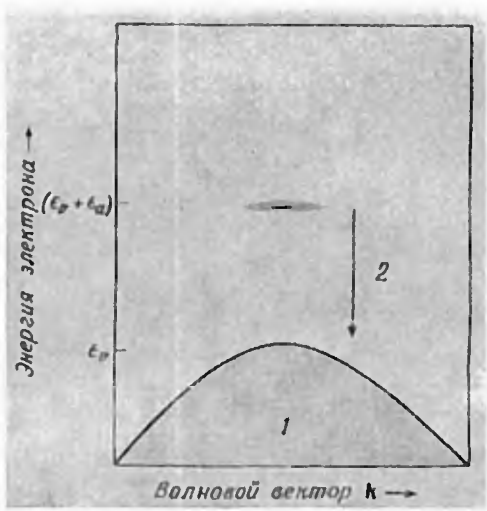


Рис. 4.15. Изображение состояния, связанного с однозарядным акцепторным центром. Нейтральный центр содержит вращающуюся вокруг него связанную дырку, и в этом случае *электронное состояние* с энергией $(\varepsilon_v+\varepsilon_a)$ является *незаполненным*. Центр становится ионизованным, когда он захватывает электрон, заполняя, таким образом, примесный уровень и оставляя пустым одно из нелокализованных состояний валентной зоны. Это не так просто понять при первом чтении! *1* — валентная зона; *2* — энергия связи дырки на акцепторе.

ких акцепторов в таких хорошо изученных полупроводниках, как германий и кремний. Только для очень небольшого числа полупроводников формы зон известны настолько хорошо, что допускают строгий расчет примесных состояний методом «эффективной массы».

Почти все посторонние атомы, входящие в полупроводник в качестве примеси, имеют энергии связи основного состояния ε_d или ε_a, *превышающие* значение, вычисленное с помощью метода эффективной массы. Добавка к энергии связи, определяемая конкретным видом примеси, называется *химическим сдвигом*. Таким образом, действительно успешная теоретическая модель примесных состояний в полупроводнике должна быть в состоянии объяснить химический сдвиг для каждого типа примесного атома, а также весь спектр возбужденных, но не ионизованных состояний для каждой примеси, поскольку каждая из них «химически сдвинута».

Причиной химического сдвига является отклонение притягивающего потенциала от величины $(e^2/4\pi\varkappa\varepsilon_0 r)$ на малых расстояниях от ядра примеси. Хотя на связанный электрон при его

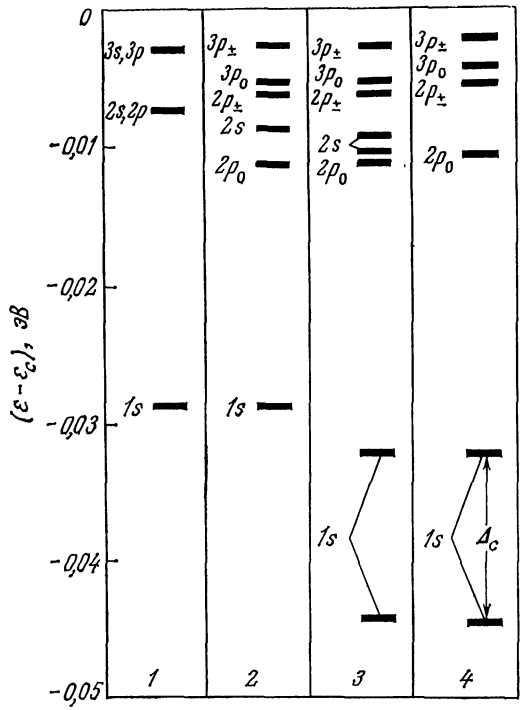


Рис. 4.16. Основное состояние и несколько первых возбужденных состояний донорного атома фосфора в кремнии, вычисленные, согласно водородной модели, с помощью теории эффективной массы Кона[11] и полученные из оптических [12] и электрических[13] экспериментов. *1* — водородная модель; *2* — теория эффективной массы· *3* — модифицированная теория эффективной массы; *4* — эксперимент.

орбитальном движении и действует кулоновское притяжение вида $(e^2/4\pi\varkappa\varepsilon_0 r_0^2)$ на расстояниях, больших по сравнению с межатомным, он испытывает гораздо более сильное притяжение на расстояниях, настолько малых, что макроскопическое понятие диэлектрической проницаемости теряет смысл. Это повышенное притяжение на малых расстояниях сильнее всего сказывается на состояниях с малым главным квантовым числом и особенно существенно для состояний с *s*-симметрией (поскольку в этом случае волновая функция не имеет узла в точке, где находится ядро).

В результате того, что эффективная масса представляет собой тензор, спектр возбужденных состояний примесного центра может иметь очень мало сходства с исправленным спектром возбужденных состояний нейтрального атома водорода. На

рис. 4.16 в левом столбце изображен водородоподобный спектр для мелкого донора в кремнии, в следующем столбце изображен спектр любого мелкого донора в кремнии, рассчитанный в модели эффективной массы Кона [11]. Эта модель (второй столбец) учитывает тензорный характер эффективной массы для зоны проводимости кремния (см. на рис. 3.50 вид поверхностей постоянной энергии вблизи дна этой зоны), однако все еще не учитывает специфики, вносимой химической природой атома (в данном случае *фосфора*). Заметим, что асимметрия эффективной массы приводит к тому, что энергия для $n=2$ или $n=3$ в значительной степени зависит от азимутального и магнитного квантовых чисел.

Третий столбец энергетических состояний, изображенных на рис. 4.16, получен с учетом того, что связанный электрон в своем орбитальном движении испытывает кулоновское притяжение на больших расстояниях и притяжение, гораздо более сильное, на расстояниях, настолько малых, что макроскопическое понятие диэлектрической проницаемости теряет смысл. Конкретная природа примесного центра определяет, в какой степени кристалл оказывается искажен его присутствием и при каком радиусе не действует представление о кулоновском притяжении в диэлектрической среде. Поэтому $2p$- и $3p$-состояния других донорных центров в кремнии довольно близки к тем, которые изображены здесь для фосфора, однако $1s$-состояния сильно различаются.

Электрон, находящийся в основном состоянии фосфорного донорного центра, обладает степенью свободы в отношении того, какую долю времени он находится вблизи ядра, и это приводит к дальнейшему расщеплению основного состояния, как показано в третьем столбце рис. 4.16 и что подтверждается экспериментальными результатами (четвертый столбец). Самое нижнее основное состояние и возбужденные состояния p-типа были определены в большом числе экспериментов по измерениям длин волн оптического поглощения. Здесь показаны энергии, полученные Аггарвалом и Рамдасом [12]. Переходы в состояния s-типа, лежащие выше основного состояния, являются оптически запрещенными, однако энергетический зазор Δ_c между основным и «верхним основным состоянием» был получен из электрических измерений Лонга и Майерса [13]. Позднее мы вернемся к их эксперименту.

Наличия тонкой структуры основного состояния и в спектре возбужденных состояний для «мелких центров» следует ожидать в большинстве полупроводников. Задача становится еще

[12] *Aggarwal R. L., Ramdas A. K.*— Phys. Rev., **137**, A602 (1965).
[13] *Long D., Myers J.*— Phys. Rev., **115**, 1119 (1959).

труднее для состояний глубоких центров, тех, которые имеют энергию связи, составляющую значительную долю от полной ширины запрещенной зоны. Основное состояние глубокого центра сильно локализовано и представление о диэлектрической проницаемости становится неприменимым *. Одной из возможных модификаций модели эффективной массы, которая представляется обнадеживающей для более глубоких примесных состояний, является *«метод квантовых дефектов»*, который приписывает каждому связанному состоянию нецелое квантовое число (обычный квантовый номер минус «квантовый дефект»)[14].

Сложность спектра многократно возрастает для *многовалентного* ** (многозарядного) центра, который может связывать два или более электронов (или дырок) с различными энергиями связи. В рамках представлений, развитых выше для атома фосфора как центра замещения в кремниевой решетке, можно представить себе, что атом серы в кристалле кремния может легко отдать один электрон и с несколько большим затруднением — другой. Процесс грубо можно сравнить с последовательными состояниями ионизации изолированного атома гелия. В то время как сера в кремнии ведет себя как *двойной донор,* такой элемент, как цинк, ведет себя как *двойной акцептор.*

Еще более сложными являются *амфотерные* центры, которые стремятся приобрести положительный заряд, когда уровень Ферми расположен низко, и проходят через нейтральное состояние к состоянию однократно или многократно заряженного отрицательного иона, когда электрохимический потенциал достаточно высок. Так происходит во многих полупроводниках при растворении в них атомов некоторых благородных металлов или переходных элементов.

Статистика заполнения донорных и акцепторных состояний

Мы воспользовались законом заполнения Ферми — Дирака (3.45), или (4.3), чтобы описать вероятность того, что *любое* блоховское состояние в зоне окажется занятым при данной

---

* В ряде работ показано, что волновая функция локализована в пределах второй координационной сферы, а на таких расстояниях диэлектрическая проницаемость почти не отличается от макроскопического значения.— *Прим. ред.*

[14] Модель квантовых дефектов и другие приближения для состояний глубоких центров обсуждаются Беббом и Чепменом. *Bebb H. B, Chapman R. A.* Proceedings 3rd International Conference on Photoconductivity, ed. E. M. Pell, Pergamon, 1971, p. 245.

** В отечественной литературе общеупотребителен термин «многозарядный центр».— *Прим. перев.*

температуре и данном значении $(\varepsilon - \varepsilon_F)$. Этот закон заполнения используется для любого состояния в зоне независимо от того, заполнено или незаполнено парное ему состояние с таким же волновым вектором и энергией, но противоположным спином. Поскольку волновые функции блоховских состояний простираются по всему кристаллу, присутствие одного электрона с определенным значением волнового вектора не оказывает сколько-нибудь заметного влияния на возможность заполнения другого (с противоположным спином) состояния с таким же волновым вектором $\mathbf{k}$; дело в том, что эти два электрона очень редко будут находиться в одной и той же области реального пространства. Поэтому для *зонных состояний* каждая точка в $\mathbf{k}$-пространстве отвечает двум определенным и порознь заполняемым состояниям, каждое из которых обладает статистическим весом $g=1$.

В отличие от этого рассмотрим локализованный примесный уровень, например, основное состояние донора. Такой уровень обычно обладает статистическим весом g, превышающим единицу. Так, для s-образного основного состояния «водородоподобного центра» существуют две возможности ориентации спина (вверх и вниз), и мы должны для этого состояния положить $g=2$. Однако *только одно из этих состояний может быть занято электроном*, поскольку связыванию второго электрона будет препятствовать кулоновское отталкивание электрона, уже находящегося в связанном состоянии с *той или иной* ориентацией спина.

Для донорного центра более общего типа полную волновую функцию электрона на центре (с учетом разных вариантов орбитального и спинового движения) можно описать g_n способами для невозбужденного нейтрального центра и g_i способами, когда электрон с центра удален. Позже в этом разделе мы обратимся к вопросу о том, что происходит, когда существует заметная вероятность того, что центр нейтрален, но находится в возбужденном состоянии. Сейчас же мы, исключая эту возможность, рассмотрим соотношение между N_{di} ионизованными центрами и N_{dn} невозбужденными нейтральными центрами, на которых связанный электрон обладает энергией $[(\varepsilon_c - \varepsilon_d) - \varepsilon_F]$, относительно энергии Ферми. Итак, при температуре T

$$\left\{\frac{N_{dn}}{N_{di}}\right\} = \left\{\frac{g_n \exp\left[\dfrac{\varepsilon_F - (\varepsilon_c - \varepsilon_d)}{k_0 T}\right]}{g_i}\right\}. \tag{4.27}$$

Помня о том, что мы пренебрегли возбужденными состояниями, можно переписать (4.27) в таком виде, чтобы выразить

часть полной концентрации доноров, которая сохраняет электрон с той или иной ориентацией спина:

$$\frac{N_{dn}}{N_{dn}+N_{dl}} \equiv \frac{N_{dn}}{N_d} = \frac{1}{1+(g_l/g_n)\exp\left[\dfrac{\varepsilon_c-\varepsilon_d-\varepsilon_F}{k_0T}\right]}. \qquad (4.28)$$

Это эквивалентно утверждению: вероятность $P_e(\varepsilon_c-\varepsilon_d)$ того, что *любой* из совокупности доноров с энергией основного состояния $(\varepsilon_c-\varepsilon_d)$ удерживает свой связанный электрон, равна

$$P_e(\varepsilon_c-\varepsilon_d) = \frac{1}{1+\left(\dfrac{g_l}{g_n}\right)\exp\left[\dfrac{\varepsilon_c-\varepsilon_d-\varepsilon_F}{k_0T}\right]} = \frac{1}{1+\beta\exp\left[\dfrac{\varepsilon_c-\varepsilon_d-\varepsilon_F}{k_0T}\right]}. \qquad (4.29)$$

Величину $\beta \equiv (g_l/g_n)$ часто называют *фактором спинового вырождения*. Для центра донорного типа величина β обычно меньше единицы, т. е. вероятность заполнения $P_e(\varepsilon_c-\varepsilon_d)$ в том случае, когда энергия $(\varepsilon_c-\varepsilon_d)$ совпадает с энергией Ферми, составляет более 50 %.

Аналогичной терминологией можно воспользоваться, чтобы выразить вероятность того, что акцепторное состояние с энергией $(\varepsilon_v+\varepsilon_a)$ является ионизованным (т. е. содержит *электрон*)

$$P_e(\varepsilon_v+\varepsilon_a) = \frac{1}{1+\beta\exp\left[\dfrac{\varepsilon_v+\varepsilon_a-\varepsilon_F}{k_0T}\right]}. \qquad (4.30)$$

Отношение β статистических весов для нейтрального и отрицательно заряженного состояний акцепторного центра в этом случае больше единицы. Вероятность того, что акцептор окажется *нейтральным* (т. е. на нем будет связанная дырка), есть

$$P_h(\varepsilon_v+\varepsilon_a) \equiv [1-P_e(\varepsilon_v+\varepsilon_a)] = \frac{1}{1+\beta^{-1}\exp\left[\dfrac{\varepsilon_F-\varepsilon_v-\varepsilon_a}{k_0T}\right]}. \qquad (4.31)$$

По аналогии с (4.29) для донорных состояний находим, что вероятность $P_h(\varepsilon_v+\varepsilon_a)$ для акцептора составляет более 50 %, когда ε_F совпадает с $(\varepsilon_v+\varepsilon_a)$.

Компенсация

Если подготовка, очистка, выращивание полупроводникового кристалла не выполнены с *чрезвычайной* тщательностью, то скорее всего в нем будут присутствовать в значительной концентрации несколько типов донорных и акцепторных центров с различным энергетическим положением связанных состояний *. В термодинамическом равновесии для описания всех электронных состояний в кристалле, локализованных и нелокализованных, соответствующих любым энергиям имеется единственный нормировочный параметр, или энергия Ферми.

Фактически, конечно, происходит то, что каждый донорный центр с относительно высокой характерной энергией стремится оказаться пустым, отдав свой электрон одному из имеющихся нижележащих акцепторов (рис. 4.17). Говорят, что присутствие состояний с низкими энергиями *компенсирует* (частично или полностью) присутствие донорных состояний с более высокими энергиями. Условие полной электрической нейтральности в состоянии равновесия имеет вид

$$(n_0 - p_0) = N_r = \sum_j N_{dj}[1 - P_e(\varepsilon_{dj})] - \sum_k N_{ak} P_e(\varepsilon_{ak}), \qquad (4.32)$$

и именно это условие требует единой энергии Ферми для каждой конкретной температуры в данном полупроводнике с данным распределением центров.

Первый член правой части уравнения (4.32) представляет собой суммарную потерю электронов j сортами доноров; второй член представляет собой суммарное число электронов, принятых k сортами акцепторов. Разница этих сумм есть величина, которую в формулах (4.23) и (4.24) мы назвали N_r, обозначающая интегральное действие примесных центров в образовании разницы между величинами n_0 и p_0. Впервые введя N_r, мы предпочли не выяснять, каким образом конкретные центры могут оказаться либо нейтральными, либо ионизованными. Однако теперь настало время рассмотреть, как степень заполнения совокупности центров может изменяться с температурой, и проследить, как это влияет на концентрацию свободных основных носителей заряда.

Конечно, донорные состояния, расположенные намного ниже энергии Ферми, дадут пренебрежимо малый вклад в уравнение (4.32), равно как и акцепторные состояния, лежащие достаточно высоко над уровнем Ферми. Донорные состояния, распо-

* Эффект компенсации сегодня используется специально в ряде полупроводниковых материалов и приборов, например, с помощью компенсации достигаются полуизолирующие свойства GaAs и GaP для подложек, применяемых в эпитаксии.— *Прим. ред.*

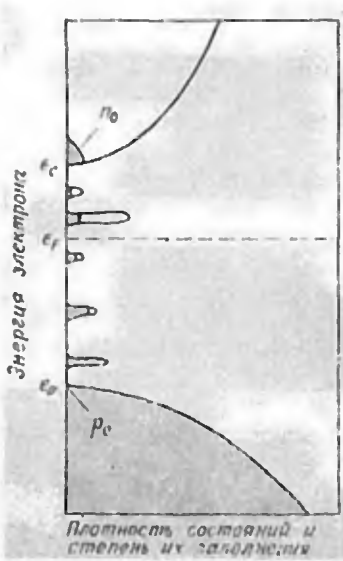


Рис. 4.17. Компенсированный полупроводник, содержащий несколько типов донорных и акцепторных центров, обладающих различными энергиями в запрещенной зоне. Донорные состояния при высоких энергиях легко ионизуются, ионизуя при этом любые акцепторные состояния с более низкой энергией. Концентрации и степень компенсации примесных состояний определяют равновесные концентрации n_0 и p_0.

ложенные высоко над ε_F, и акцепторные состояния, лежащие гораздо ниже ε_F, дают значительный, но не зависящий от температуры вклад в это уравнение. Обычно существует реальная возможность исследовать температурную зависимость заполнения состояний для центров, ближайших по энергии к энергии Ферми; при этом предполагается, что заполнение всех прочих примесных состояний относительно нечувствительно к изменениям как T, так и ε_F. Такое предположение не внесет заметной ошибки, если все эти примесные состояния расположены по крайней мере на $3k_0T$ выше или ниже энергии Ферми.

Концентрация свободных электронов в компенсированном полупроводнике

В этом разделе мы хотим рассмотреть уравнение (4.32) в условиях, когда в достаточно широком интервале температур изменяется заполнение только *одной* зоны и только *одного* сорта примесных центров. Так может происходить в ситуации, представленной на рис. 4.18. Эта ситуация задана действием N_d донорных примесных центров с уровнем энергии на ε_d ниже зоны проводимости. Действие всех прочих примесных центров, отличных от данного вида, описывается «полной концентрацией компенсирующей примеси» N_a. Для простоты предполагается, что ширина запрещенной зоны очень велика и $p_0=0$ во всем

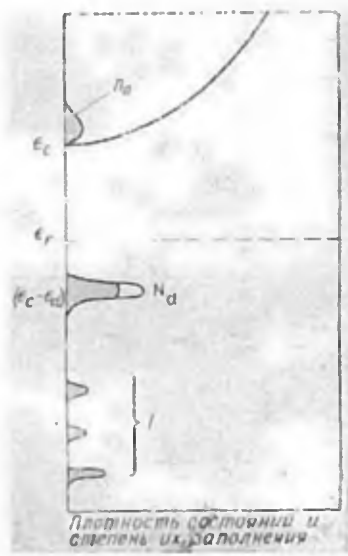


Рис. 4.18. Упрощенная диаграмма состояний для примесного полупроводника n-типа, в котором действуют донорные центры одного типа, частично компенсированные. *1* — полная концентрация всех акцепторов N_a.

рассматриваемом интервале температур. Тогда уравнение (4.32) упрощается к виду

$$n_0 = N_d\,[1 - P(\varepsilon_d)] - N_a. \tag{4.33}$$

Это выражение можно переписать следующим образом:

$$(N_d - N_a - n_0) = N_d P_e(\varepsilon_d) = \frac{N_d}{1 + \beta \exp\left[\dfrac{\varepsilon_c - \varepsilon_d - \varepsilon_F}{k_0 T}\right]}. \tag{4.34}$$

Далее предположим, что число свободных электронов, обусловленных примесными уровнями, достаточно мало, чтобы неравенство $n_0 \ll N_c$ выполнялось при всех температурах. (N_c есть эффективная плотность состояний в зоне проводимости.) Тогда можно воспользоваться очень простой связью (4.9) между концентрацией свободных электронов и соответствующей энергией Ферми:

$$n_0 = N_c \exp\left(\frac{\varepsilon_F - \varepsilon_c}{k_0 T}\right). \tag{4.35}$$

При одновременном выполнении соотношений (4.34) и (4.35) получаем два условия, связывающие n_0 и ε_F. Совместное их решение приводит к квадратному уравнению либо для n_0, либо для $(\varepsilon_F/k_0 T)$. Выбирая первое, мы можем сказать, что

$$\frac{n_0 (N_a + n_0)}{(N_d - N_a - n_0)} = \beta N_c \exp(-\varepsilon_d/k_0 T), \quad n_0 \ll N_c. \tag{4.36}$$

Аналогичное квадратное уравнение для концентрации свободных дырок, возбужденных с N_a акцепторов, обладающих энергией ионизации ε_a и частично компенсированных присутствием N_d доноров, имеет вид

$$\frac{p_0(N_d + p_0)}{(N_a - N_d - p_0)} = (N_v/\beta)\exp(-\varepsilon_a/k_0T), \quad p_0 \ll N_v. \qquad (4.37)$$

Уравнения (4.36) и (4.37) совершенно эквивалентны в отношении их зависимости от температуры и степени компенсации (N_a/N_d), и оба уравнения широко используются. В последующем обсуждении мы будем говорить об электронах и уравнении (4.36), поскольку именно этот случай изображен на рис. 4.18.

В полупроводнике при тепловом равновесии соблюдается *детальное равновесие* между любым процессом превращения энергии и обратным ему процессом. В результате процессов превращения, приводящих к обмену энергией между электронами и колебаниями решетки (фононами) и фоновым излучением черного тела (фотонами) в кристалле, электроны постоянно возбуждаются в зону проводимости и постоянно падают из этой зоны. Скорости *генерации электронов* и *рекомбинации электронов* в тепловом равновесии равны, и поэтому наблюдаемая величина n_0 не зависит от времени. Если в полупроводнике действуют компенсирующие друг друга центры, то величина n_0 будет удовлетворять уравнению (4.36). Уравнение (4.36) для электронов и соответствующее ему уравнение (4.37) для дырок называют уравнениями *действующих масс*, потому что они записаны в такой форме, которая ясно отражает закон действующих масс и принцип детального равновесия при установлении динамического равновесия между генерацией и рекомбинацией. Более подробно обсуждение процессов генерации и рекомбинации будет проведено в разд. 4.4.

Соотношения (4.36) и (4.37) представляют собой относительно n_0 и p_0 соответственно легко разрешаемые квадратные уравнения, решение которых позволяет узнать, как меняется концентрация свободных носителей с температурой. Тогда обращение выражения (4.35) позволяет проследить движение энергии Ферми с температурой. Задача 4.6 представляет собой простое упражнение по определению n_0 и ε_F в зависимости от температуры для случая невырожденного примесного полупроводника и напоминает о том, что предельный случай отсутствия вырождения, которого требует уравнение (4.36), может перестать выполняться.

Рис. 4.19 иллюстрирует изменение n_0 с обратной температурой для частично компенсированного полупроводника n-типа. На этом рисунке представлены данные, относящиеся к образ-

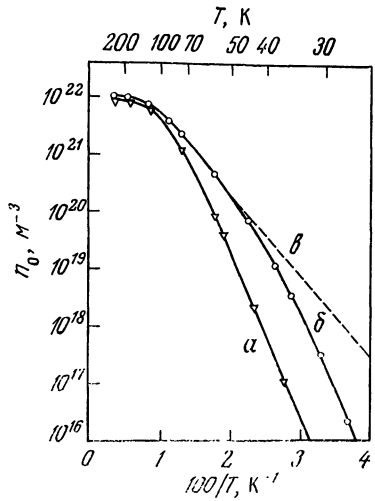


Рис. 4.19. Характер изменения концентрации свободных электронов с температурой в примесной ситуации с компенсацией, определяемый соотношением (4.36). Экспериментальные точки для образцов *а* и *б* получены из измерений эффекта Холла на образцах кремния, легированных донорной примесью мышьяка и содержащих меньшие количества неизвестных компенсирующих центров. Кривые удовлетворяют уравнению (4.36) с правой частью $2{,}7\times\times10^{21}\ T^{3/2}\exp(-615/T)$ м$^{-3}$, что соответствует $m_c=1{,}08m$, $\beta=1/2$ и $\varepsilon_d=0{,}053$ эВ. Подгонка для образца *а* требует $N_d=1{,}1\cdot10^{22}$ м$^{-3}$ и $N_a=2\times\times10^{21}$ м$^{-3}$. Для менее компенсированного образца *б* необходимые параметры составляют $N_a=1{,}02\cdot10^{22}$ м$^{-3}$ и $N_a=3\cdot10^{19}$ м$^{-3}$. Задача 4.7 касается температурной зависимости энергии Ферми для образцов *а* и *б*.

цам кремния, легированного мышьяком. Каждый из образцов содержит около 10^{22} м$^{-3}$ донорных центров мышьяка, однако в образце *а* больше компенсирующих центров, чем в образце *б*. Сплошные линии на рис. 4.19 отвечают решению уравнения (4.36) для значений N_a и N_d, приведенных в подписи к рисунку; пунктирная кривая *в* представляет собой экстраполяцию для образца, аналогичного *б*, но с *нулевой* компенсацией.

В задаче 4.7 исследуется вопрос о том, как заполнение связанных состояний донора, лежащих выше основного состояния, *могло бы* в некоторой части температурного интервала оказать небольшое влияние на результаты рис. 4.19. Однако в общем сравнении экспериментальных кривых на рисунке с их ходом, предсказанным уравнением (4.36), этим осложнением можно пренебречь.

Уравнение (4.36) принимает несколько важных предельных форм, которые можно проиллюстрировать с помощью рис. 4.19. Прежде всего можно видеть, что n_0 приближается к предель-

ному максимальному значению $(N_d - N_a)$ для температур больших, чем $(\varepsilon_d/3k_0)$. (На рис. 4.19 это имеет место для температур выше 200 К.) Высокотемпературная область обычно называется *областью истощения* системы, поскольку никакого дальнейшего изменения в концентрации электронов не может произойти, если только уровень Ферми не опустится настолько низко, что начнется опустошение других типов центров или интенсивное собственное возбуждение.

Для противоположного случая очень низких температур выражение (4.36) упрощается к виду

$$n_0 = [\beta N_c (N_d - N_a)/N_a] \exp(-\varepsilon_d/k_0 T), \text{ если } n_0 \ll N_a < N_d. \tag{4.38}$$

Условия, необходимые для этой асимптотической формы, реализуются ниже 70 К для образца *а* на рис. 4.19 и ниже 35 К для образца *б*. В этом низкотемпературном режиме $\ln(n_0)$ представляет собой в основном линейную функцию от $(1/T)$ с наклоном $-\varepsilon_d/k_0$. Более слабая часть температурной зависимости обусловлена зависимостью $T^{3/2}$ для величины N_c.

Промежуточный температурный интервал, в котором $\ln(n_0)$ также изменяется почти линейно с $(1/T)$, но с энергией активации, составляющей только $(\varepsilon_d/2)$, можно видеть для образца *б* на рис. 4.19. Анализ уравнения (4.36) показывает, что для образца с очень слабой степенью компенсации *должен* существовать интервал температур, для которого

$$n_0 = (\beta N_c N_d)^{1/2} \exp(-\varepsilon_d/2k_0 T), \text{ если } N_a \ll n_0 \ll N_d, \tag{4.39}$$

и образец *б* удовлетворяет обоим этим неравенствам в интервале температур примерно от 45 до 90 К. Пунктирная кривая *в* на рис. 4.19 показывает ход зависимости (4.39), продолженной в область низких температур, как это было бы в идеализированном случае кремниевого образца, легированного мышьяком, совсем не содержащего компенсирующих центров.

Для полупроводникового образца с относительно сильной компенсацией основных центров температурный интервал, к которому может быть применимо выражение (4.39), исчезает полностью. Действительно, для сильно компенсированного кристалла выражение

$$n_0 \approx \frac{(N_d - N_a)}{1 + (N_a/\beta N_c) \exp(\varepsilon_d/k_0 T)} \tag{4.40}$$

становится удовлетворительным приближением для *всего* температурного интервала.

Соотношение между связанными электронами и электронами проводимости будет изменяться от так называемых *условий*

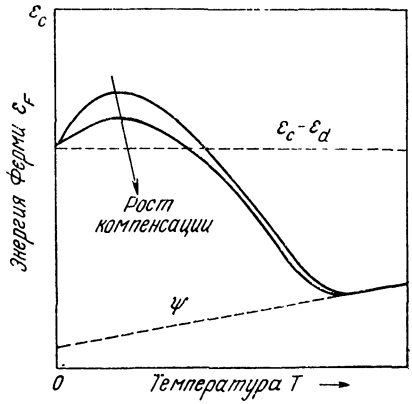


Рис. 4.20. Зависимость электрохимического потенциала ε_F от температуры, отражающая уменьшение заселенности одного сорта частично компенсированных доноров, отдающих свои электроны в зону проводимости. Возникающая в результате этого концентрация свободных электронов дается уравнением (4.36), а изменение ε_F во всем интервале температур может быть описано уравнением (4.41). Подъем в левой части рисунка соответствует выражению (4.42), а последующий спад ε_F с ростом температуры соответствует области истощения, описываемой выражением (4.43). (Стрелка показывает направление сдвига зависимости при увеличении числа компенсирующих центров). Конечно, уровень Ферми должен совпадать с положением уровня Ферми ψ в собственном полупроводнике при достаточно высоких температурах.

резервуара $[n_0 \ll N_d - N_a]$ при низких температурах до *условий истощения* $[n_0 \approx N_d - N_a]$ при достаточно высоких температурах в зависимости от того, мала или велика компенсация в кристалле. Увеличение концентрации свободных электронов при нагревании, конечно, может быть равным образом описано как температурное изменение энергии Ферми, которую, обращая равенство (4.35), можно представить в виде

$$\varepsilon_F = \varepsilon_c - k_0 T \ln (N_c/n_0). \tag{4.41}$$

Можно конечно, решая совместно (4.34) и (4.35), получить квадратное уравнение относительно $\exp(\varepsilon_F/k_0T)$, как это сделано в книге «Статистика полупроводников». Однако в большинстве случаев ε_F определяют с помощью соотношения (4.41) после того, как величина n_0 найдена из уравнения (4.36) или непосредственно из экспериментальных данных. Последняя процедура предлагается в задаче 4.7 для определения максимальной энергии Ферми при изменении температуры для образцов рис. 4.19.

Общий характер изменения ε_F с температурой, когда опустошается определенный тип центров, схематически изображен на рис. 4.20. Имеют место два предельных вида поведения ε_F:

$$\varepsilon_F = (\varepsilon_c - \varepsilon_d) + k_0 T \ln[\beta (N_d - N_a)/N_a]$$
при малых T, $n_0 \ll N_a < N_d$ (4.42)

и

$$\varepsilon_F = \varepsilon_c - k_0 T \ln[N_c/(N_d - N_a)] \text{ при больших } T,\ n_0 \approx N_d - N_a, \quad (4.43)$$

однако простого и удобного выражения, пригодного во всем температурном диапазоне, не существует. Этот вопрос обсуждается в «Статистике полупроводников» более детально, включая те осложнения, которые возникают, когда величина n_0 становится сравнима с N_c и ситуация в зоне проводимости становится «полувырожденной».

Сложности в анализе процесса заполнения примесных уровней

По мере повышения температуры полупроводника электроны энергетически поднимаются из примесных состояний в состояния проводимости и в конце концов устанавливается собственная проводимость, если полупроводник до этого не расплавится. В последнем разделе описан простейший вариант такого термостимулированного роста концентрации свободных носителей. Для многих систем примесных центров в полупроводнике этот рост происходит совершенно иначе из-за множества осложняющих факторов. Некоторые из них затронуты в данном разделе, который может быть опущен читателем, чьи интересы не лежат в данной области. Последующие разделы, посвященные полупроводникам, не связаны с этим материалом.

Влияние возбужденных состояний

Одним из таких осложняющих факторов может быть существование большого числа центров, находящихся в некоторой области температур в состоянии возбуждения, но не ионизации. На рис. 4.21 изображен полупроводник n-типа, в котором действует N_d донорных центров и N_a компенсирующих акцепторов разной природы. На этом же рисунке показано несколько возбужденных состояний электронов, связанных на донорах. Донорный центр может быть нейтрален, если его электрон находится в одном из β_1 состояний, отвечающих энергии $(\varepsilon_c - \varepsilon_d)$, но он может быть нейтрален и в случае, когда тот же электрон находится в одном из β_r состояний, отвечающих энергии ε_{r1},

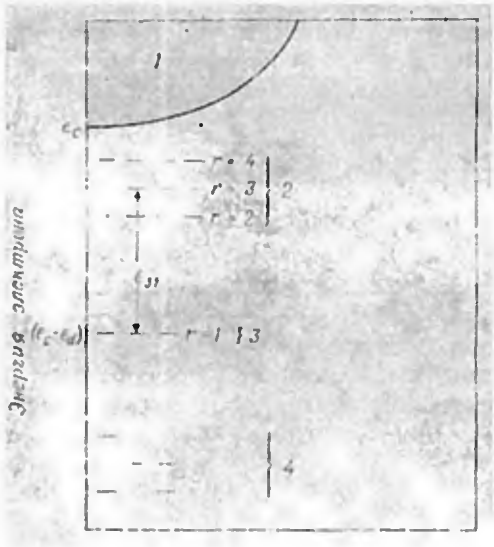


Рис. 4.21. Несколько более сложная, чем на рис. 4.18, модель спектра энергетических состояний для полупроводника n-типа, в котором действует N_d донорных центров одного типа и N_a компенсирующих центров разных типов. На рисунке также показано несколько первых возбужденных состояний каждого донорного центра. Когда электрон связан на доноре в *любом* из β_1 основных состояний или в *любом* из β_r возбужденных состояний, находящихся при энергии $(\varepsilon_c - \varepsilon_d + \varepsilon_{r1})$, кулоновское отталкивание препятствует заполнению любого из оставшихся состояний вторым электроном. *1* — зона проводимости; *2* — возбужденные состояния донора; *3* — основное состояние донора; *4* — уровни компенсирующих акцепторов с суммарной концентрацией N_a.

расположенной выше основного состояния. (Здесь r есть целое число, нумерующее при движении вверх, начиная с $r=2$, различные возбужденные состояния центра.)

Если учесть различные возбужденные состояния, равновесный статистический вес видоизменяется таким образом, что при данной энергии Ферми доля нейтральных центров может стать больше. Только в том случае, когда каждое возбужденное состояние расположено по крайней мере на $3k_0T$ выше уровня Ферми, существованием этих состояний можно пренебречь. Как указано в «Статистике полупроводников», весовой вклад более высоких связанных состояний следующим образом видоизменяет уравнение (4.36):

$$\frac{n_0\,(N_a + n_0)}{N_d - N_a - n_0} = \frac{\beta_1 N_c \exp\left(-\varepsilon_d/k_0T\right)}{1 + \sum\limits_{r=2}^{\infty} (\beta_1/\beta_r) \exp\left(-\varepsilon_{r1}/k_0T\right)}. \qquad (4.44)$$

Знаменатель правой части этого уравнения всегда больше единицы за исключением случая, когда *все* возбужденные состояния лежат на несколько k_0T выше энергии Ферми и становится справедлив простой анализ предыдущего раздела. Когда знаменатель в правой части (4.44) значительно больше единицы, концентрация свободных носителей становится меньше, чем следовало бы ожидать из уравнения (4.36) при той же температуре.

Расщепление основного состояния кристаллическим полем

Из третьего и четвертого столбцов (рис. 4.16) видно, что донорные атомы фосфора в кремнии существуют с двумя возможными энергиями основного 1*s*-состояния, разделенными энергией Δ_c. Верхние основные состояния действуют точно так же, как возбужденные состояния с главным квантовым числом $n>1$, давая донору возможность становиться возбужденным, но не ионизованным. Этот эффект опять приводит к тому, что при данной температуре величина n_0 оказывается меньше, чем она была бы в отсутствие каких-либо связанных состояний, кроме одного нижнего. Уравнение (4.36) должно быть теперь заменено некоторой модификацией уравнения (4.44):

$$\frac{n_0\,(N_a+n_0)}{(N_d-N_a-n_0)}=\frac{\beta_1 N_c \exp\,(-\varepsilon_d/k_0T)}{1+(\beta_1/\beta_{ug})\exp\,(-\Delta_c/k_0T)+\Sigma}, \qquad (4.45)$$

где знак суммы в знаменателе правой части означает такое же суммирование, как в (4.44). Новый член в знаменателе обусловлен вкладом верхних основных состояний.

Расщепление основного состояния, подобное тому, которое имеет место для фосфора в кремнии, наблюдается для многих центров в полупроводниках. На рис. 4.22 показаны экспериментальные данные для Si — P. Из кривых на этом рисунке видно, как верхние основные состояния понижают величину n_0 в широком интервале температур. Этими состояниями можно пренебречь при очень низких температурах, поскольку в этом случае почти все электроны «заморожены» на самом нижнем из возможных связанных состояний. Аналогичным образом и верхнее и нижнее основные состояния (как и все возбужденные) оказываются очень слабо заполненными при высоких температурах, поскольку все связанные состояния в этом случае уже не играют роли на фоне большого числа ставших доступными для электронов состояний в зоне проводимости. Это создает ситуацию истощения, когда $n_0\approx(N_d-N_a)$. Соотношения (4.36) и (4.45) приводят к различающимся результатам для промежуточных температур, для которых имеется частичное заполнение

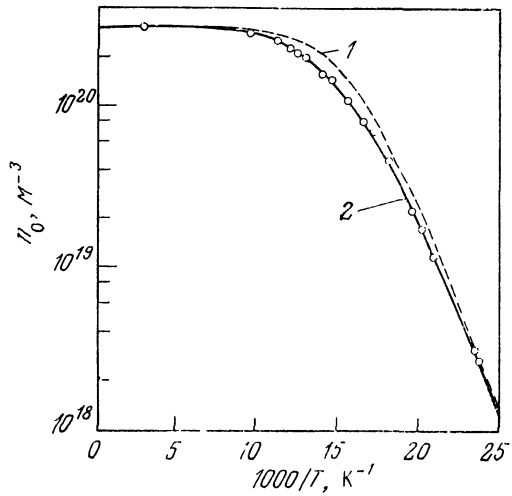


Рис. 4.22. Температурная зависимость концентрации свободных электронов в частично компенсированном кристалле кремния, легированного фосфором, по данным Лонга и Майерса [Long D., Myers J.— Phys. Rev, 115, 1119 (1959)]. Кривая *2* обнаруживает соответствие соотношению (4.45) для $N_d = 6{,}9 \cdot 10^{20}$ м$^{-3}$, $N_a = 3{,}8 \cdot 10^{20}$ м$^{-3}$ и $\varepsilon_d = 0{,}0435$ эВ, $\Delta_c = 0{,}010$ эВ. Кривая *1* вычислена в соответствии с (4.36) для тех же значений N_d, N_a и ε_d.

верхних 1*s*-состояний за счет состояний зоны проводимости. Лонг и Майерс [13] интерпретировали свои экспериментальные результаты для n_0, найденные из эффекта Холла (рис. 4.22), с помощью соотношения (4.45) и смогли определить значение Δ_c, которое удовлетворительно согласуется с оптическими измерениями этой величины [12].

Таким образом, распределение данного числа электронов между примесными центрами и зоной проводимости может стать более сложным, если оказываются заметно заполнены возбужденные состояния центров или если кристаллическое поле снимает вырождение основных состояний. Аналогичное расщепление может происходить в сильном магнитном поле или при анизотропной деформации полупроводника. Для большинства (но не для всех) полупроводников изотропная деформация не приводит к сильному изменению структуры примесных уровней.

Влияние двух или большего числа типов центров

Следует рассмотреть различного рода осложнения, которые возникают, когда вклад двух или нескольких видов примесных центров в заселенность зоны проводимости свободными элек-

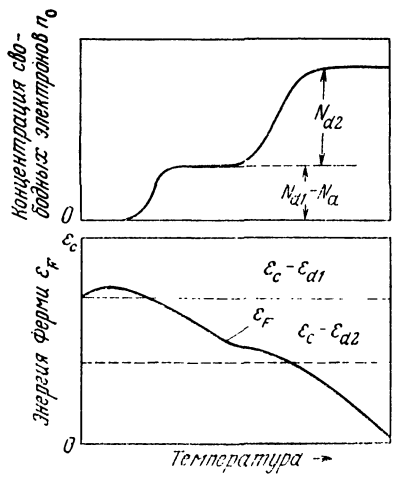


Рис. 4.23. Схема зависимости n_0 и ε_F от температуры для примесного полупроводника n-типа, в котором действуют два типа донорных центров, из которых более мелкие частично компенсированы.

тронами изменяется с изменением температуры. Итак, предположим, что в полупроводнике n-типа имеется N_{d1} примесных донорных состояний с энергией ионизации ε_{d1} и N_{d2} совершенно других донорных центров с несколько большей энергией связи ε_{d2}. Как обычно, в нем присутствует некоторая концентрация N_a различных компенсирующих центров, и для полупроводника n-типа мы требуем, чтобы N_a было меньше, чем $(N_{d1}+N_{d2})$[15]. Тогда распределение всех имеющихся электронов в тепловом равновесии удовлетворяет равенству

$$n_0 = N_c \exp(\varepsilon_F/k_0T) = N_{d1}[1-P_e(\varepsilon_{d1})] + N_{d2}[1-P_e(\varepsilon_{d2})] - N_a \tag{4.46}$$

при условии, что $n_0 \ll N_c$ и что возбужденными состояниями различных центров можно пренебречь. Формально уравнение (4.46) представляет собой кубическое уравнение относительно n_0 или $\exp(\varepsilon_F/k_0T)$, однако оно упрощается к квадратному, если зазор между двумя энергиями примесных центров велик по сравнению с k_0T. На рис. 4.23 изображен типичный характер возрастания n_0 при нагревании (сопровождаемого температурным изменением электрохимического потенциала) для двух

[15] Величина N_a должна быть меньше, чем $(N_{d1}+N_{d2})$, чтобы вообще поведение полупроводника определялось донорами. Однако чтобы наблюдалось поведение, показанное на рис. 4.23, должно выполняться более жесткое требование $N_a < N_{d1}$. Дело в том, что когда число компенсирующих центров превышает число самых мелких доноров, уровень Ферми при низких температурах оказывается «привязанным» к энергии *более глубокого* донора. Формально ситуация такая же, как для одного вида доноров N_{d2} с $(N_a - N_{d1})$ компенсирующими центрами.

типов центров, хорошо разделенных по энергии. «Истощение» одного из них наступает раньше, чем может начаться сколько-нибудь заметная ионизация второго.

Многовалентные центры

Мы уже говорили о том, что некоторые типы центров обладают способностью отдавать или принимать на себя более чем один электрон. Температурная зависимость примесной концентрации свободных электронов, когда поведение полупроводника определяется многовалентными донорами, имеет некоторое сходство с только что обсуждавшейся моделью (с двумя типами различных одновалентных доноров), однако существуют и интересные отличия.

На рис. 4.24 условно изображен электронный спектр двухвалентных доноров. В левом столбце показана ситуация, когда уровень Ферми расположен близко к краю зоны проводимости. В этом случае каждый донор нейтрален и удерживает два сравнительно слабо связанных электрона с *общей* энергией связи $(\varepsilon_1+\varepsilon_2)$. Фактически эти электроны могут иметь одинаковые орбиты. Однако на рис. 4.24 отмечены две разные энергии связи ε_1 и ε_2, потому что *центр* должен получить энергию ε_1, чтобы стать однократно ионизованным, тогда второй электрон удерживается на орбите с энергией связи ε_2. Большинство двухвалентных донорных центров будет находиться в состоянии однократной ионизации, если $\varepsilon_1<(\varepsilon_c-\varepsilon_F)<\varepsilon_2$, как показано во втором столбце рис. 4.24. Каждый однократно ионизованный донор предоставляет пустые электронные состояния с энергией $(\varepsilon_c-\varepsilon_1)$ и статистическим весом β_1^{-1}. (Помните, что $\beta<1$ для центра донорного типа.)

Когда уровень Ферми посредством компенсации или путем повышения температуры опускается еще ниже, каждый донор становится дважды ионизованным и β_2^{-1} электронных состояний на каждый донор теперь открываются при энергии $(\varepsilon_c-\varepsilon_2)$. Это проиллюстрировано третьим столбцом рис. 4.24. Заметим, что на верхнем уровне никаких состояний не показано, поскольку эти состояния *не существуют* до тех пор, пока на центре не появится первый электрон. Как только этот первый электрон получен центром, возникают верхние состояния и второй электрон может быть захвачен только на одно из них.

В общем случае как однократно, так и двукратно заряженные центры могут захватывать электрон и в основное, и в возбужденные состояния. Однако было бы неправильно представлять себе размещение на центре электрона с энергией $(\varepsilon_c-\varepsilon_1)$ как возбужденное состояние однозарядной конфигурации;

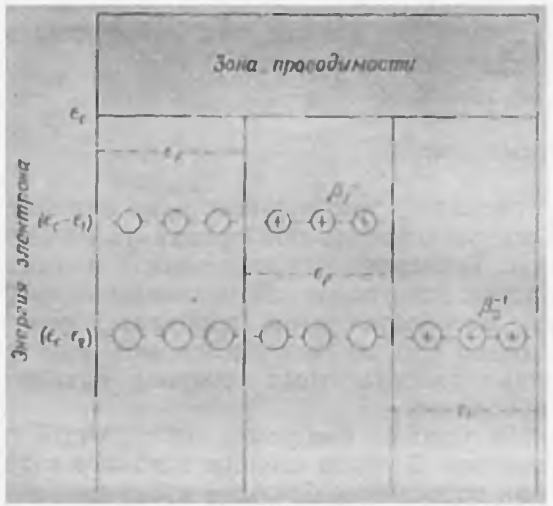


Рис. 4.24. Три возможных зарядовых состояния двухвалентных донорных центров в полупроводнике. Показано положение энергии Ферми, соответствующее каждому виду ионизации.

истинные возбужденные состояния для однократно ионизованного центра обычно бывают совершенно другими.

Уравнения для распределения электронов между зоной проводимости, многовалентными донорными центрами одного вида и разными компенсирующими центрами чрезвычайно громоздки, но обычно они могут быть разбиты на вполне разрешимые блоки, когда возможные энергетические состояния центров разделены величиной порядка нескольких k_0T. (В «Статистике полупроводников» обсуждаются некоторые решаемые ситуации). Простейшим случаем многовалентного центра является *двухвалентный донор* с тремя зарядовыми состояниями, обсуждавшимися в связи с рис. 4.24. В этом случае концентрация электронов проводимости должна быть равна $n_0 = (2N_2 + N_1 - N_a)$, где N_a — число компенсирующих центров, а $N = (N_0 + N_1 + N_2)$ — полное число доноров, из которых N_0 нейтральны, N_1 однократно ионизованы и N_2 двукратно ионизованы. Поведение полупроводника определяется состояниями при ε_1, когда $(n_0 + N_a)$ гораздо меньше N, и более глубоким уровнем, когда высокая температура или сильная компенсация, действуя одновременно или порознь, приводят к соотношению $N < (n_0 + N_a) < 2N$.

Последнее утверждение иллюстрируется кривыми рис. 4.25 для концентрации свободных дырок в двух образцах германия,

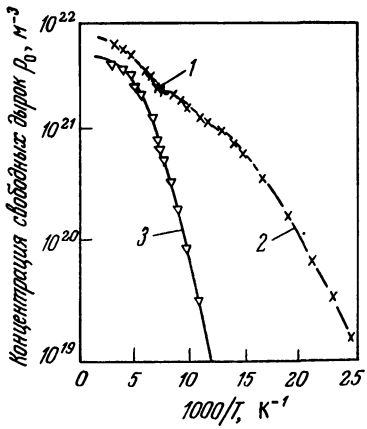


Рис. 4.25. Зависимость концентрации свободных дырок от температуры в кристаллах германия, *p*-тип проводимости которых обусловлен присутствием в них двухвалентных акцепторных атомов цинка. Результаты получены Тайлером и Вудбери [*Tyler W. W., Woodbury H. H.*— Phys. Rev., **102**, 647 (1956)]. Для образца *A* концентрация компенсирующих центров N_d меньше, чем концентрация цинка N_a, в то время как для образца *B* должно выполняться условие $N_a < H_d < 2N_a$. *1* — заселение второго уровня; *2* — образец *A*; значение ε_F привязано к первому уровню, наклон $\varepsilon_1 = 0{,}03$ эВ; *3* — образец *B*; значение ε_F привязано ко второму уровню, наклон $\varepsilon_2 = 0{,}09$ эВ.

поведение которых определяется двухвалентными *акцепторами,* которыми являются атомы цинка. Образец *A* компенсирован довольно слабо и при низких температурах ни один из акцепторных атомов цинка не является двукратно ионизованным. При более высоких температурах концентрация свободных дырок увеличивается за счет получения от каждого акцептора в этом образце второй свободной дырки.

Более сильная компенсация образца *B* препятствует *какому-либо* из акцепторных атомов цинка достигнуть при низкой температуре нейтрального состояния. Больший наклон ε_2 экспериментальной кривой для образца *B* демонстрирует присутствие смеси однократно и двукратно ионизованных акцепторов при всех температурах.

Амфотерные центры

Амфотерным является такой примесный центр, нейтральное состояние которого наиболее вероятно, когда уровень Ферми находится при некоторой энергии ε_0 *внутри* запрещенной зоны. Для меньших значений электрохимического потенциала центр, очевидно, будет отдавать один или более электронов (что соответствует донорному центру); когда энергия Ферми расположена выше ε_0, центр может переходить в состояние однократной или многократной отрицательной ионизации, т. е. ведет себя как акцептор. Такое поведение обнаруживают центры золота в германии, и спектр известных примесных состояний этого центра показан на рис. 4.26. В германии сильно выраженного *p*-типа золото может действовать как однозарядный донор, однако в германии сильно выраженного *n*-типа каждый атом золота может быть тройным акцептором. Математические

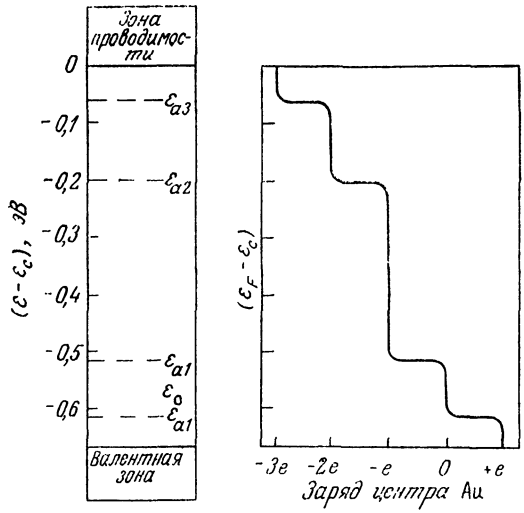


Рис. 4.26 Энергетические уровни амфотерного центра золота в германии. Правая часть рисунка показывает, как зависит заряд, приходящийся на один атом золота, от энергии Ферми при достаточно низкой температуре. Основано на сообщении: *Dunlap W. C.*— In: Progress in Semiconductors, v. 2, Wiley, 1957, p. 165.

соотношения, необходимые для описания того, сколько центров находится в каждом из зарядовых состояний, представляют собой распространение на этот случай системы уравнений, используемой для «обычных» многовалентных центров.

4.2. Явления переноса в полупроводнике

Средняя длина свободного пробега, подвижность, электропроводность

По-видимому, самым простым исследованием свойств полупроводника является измерение его электропроводности. Результат этого измерения зависит от числа подвижных носителей заряда (электронов, дырок или тех и других), от распределения их тепловых скоростей и отклонения от равновесного распределения, которое вызывается приложенным электрическим полем. Используемая теория обычно основана на приближении кинетического уравнения Больцмана (разд. 3.2).

Всегда можно с уверенностью считать, что в металле любое приложенное электрическое поле производит сдвиг распределения электронных скоростей, представляющий собой малое воз-

мущение равновесного распределения. Такое же предположение делается при рассмотрении «обычной», или «квазиравновесной», проводимости в полупроводнике, находящемся под воздействием не слишком большого поля. Однако в полупроводнике можно приложить электрическое поле, которое достаточно велико, чтобы радикально изменить распределение скоростей. О возникающей при этом проводимости (о которой опять пойдет речь ниже в этом разделе) говорят, что она определяется *горячими носителями заряда*. Эффекты разогрева полем (горячих носителей) определяют работу многих полупроводниковых приборов на основе монокристаллов и чрезвычайно важны для приложений к аморфным полупроводникам.

Рассмотрим обычную электропроводность (в слабом поле), возникающую в результате переноса носителей в полупроводнике n-типа, имеющем в равновесии n_0 подвижных электронов в единице объема. Мы будем оперировать скалярной эффективной массой m_c, которая представляет собой результат подходящего усреднения по анизотропии поверхностей постоянной энергии, лежащих непосредственно над ε_c. Тогда по аналогии с формализмом разд. 3.2 электропроводность должна иметь вид

$$\sigma = \frac{n_0 e^2 \tau_m}{m_c} = \frac{n_0 e^2 \bar{\lambda}}{m_c \bar{s}} = n_0 e \mu_n. \tag{4.47}$$

Здесь τ_m — среднее время свободного пробега между рассеивающими соударениями, должным образом усредненное по всему распределению электронов, а $\bar{\lambda} = \bar{s}\tau_m$ является соответствующей средней длиной свободного пробега для распределения электронов, обладающего средней скоростью теплового и нулевого (в квантовомеханическом смысле) движения $\bar{s}$. Величина

$$\mu_n = \frac{e\tau_m}{m_c} = \frac{e\bar{\lambda}}{m_c \bar{s}} \tag{4.48}$$

представляет собой *подвижность электронов*, и все свойства переноса электронного газа в полупроводнике наиболее удобно могут быть выражены с помощью единственной величины μ_n. Подвижность — это отношение дрейфовой скорости (в м/с) к приложенному полю (в В/м) и выражается соответственно в единицах $м^2/В \cdot с$. Эта единица совпадает с $(Тл)^{-1}$.

Перенос дырок в полупроводнике p-типа описывается аналогичной величиной — *дырочной подвижностью* μ_p. В случае *амбиполярной проводимости* собственного полупроводника или близкого к собственному в выражение для полной электропроводности полупроводника входят обе подвижности — и дырочная и электронная:

$$\sigma = e(n\mu_n + p\mu_p). \tag{4.49}$$

Характер изменения электронной подвижности с температурой зависит от следующих причин: 1) вырожден или нет электронный газ, 2) от конкретного вида связи энергии с волновым вектором для самых нижних состояний зоны проводимости полупроводника (которую мы пытаемся как-то описать, вводя некоторую «усредненную» эффективную массу в том случае, когда тензор масс фактически анизотропен) и 3) от того, какие процессы определяют рассеяние. По-видимому, эти процессы включают рассеяние акустическими фононами, оптическими фононами, нейтральными и ионизованными центрами и дислокациями. Когда два или более процессов рассеяния имеют сравнимые скорости, *необходимо* вычислять подвижность (а следовательно, и электропроводность), определяя величину

$$\tau = [(1/\tau_a) + (1/\tau_b) + (1/\tau_c) + \ldots]^{-1} \qquad (4.50)$$

как функцию энергии электрона, а затем уже выполняя операцию усреднения по всем скоростям электрона или энергиям. Только в этом случае мы найдем правильное значение τ_m. Практически установлено, что ошибка оказывается не слишком велика, если каждый процесс рассеяния рассматривать отдельно, вычисляя отвечающую ему усредненную по всем скоростям длину свободного пробега, а затем объединить эти длины согласно соотношению

$$\bar{\lambda} = [(1/\lambda_a) + (1/\lambda_b) + (1/\lambda_c) + \ldots]^{-1}. \qquad (4.51)$$

Последняя (несколько сомнительная) процедура фактически означает объединение «подвижностей» для разных процессов рассеяния, полученных в пренебрежении всеми прочими в соответствии с формулой

$$\mu_n = [(1/\mu_a) + (1/\mu_b) + (1/\mu_c) + \ldots]^{-1}. \qquad (4.52)$$

Только в отдельных случаях и при очень тщательных расчетах можно обнаружить, что процессы рассеяния объединяются в соответствии с (4.50), а не (4.52) *.

Рис. 4.27—4.29 показывают, как подвижность дырок (рис. 4.27) или электронов (рис. 4.28 и 4.29) изменяется с температурой для трех типичных полупроводниковых монокристаллов. Форма кривых определяется конкуренцией нескольких процессов рассеяния, причем эта конкуренция сильно зависит от температуры. Мы еще вернемся к этим рисункам, когда будем обсуждать различные механизмы рассеяния.

---

* Это утверждение автора сомнительно. Ошибка в вычислении результирующей подвижности не по обратным временам релаксации, а по обратным подвижностям особенно заметна в случае компенсированных полупроводников. — *Прим. ред.*

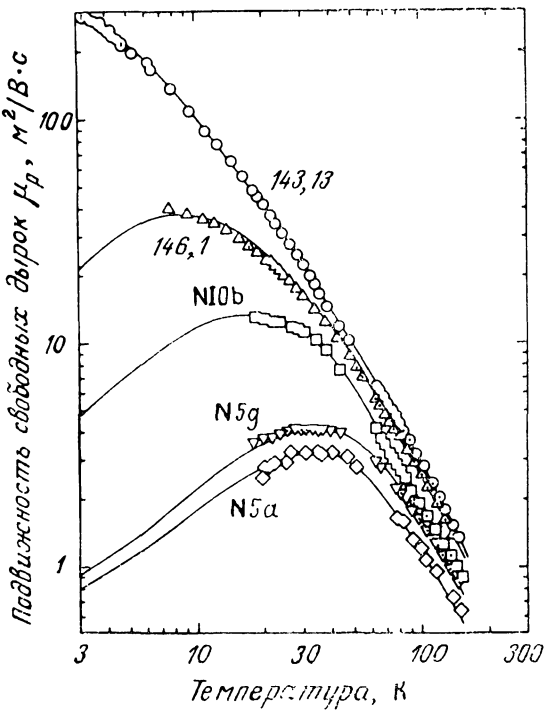


Рис. 4.27. Зависимость подвижности свободных дырок от температуры в кристаллах германия, легированных акцепторами меди (и в гораздо меньшей концентрации компенсирующими донорами). Концентрация меди лежит в пределах от $\sim 10^{19}$ м$^{-3}$ для образца 143,13 до $\sim 10^{22}$ м$^{-3}$ для образца *N5a* [*Norton P., Levinstein H.*— Phys. Rev, **B6**, 470 (1972)].

Вырожденное состояние возникает, когда n_0 имеет большую величину, а температура довольно низка. Можно ожидать, что в этих условиях процесс проводимости сходен с проводимостью в металле, где μ_n имеет ту же температурную зависимость, что и электропроводность, кривая для которой показана на рис. 3.1. При этом аналогия между металлом и вырожденным полупроводником несколько преувеличена, поскольку поведение электронов в полупроводнике в гораздо большей степени зависит от статических нарушений (дефектов).

Для вырожденного электронного распределения в полупроводнике усреднение длин и времен свободного пробега производится для электронов, энергии которых расположены в пределах 1-2 k_0T от энергии Ферми. Скорость электрона, подставляемая в знаменатель выражений (4.47) и (4.48), в этом случае есть фермиевская скорость $s_F=[2(\varepsilon_F-\varepsilon_c)/m_c)]^{1/2}$. Можно

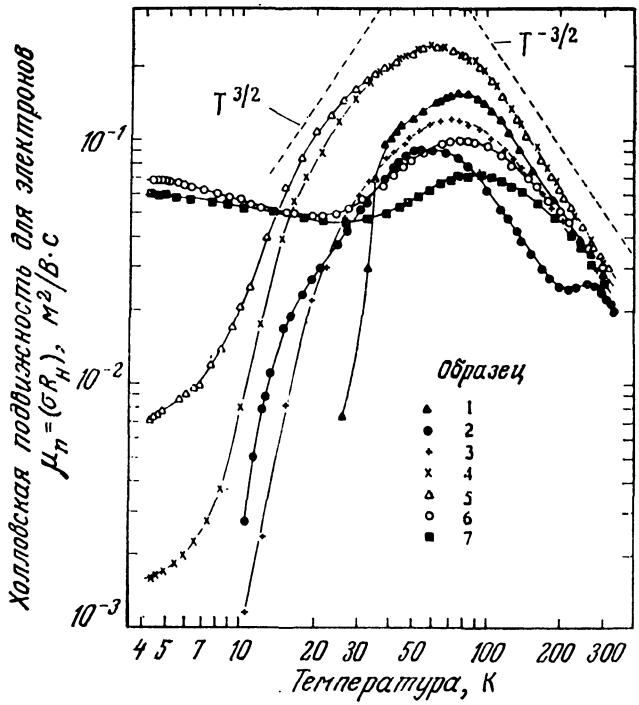


Рис. 4.28. Температурная зависимость холловской подвижности электронов в монокристаллах Mg_2Ge n-типа [*Li P. W., Lee S. N., Danielson G. C.*— Phys. Rev., **B6**, 442 (1972)]. Штриховые линии условно изображают зависимости, ожидаемые для решеточного рассеяния ($T^{-3/2}$) и для рассеяния на ионизованных примесях ($T^{3/2}$). Для нескольких образцов «холловская подвижность» аномально падает при низких температурах из-за проводимости по примесям.

вычислить, как величина μ_n *должна* изменяться с температурой для такого вырожденного распределения, при котором преобладает рассеяние на фононах, ионизованных центрах и т. д.[16]. Однако практически подвижность в полупроводниковом образце с большим числом свободных электронов обычно слабо изменяется в широком интервале температур. Примером этого являются две почти совершенно плоские зависимости для образцов InSb XS22 XS25 на рис. 4.29.

---

[16] Рассеяние электронов для вырожденных электронных распределений обсуждается, например, в книгах Смита и Путли, указанных в библиографии в конце этой главы. [См. также: *Фистуль В. И.* Сильно легированные полупроводники.— М.: Наука, 1967. — *Прим. ред.*]

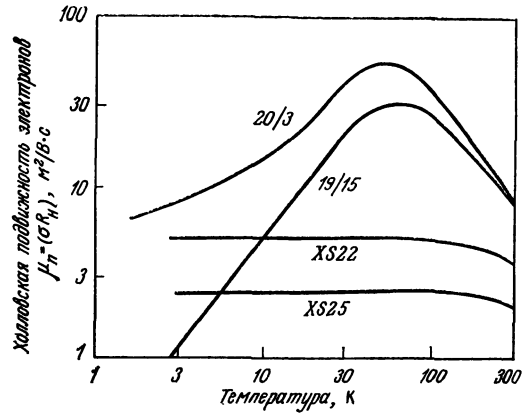


Рис. 4.29. Подвижность электронов в образцах InSb n-типа, полученная из измерений электропроводности и эффекта Холла [*Putley E. H.*— Proc. Phys. Soc., 73, 280 (1959)]. В образцах XS22 и XS25 с $n_0 \sim 10^{22}$ м$^{-3}$ электронный газ вырожден. Для образца 20/3 $n_0 \sim 3 \cdot 10^{20}$ м$^{-3}$, и полная концентрация ионизованных центров, которые могут принимать участие в рассеянии, не намного превышает это значение. Образец 19/15 обладает примерно такой же концентрацией свободных электронов, однако он сильнее компенсирован и поэтому содержит больше ионизованных центров, уменьшающих подвижность при низких температурах.

Гораздо более интересное многообразие вида температурных кривых подвижности наблюдается, когда электроны в полупроводнике имеют невырожденное (больцмановское) распределение. Когда концентрация электронов достаточно мала, а температура достаточно велика, чтобы обеспечить справедливость невырожденного приближения, скорость $\bar{s}$, входящая в уравнение (4.47) и (4.48) равна

$$\bar{s} = (8k_0T/\pi m_c)^{1/2}. \tag{4.53}$$

Для такого случая полного отсутствия вырождения предположим, что зависимость длины свободного пробега от температуры и от энергии электрона (т. е. от квадрата его скорости) может быть аппроксимирована выражением

$$\bar{\lambda} = AT^p(\varepsilon - \varepsilon_c)^q. \tag{4.54}$$

Числа p и q зависят от того, на чем (фононах, примесных центрах и т. д.) фактически рассеивается бо́льшая часть электронов. Тогда при использовании этой длины свободного пробега

в кинетическом уравнении Больцмана она усредняется по больцмановскому распределению в духе равенства (3.33):

$$\bar{\lambda} = \frac{\int_{\varepsilon_c}^{\infty} \lambda (\varepsilon - \varepsilon_c) \exp(-\varepsilon/k_0T)\, d\varepsilon}{\int_{\varepsilon_c}^{\infty} (\varepsilon - \varepsilon_c) \exp(-\varepsilon/k_0T)\, d\varepsilon}. \tag{4.55}$$

Следовательно,

$$\bar{\lambda} = AT^p (k_0T)^{-2} \int_{\varepsilon_c}^{\infty} (\varepsilon - \varepsilon_c)^{q+1} \exp\left(\frac{\varepsilon_c - \varepsilon}{k_0T}\right) d\varepsilon =$$

$$= A\Gamma(q+2)\, T^p (k_0T)^q, \tag{4.56}$$

где температурная зависимость содержит характерным образом обе величины p и q. Если подставить температурные зависимости (4.53) и (4.56) в выражение (4.48), подвижность должна принять вид

$$\mu_n = BT^{p+q-1/2}, \tag{4.57}$$

где постоянная B зависит от выбора соответствующей микроскопической модели электронного рассеяния.

Рассеяние на колебаниях решетки

В достаточно чистом полупроводниковом кристалле, в котором преобладают ковалентные связи, электроны преимущественно рассеиваются продольными акустическими (LA) фононами. LA-моды производят в кристалле последовательность сжатий и растяжений, создающую локальную модуляцию диэлектрической проницаемости, а также энергий ε_c и ε_v, отвечающих верхней и нижней границам запрещенной зоны. Чтобы описать, в какой мере минимум зоны проводимости «чувствует» растяжение, используют *константу потенциала деформации*[17]

$$\varepsilon_1 = -V(\partial\varepsilon_c/\partial V), \tag{4.58}$$

которая составлена так, что имеет размерность энергии. Если величина ε_1 отлична от нуля, то любые LA-фононы с соответствующим волновым вектором будут создавать подъемы и впа-

[17] Приближение «потенциала деформации» для рассеяния электронов на волнах сжатия (продольных фононах) в полупроводнике было предложено Шокли и Бардином [*Shockley W., Bardeen J.*— Phys., Rev., **80**, 72 (1950); это приближение детально обсуждается в книге Шокли (см. библиографию). Интересное и простое описание квантовомеханического представления процесса рассеяния содержится в разд. 9.1 книги Мак-Келви (см. библиографию).

дины минимальной энергии зоны проводимости, каждый из которых может приводить к рассеянию электрона. Когда электрон встречает ряд сжатий и разрежений в таком направлении, что угол между направлением движения электрона и нормалью к плоскостям сжатия составляет ϕ, то наиболее вероятный угол рассеяния составляет $(\pi - 2\phi)$.

LA-фононы, которые могут рассеять электрон, должны обладать длиной волны, во всяком случае, не меньшей, чем длина волны электрона. Простой расчет показывает, что рассеяние электрона с энергией k_0T LA-фононом приводит к значительному изменению направления импульса и к очень малому изменению энергии электрона. (Это подтверждается в задаче 4.8). Такие относительно длинноволновые и низкоэнергетичные фононы сохраняются в твердом теле в большом количестве и при уменьшении температуры значительно ниже температуры Дебая. При очень низких температурах (представляющих меньший интерес для решеточного рассеяния, поскольку на него налагаются другие эффекты при малых T), рассеяние LA-фононами становится анизотропным, с предпочтительным рассеянием в прямом направлении, однако, как было показано Шокли и Бардином [17], в большей части температурного диапазона LA-фононы рассеивают электроны изотропно. Эти авторы показали, что длина свободного пробега между соударениями может быть с помощью потенциала деформации представлена выражением

$$\lambda = \frac{4\rho v_{LA}^2}{k_0 T}\left(\frac{\hbar^2}{2m_c\varepsilon_1}\right)^2 \tag{4.59}$$

для довольно идеализированного случая зоны проводимости с изотропной и не зависящей от энергии эффективной массой. В этом выражении v_{LA} обозначает скорость длинноволновых LA-фононов, и ρ — плотность твердого тела.

Сравнение выражений (4.54) и (4.59) показывает, что рассеяние на LA-фононах соответствует значениям $p = -1$ и $q = 0$ в выражении (4.54). Соответственно для *решеточной подвижности*, согласно (4.57), следует ожидать зависимости

$$\mu_L = BT^{-3/2}. \tag{4.60}$$

Для этого простого вида рассеяния величина B может быть легко рассчитана с помощью величин, входящих в выражения (4.53)—(4.59). Пример такого расчета предлагается читателю в задаче 4.9.

Отрицательный показатель степени в температурной зависимости подвижности наблюдается в наиболее чистых образцах при высоких температурах на рис. 4.27—4.29. Зависимость близка к $T^{-3/2}$ для электронов в Mg_2Ge и InSb, но более близка к $T^{-1,9}$ для дырок в образце германия 143,13. Отклонения от

точной зависимости $T^{-3/2}$ для решеточной подвижности в большей или меньшей степени провляются почти во всех полупроводниках. Причиной этого является множество конкурирующих эффектов, включающих несферическую форму поверхностей постоянной энергии для электронов проводимости в таких полупроводниках, как кремний или германий, зависимость эффективной массы от энергии в таких полупроводниках, как InSb, рассеяние, обусловленное поперечными акустическими колебаниями (*TA*-фононами), и рассеяние более высокоэнергетичных электронов при высоких температурах фононами оптических ветвей. Этот последний вид фононов вне зависимости от того, насколько сильно изменилось направление движения электрона, способен, конечно, сильно изменить энергию электронов.

Рассеяние на ионизованных примесных центрах

На кривых рис. 4.27—4.29 можно заметить, что подвижность свободных носителей в невырожденном образце при охлаждении не возрастает неограниченно. Действительно, для ряда приведенных образцов подвижность при более низких температурах опять начинает уменьшаться. Этого можно ожидать в кристалле, когда число фононов в нем сильно уменьшено охлаждением и средние длина и время свободного пробега определяются *рассеянием на ионизованных примесных центрах.*

Каждый ионизованный центр в кристалле представляет собой неподвижный положительный или отрицательный заряд, который может отклонить траекторию пролетающего электрона, как это показано на рис. 4.30. Результирующая орбита как в случае притяжения, так и в случае отталкивания будет гиперболической, и необходимый математический вывод идентичен тому, который был использован Резерфордом в его модели рассеяния α-частиц атомными ядрами [18]. Только в нашем случае выражение Резерфорда для угла рассеяния должно быть пересчитано с учетом диэлектрической проницаемости среды $\varkappa$.

Средняя длина свободного пробега, связанная с резерфордовским рассеянием, не зависит явно от температуры системы, однако она пропорциональна квадрату энергии электрона. В соответствии с этим результирующая подвижность μ_i, согласно (4.57), должна изменяться как $T^{3/2}$. В основном этот вывод является правильным, однако имеется одна осложняющая особенность, которую необходимо рассмотреть.

---

[18] Вывод формулы рассеяния Резерфорда, впервые приведенный в работе *Rutherford E.*— Phil. Mag., **21**, 669 (1911), имеется в любом учебнике современной атомной физики.

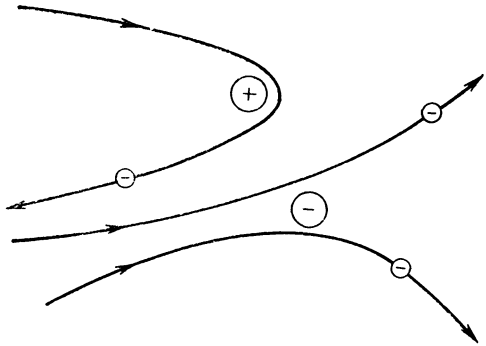


Рис. 4.30. «Резерфордовское рассеяние» электронов положительно и отрицательно заряженными центрами.

Максимальный угол рассеяния при взаимодействии электрона с ионизованным центром составляет 180°, что соответствует лобовому столкновению электрона с заряженным центром. Конуэлл и Вайскопф [19] предположили, что нижний угловой предел интеграла Резерфорда для полного сечения рассеяния должен быть равен

$$\phi_{\text{мин}} = (Ze^2 N_i^{1/3}/\pi\varkappa\varepsilon_0 m_c s^2) \tag{4.61}$$

для электрона со скоростью s. Здесь N_i — концентрация заряженных центров, каждый из которых имеет эффективный заряд $\pm Ze$. Соотношение (4.61) отражает тот факт, что электрон не может быть удален от ионизованного центра на расстояние, большее $(8N_i)^{-1/3}$, в соответствии с чем эффективная сила кулоновского взаимодействия между электроном и центром полагается на этом расстоянии равной нулю [20].

В виде, который приняли Конуэлл и Вайскопф, резерфордовское рассеяние приводит к следующему выражению для электронной подвижности:

$$\mu_I = \frac{(64N_i)\,(\varkappa\varepsilon_0/Z)^2\,(\pi/m_c)^{1/2}\,(2k_0T/e^2)^{3/2}}{\ln\left[1+\left(12\pi\varkappa\varepsilon_0 k_0 T/Ze^2 N_i^{1/3}\right)^2\right]}, \tag{4.62}$$

---

[19] *Conwell E. M., Weisskopf V. F.*— Phys. Rev., 77, 388 (1950).

[20] Предложенная Конуэлл и Вайскопфом процедура исключения малоуглового рассеяния, соответствующего траекториям электрона с большим *прицельным расстоянием* относительно заряженного центра, не является единственно возможной. Альтернативный метод, основанный на экранировании свободными носителями кулоновского поля ионизованной примеси, был развит в 50-х годах Бруксом и Херрингом, а также Динглом. Обзор различных моделей рассеяния на ионизованной примеси содержится в работе: *Chattopadhyay D*, Queisser H. J.— Rev. Mod., Phys., 53, 745 (1981).

турах, по-видимому, оказывает значительное воздействие *междолинное рассеяние*, как было показано Херрингом[21].

Другой важный механизм рассеяния для электронов и дырок особенно при достаточно высокой температуре в кристалле со значительной долей ионной связи обусловлен фононами *оптических мод*. В то время как фононы акустических мод обычно обладают энергией, очень малой по сравнению с энергией электронов, которые ими рассеиваются (что выясняется в задаче 4.8), фононы оптических мод обладают большой энергией $\hbar\omega_{\text{опт}}$ независимо от величины их волнового вектора. Рассеяние на оптических фононах (которые часто называют фононами *поляризационной моды*) особенно существенно для твердого тела с полярной (частично или полностью ионной) решеткой, поскольку оптический фонон создает в таком твердом теле электрическое поле дипольного типа.

Можно ожидать, что рассеяние поляризационной модой окажется несущественным при низких температурах, однако оно может в некоторых материалах стать преобладающим, когда величина k_0T становится сравнимой с энергией образования оптических фононов $\hbar\omega_{\text{опт}}$.

Петриц и Склэнлон[22] обсуждают модель такого рассеяния, для которого подвижность изменяется с температурой как $\exp(\hbar\omega_{\text{опт}}/k_0T)$.

Электрон в полярном кристалле создает вокруг себя определенное пространственное распределение поляризации решетки; каждый ион оказывается несколько смещенным из-за присутствия вблизи него электрона. Такое *поляризационное облако*, окружающее каждый электрон, обладает запасом электростатической энергии. Когда электрон движется по кристаллу, его поляризационное облако должно следовать за ним. Движение электрона вместе с создаваемым им поляризационным облаком по полярной решетке называют движением *полярона*[23]. В свое время Ландау (1933) и Пекар (1946) высказывали предположение, что электрон может самозахватиться той деформационной ямой, которую он сам же создает в ионном кристалле, так что полярон окажется совершенно неподвижным. Столь радикальная точка зрения о поляроне теперь уже не выдвигается, однако действительно следует ожидать, что когда на энергети-

---

[21] *Herring C.*— J. Bell System Tech., **34**, 237 (1955).

[22] *Petritz R. L., Scanlon W. W.*— Phys. Rev., **97**, 1629 (1955). В этой работе рассмотрено рассеяние поляризационной модой для сульфида свинца. В ней приведены ссылки на несколько более ранних моделей рассеяния на оптических фононах.

[23] Обширная литература, касающаяся поляронов в ионных кристаллах, содержится в лекциях Polarons and Exitons, eds. C. G. Kuper, G. D. Whitfield, Plenum, 1963.

ческие состояния электрона существенное воздействие оказывает образование полярона, эффективная масса должна быть большой (а подвижность соответственно малой).

Как в ионных, так и в ковалентных кристаллах электроны должны рассеиваться также и нейтральными примесными центрами. Простая модель такого рассеяния была предложена Эргинсоем[24]. Подгоняя к этому случаю модель рассеяния медленных электронов нейтральными атомами газа, Эргинсой отметил, что для электрона с эффективной массой m_c и скоростью s атом, в котором самый удаленный от ядра электрон обладает волновой функцией с характерным радиусом a_0, представляет собой мишень с поперечным сечением

$$\sigma_e \approx (20\hbar a_0/sm_c). \tag{4.64}$$

[Следует также предположить, что электрон обладает кинетической энергией $(m_c s^2/2)$, значительно меньшей, чем первый потенциал ионизации атома.] Тогда для N_n нейтральных рассеивающих центров в единице объема время релаксации, обусловленное столкновениями, есть

$$\tau_m = (s\sigma_e N_n)^{-1} = (m_c/20\hbar a_0 N_n), \tag{4.65}$$

в соответствии с чем подвижность равна

$$\mu_N = (e/20\hbar a_0 N_n). \tag{4.66}$$

Эта подвижность, обусловленная нейтральными примесными центрами, не зависит от температуры, а эффективность рассеяния не зависит от распределения электронных скоростей. Практически рассеяние на нейтральных центрах особенно эффективно при низких температурах для самых быстрых электронов из теплового распределения, тогда как процесс рассеяния на ионизованных центрах наиболее вероятен для медленных электронов. В 1972 г. Нортон и Левинштейн, объясняя подвижность дырок в германии, легированном медью (см. рис. 4.27), предположили, что «глубокие» акцепторы, создаваемые атомами меди, в качестве нейтральных рассеивающих центров менее эффективны, чем следовало бы ожидать из простой модели Эргинсоя.

В дополнение ко всем видам рассеяния, которые имеют место в тех случаях, когда объекты атомного размера нарушают периодичность решетки, электроны могут рассеиваться объектами, протяженными в одном или нескольких измерениях. Линия дислокации, пересекающая кристалл, не слишком эффективно влияет на рассеяние, если она электрически нейтральна, однако она может быть очень эффективным рассеивателем, если

[24] *Erginsoy C.*— Phys. Rev., 79, 1013 (1950).

состоит из цепочки отрицательно заряженных акцепторов[25]. Границы зерен в поликристаллическом образце, несомненно, оказывают воздействие на подвижность носителей и могут воспрепятствовать перемещению электрона из одного микрокристалла в другой, если на поверхности раздела имеется потенциальный барьер. Поэтому свойства полупроводника в монокристаллической форме могут быть описаны с гораздо большей надежностью, чем свойства того же материала в виде поликристаллического слитка, поликристаллической пленки или прессованного порошка.

Эффект Холла и магнетосопротивление

Предположим, что электроны в зоне проводимости полупроводника движутся, находясь под одновременным воздействием электрического поля, магнитного поля и какого-то достаточно эффективного механизма рассеяния. Если зона проводимости может быть охарактеризована изотропной скалярной эффективной массой, то с помощью кинетического уравнения Больцмана можно прийти к соотношению

$$\mathbf{E} = \sigma^{-1}\{\mathbf{J} - [\sigma R_H]\,\mathbf{J} \times \mathbf{B}\}, \tag{4.67}$$

которое уже фигурировало раньше под номером (3.37) в классической модели металла Лоренца. В соответствии с этой моделью постоянная Холла (которая определяет величину отклика в направлении, перпендикулярном как току, так и полю) равна

$$R_H = -\frac{\langle\tau^2\rangle}{\langle\tau\rangle^2 n_0 e} = \frac{-r}{n_0 e}. \tag{4.68}$$

Безразмерная величина r называется *холл-фактором*. Она зависит от комбинации процессов рассеяния, эффективных в заданных условиях, и от того, как меняется среднее время между столкновениями с энергией электрона. Как правило холл-фактор по величине близок к единице. Так, $r=1$, если все электроны движутся с одинаковой скоростью, как это предполагается в классической модели Друде, или для вырожденного электронного газа (в котором все электроны движутся с одинаковой скоростью Ферми). Мы находим $r=(3\pi/8)=1{,}18$, если электронный газ невырожден, а рассеяние осуществляется только LA-фононами. Когда существенно рассеяние на заряженных центрах, r для изотропной зоны может стать больше, чем 1,9; с другой стороны, r может уменьшиться до 0,7, если поверхности постоянной энергии электрона существенно откло-

[25] Модель рассеяния электронов *электрически заряженной* дислокацией была предложена Ридом [*Read W. T.*— Phil. Mag., **45**, 775 (1954)].

няются от сферической формы. В книгах Смита и Патли, цитированных в библиографии, обсуждаются некоторые критерии для определения r.

Из выражений (4.47) и (4.68) следует, что электропроводность и постоянная Холла связаны следующим, образом: $(-\sigma R_H) = r\mu_n \equiv \mu_H$. Величину μ_H часто называют *холловской подвижностью*. Подвижность, показанная на рис. 4.28 и 4.29, представляет собой холловскую подвижность $\mu_H = (-\sigma R_H)$, а не *подвижность электропроводности*, или *дрейфовую подвижность* $\mu_n = (\mu_H/r)$. Для некоторых полупроводников оказалось возможным измерить дрейфовую подвижность независимым методом, однако часто приходится полагаться исключительно на данные по холловской подвижности.

Безразмерную величину $(B\mu_H)$ иногда называют *углом Холла,* поскольку тангенс этого угла есть отношение поперечной составляющей электрического поля к продольной в плоскости, перпендикулярной магнитному полю. (Это обсуждается в задаче 4.10.)

Для электронов в полупроводнике, так же как и для электронов в классической модели металла Лоренца, электропроводность в магнитном поле *должна* уменьшаться в соответствии с формулой (3.39). Следовательно, полупроводник имеет максимальное магнетосопротивление (пропорциональное B^2), когда **B** и **J** взаимно перпендикулярны, и его магнетосопротивление обращается в нуль, когда **B** и **J** параллельны. На практике ситуация редко бывает столь простой. Для многих полупроводников все же существует значительное магнетосопротивление в продольном направлении, обусловленное анизотропией тензора эффективной массы и анизотропией времени релаксации[26]. Одновременное действие нескольких конкурирующих механизмов рассеяния может привести как к более быстрому, так и более медленному возрастанию сопротивления, чем точная зависимость B^2. И, конечно, для больших магнитных полей и слабого рассеяния существует возможность наблюдения осцилля-

[26] В теории магнетосопротивления имеется ряд интересных особенностей, присущих зоне, для которой энергия вблизи минимума описывается в k-пространстве набором эллипсоидальных изоэнергетических поверхностей (подобных тем, которые изображены на рис. 3.50). Соответствующая теория обсуждается в книгах Смита и Патли, приведенных в библиографии. Ранние измерения «аномального» магнетосопротивления в германии [*Pearson G. L., Suhl H.*— Phys. Rev., **83**, 768, (1951)] и в кремнии [*Pearson G. L., Herring C.*— Physica, **20**, 975, (1954)] и расчеты зонной структуры с помощью метода ОПВ [*Herman.*— Phys. Rev., **93**, 1214 (1954)] прекрасно совпали между собой в предсказании ряда особенностей зонной структуры, которые были подтверждены последующими измерениями. Обзор поведения магнетосопротивления в германии и кремнии выполнен Гликсманом [*Glicksman M.* Progress in Semiconductors, v. 3, Wiley, 1958].

ционного эффекта Шубникова — де Гааза, как показано на рис. 3.66. Этим осложнениям следует скорее радоваться, чем огорчаться, так как они позволяют детально разобраться в том, как в зависимости от их энергии и волнового вектора электроны распределяются по зонам в твердом теле.

Явления переноса с участием нескольких зон и в случае собственной проводимости

Электропроводность собственного полупроводника представляет собой сумму электропроводностей по двум зонам:

$$\sigma_i = en_i(\mu_n + \mu_p). \tag{4.69}$$

В полупроводнике, который не является полностью собственным, но для которого обе зоны дают вклад в проводимость, это выражение должно быть записано в более общем виде

$$\sigma = e(n_0\mu_n + p_0\mu_p), \tag{4.70}$$

где n_0 и p_0 связаны требованием, чтобы их произведение составляло n_i^2 (для невырожденной статистики), а их разность равнялась разности концентраций ионизованных доноров и ионизованных акцепторов. Для полупроводника n-типа, как мы уже видели в (4.24), эти требования приводят к соотношению

$$n_0 = (p_0 + N_{di}) = \{(N_{di}/2) + [(N_{di}/2)^2 + n_i^2]^{1/2}\}, \tag{4.71}$$

где N_{di} обозначает избыток ионизованных доноров над ионизованными акцепторами. Для полупроводника p-типа можно аналогично записать

$$p_0 = (n_0 + N_{ai}) = \{(N_{ai}/2) + [(N_{ai}/2)^2 + n_i^2]^{1/2}\}. \tag{4.72}$$

Последние выражения пригодны для анализа кривых электропроводности, подобных изображенным на рис. 4.31. Три образца InSb из изображенных на этом рисунке по мере охлаждения и удаления от области собственной проводимости приобретают проводимость n-типа, остальные четыре образца приобретают проводимость p-типа. Заметим, что, поскольку отношение $(\mu_n/\mu_p) \equiv b$ для данного полупроводника значительно больше единицы, величина электропроводности, следующая из выражений (4.70) и (4.72), в промежуточной области температур для образцов p-типа *меньше*, чем для образца собственной проводимости при той же температуре.

В случае носителей заряда обоих типов эффект Холла также усложняется, поскольку сила Лоренца вынуждает и электроны, и дырки смещаться в одном и том же направлении. (Если электрическое поле заставляет электроны двигаться слева направо, а магнитное поле отклоняет их по часовой стрелке, то дырки

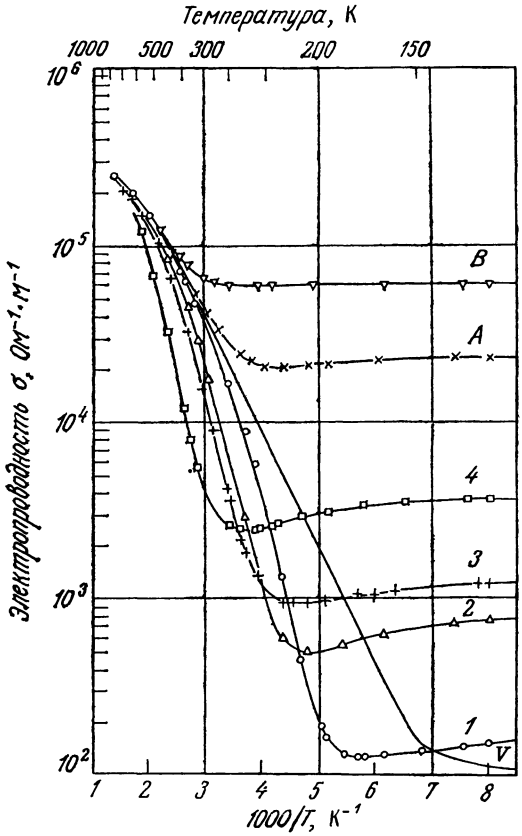


Рис. 4.31. Электропроводность, построенная в логарифмическом масштабе, как функция обратной температуры для образцов антимонида индия. Образцы *V*, *A* и *B* *n*-типа. Образцы *1*, *2*, *3* и *4* *p*-типа при низких температурах (*Madelung O.* Physics of III — V Compounds, Wiley, 1964. [Имеется перевод: *Маделунг О.* Физика полупроводниковых соединений элементов III и V групп.— М.: Мир, 1967.]

должны в том же электрическом поле дрейфовать справа налево и отклоняться магнитным полем против часовой стрелки.) Поэтому вклады электронов и дырок в эффект Холла стремятся скомпенсировать друг друга (см. задачу 4.11), так что

$$R_H = \frac{r\left(p_0\mu_p^2 - n_0\mu_n^2\right)}{e\left(p_0\mu_p - n_0\mu_n\right)^2}\,. \tag{4.73}$$

Примеры различного вида зависимостей, определяемых этим соотношением, можно видеть на рис. 4.32, где показаны результаты для тех же образцов, что и на рис. 4.31. Можно видеть, что

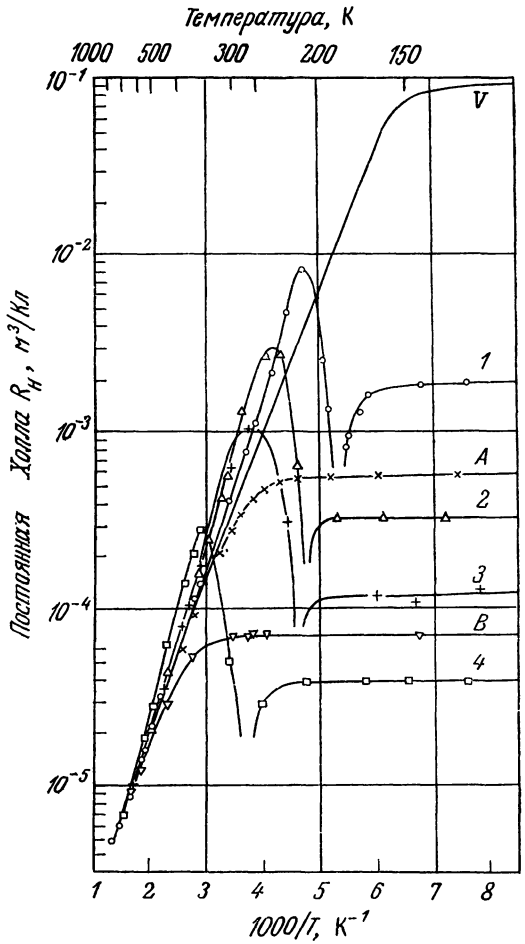


Рис. 4.32. Постоянная Холла как функция обратной температуры для семи образцов InSb, для которых температурная зависимость электропроводности представлена на рис. 4.31 [*Madelung O.* Physics of III—V Compounds, Wiley, 1964].

образцы p-типа имеют положительный коэффициент Холла при низких температурах и отрицательный при температурах выше той, при которой $n_0\mu_n{}^2 = p_0\mu_p{}^2$. Более того, отрицательная ветвь коэффициента Холла для образцов, которые первоначально обладали p-типом проводимости, поднимается выше кривой собственной проводимости, подобно тому, как электропроводность в тех же образцах падала ниже собственной электропроводности.

Читатель может много узнать о процессе проводимости в полупроводнике и о взаимодействии явлений, определяемых электронами и дырками, путем анализа кривых, подобных приведенным; скромный шаг в этом направлении предложен в задаче 4.12.

Примесная проводимость

До сих пор, говоря о примесных центрах, мы всегда характеризовали их локализованными волновыми функциями для связанных состояний, обладающих конечной энергией связи. Исходя из этого, можно было бы ожидать, что охлаждение примесного полупроводника до абсолютного нуля приведет к захвату центрами *каждого* подвижного электрона (или дырки), т. е. к бесконечному удельному сопротивлению. Такая картина верна, *если* среднее расстояние между центрами очень велико по сравнению с радиусом a_d волновой функции основного состояния центра. Однако полупроводник при низких температурах ведет себя иначе, а именно так, как если бы у электронов существовала возможность переходить с одного центра на другой, не поднимаясь в зону проводимости. Такой перенос известен как *примесная проводимость* *.

Когда примесная проводимость осуществляется между донорами в полупроводнике n-типа, полная электропроводность приближенно может быть записана в виде

$$\sigma = n_0 e\mu_n + \sigma_2 \exp(-\varepsilon_2/k_0T) + \sigma_3 \exp(-\varepsilon_3/k_0T), \qquad (4.74)$$

где второй и третий члены относятся к двум различным видам переноса электронов между донорами. Эти процессы примесной проводимости влияют на величину ЭДС Холла, измеряемой в направлении, перпендикулярном магнитному полю. Следовательно, «холловская подвижность» σR_H в том случае, когда примесная проводимость дает вклад в полный перенос электронов, не соответствует истинной подвижности μ_n. Это приводит к тому, что кривые холловской подвижности для некоторых образцов на рис. 4.28 при охлаждении резко падают вниз в той области температур, где имеет место примесная проводимость.

Прежде чем обсуждать два механизма примесной проводимости, рассмотрим влияние расстояния между донорами на

* В отечественной литературе обычно говорят о «проводимости по примесной зоне», поскольку термином «примесная проводимость» принято называть обычную проводимость полупроводника, в котором присутствие свободных носителей заряда в проводящих зонах обусловлено примесными центрами. Подробному обсуждению вопросов, излагаемых в данном разделе, посвящена монография: *Шкловский Б. И., Эфос А. А.* Электронные свойства легированных полупроводников.— М: Наука, 1979.— *Прим. перев.*

энергию ионизации донора ε_d. Существуют по крайней мере три причины, по которым сильное легирование должно приводить к уменьшению ε_d: результирующая потенциальная энергия, которой обладает электрон по отношению к породившему его донору, если этот электрон не может уйти от донора достаточно далеко, не встречая при этом другие доноры; экранирование, производимое электронным газом, которое всегда сопутствует высокой концентрации ионизованных центров, и поляризация нейтральных центров. В результате величина ε_d уменьшается по мере того, как доноры оказываются во все большей близости друг к другу, и обращается в нуль для концентраций, превышающих некоторую критическую величину $N_{\text{крит}}$, которая характерна для данного вида доноров и данного конкретного полупроводника. Когда концентрация доноров достигает $N_{\text{крит}}$, то обычно оказывается[27], что *среднее* расстояние между соседними донорами составляет от $5a_d$ до $6a_d$; при этом имеется значительная вероятность того, что в окрестности данного донора на расстоянии в пределах от $3a_d$ до $4a_d$ окажется еще один или два донора. В задачах 4.13 и 4.14 рассматриваются расстояние между донорами в полупроводнике n-типа и использование этой величины для определения $N_{\text{крит}}$.

На рис. 4.33 показано изменение энергии ионизации донорных атомов мышьяка в германии в зависимости от концентрации ионизованных доноров N_{di}. Можно видеть, что экспериментальная кривая хорошо описывается зависимостью вида

$$\varepsilon_d = \varepsilon_{d0}\,[1-(N_{di}/N_{\text{крит}})^{1/3}], \qquad (4.75)$$

по которой уменьшение энергии ионизации обратно пропорционально расстоянию между *ионизованными* центрами.

При концентрации центров, превосходящей $N_{\text{крит}}$, электроны, связанные с примесными центрами, уже не требуют термической активации и полупроводник становится *примесным металлом*. В таком сильно легированном полупроводнике концентрация свободных носителей и, по-видимому, электропроводность совершенно нечувствительны к изменениям темпера-

[27] «Среднее расстояние между донорами» и «среднее расстояние до ближайшего соседа» совпадали бы только в том случае, если бы доноры были размещены по правильной сверхрешетке; тогда каждый донор имел бы 12 соседей на одном и том же расстоянии $(2/N_d^2)^{1/3}$. Поскольку примесь в реальном кристалле распределена хаотически, каждый донор взаимодействует главным образом с одним, двумя или тремя другими донорами, обладающими наибольшей вероятностью оказаться в пределах расстояния $(^3/_4\pi N_d)^{1/3}$. Мотт и Хаббард нашли, что расстояние, на котором с наибольшей вероятностью происходит «коллапс» ε_d и возникает квазиметаллическое состояние, соответствует условию $a_d N_{\text{крит}}^{1/3} \approx 0{,}25$. Литература по этому вопросу обобщена Моттом [*Mott N. F.* Metal-Insulator Transitions, Taylor and Francis, 1974].

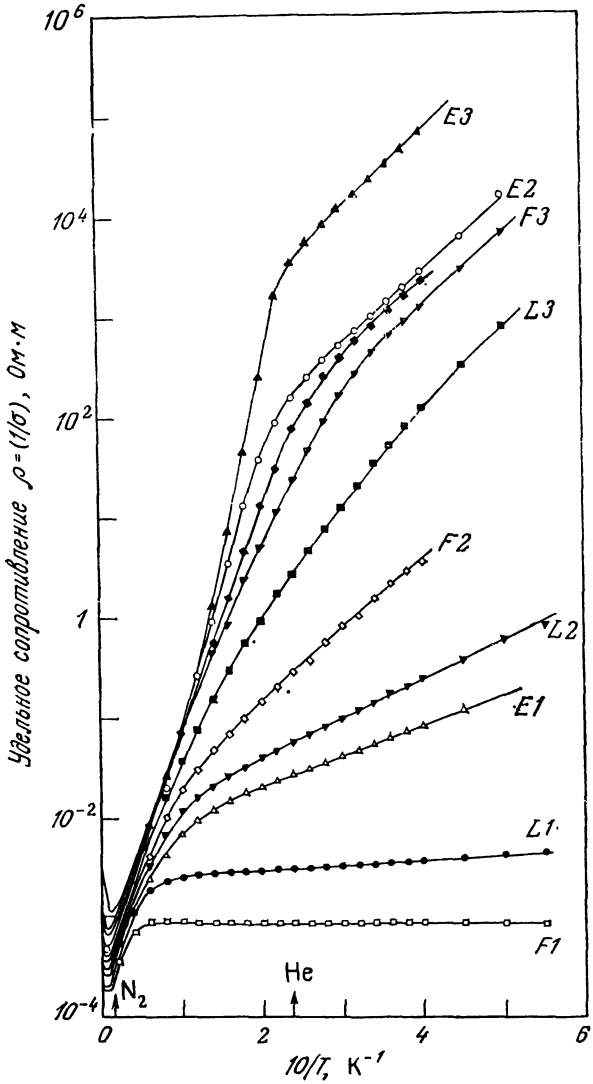


Рис. 4.34. Температурная зависимость удельного электрического сопротивления образцов германия, легированных главным образом сурьмой в концентрациях, изменяющихся в пределах от $\sim 7 \cdot 10^{21}$ м$^{-3}$ для образца *E3* до $\sim 2 \cdot 10^{23}$ м$^{-3}$ для образца *F1* *Davis E. A., Compton W. D.*— Phys. Rev., **140**, A2183 (1965)].

яние между донорами $<r>$ оказывается в пределах от $10a_d$ до $20a_d$. Туннелирование происходит по основным состояниям доноров, заполняющим энергетический интервал $\varepsilon_3 \approx \approx (e^2/4\pi\varkappa\varepsilon_0<r>)$, в результате действия беспорядочных кулоновских полей, источниками которых служат ионизованные доноры и ионизованные компенсирующие акцепторы. Поэтому туннельные переходы должны происходить с участием фононов и обнаруживать термическую энергию активации. Когда при очень низких температурах становится предпочтительным туннелирование на *более удаленные*, но *энергетически* более выгодные доноры, согласно выводу Остина и Мотта [29], член, отвечающий прыжковой проводимости, должен принять вид $\sigma_3' \exp(-A/T^{1/4})$.

Образцы с номерами от *E3* до *F3* на рис. 4.34 обнаруживают прыжковую проводимость вида $\sigma_3 \exp(-\varepsilon_3/k_0T)$ при температурах ниже примерно 5 К; по-видимому, этот член становится доминирующим для образца *L3* чуть ниже интервала температур, приведенного на рис. 4.34. Для образцов с номерами от *L3* до *E1* концентрация доноров достаточно велика, чтобы в большей части температурного интервала на рис. 4.34 их проводимость определялась механизмом $\sigma_2 \exp(-\varepsilon_2/k_0T)$. Примесная проводимость такого типа требует более близкого расстояния между донорами и не нуждается в присутствии акцепторов, обеспечивающих частичную ионизацию доноров. Принято считать, что ε_2 — это энергия, необходимая для того, чтобы поместить второй электрон на донор, на котором уже имеется один электрон (т. е. на нейтральный донор). Таким образом, «зона», по которой осуществляется проводимость,— это зона отрицательно заряженных доноров (D^--зона). Был предложен ряд других объяснений смысла ε_2, однако концепция проводимости по D^--состояниям тем не менее представляется наиболее вероятной [28].

Все состояния, связанные с донорами, становятся полностью делокализованной частью зоны проводимости, когда $N_d > N_{\text{криг}}$. Однако для локализованных состояний, которые существуют, пока $N_d < N_{\text{крит}}$, следует подчеркнуть важное различие между проводимостью за счет электронов, которые должны быть возбуждены в зону проводимости, и проводимостью за счет электронов, переходящих с одного примесного атома на другой. Для описания переноса электронов, находящихся в делокализованных блоховских состояниях, мы использовали уравнение Больцмана и обсудили связь между измеряемой подвижностью электронов и эквивалентными средним временем и средней длиной свободного пробега между рассеивающими столкновениями.

[29] *Austin I. G., Mott N. F.*— Advanc. Phys., 18, 41 (1969).

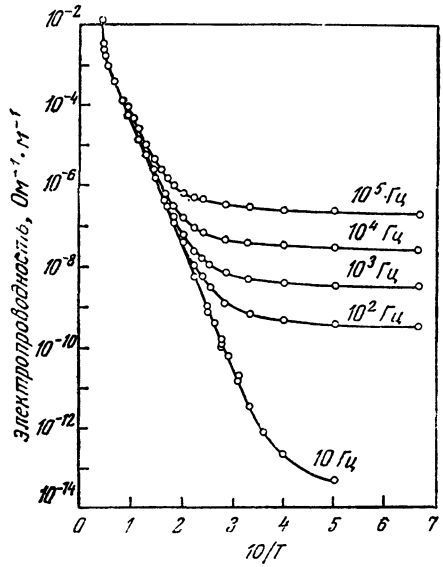


Рис. 4.35. Электропроводность на постоянном и переменном токе для слаболегированного и слабокомпенсированного кремния *n*-типа, обнаруживающая резкое возрастание возможности низкотемпературной прыжковой проводимости на больших частотах [*Pollak M., Geballe T. H.*— Phys. Rev., **122**, 1742 (1961)]

В случае переходов из одного локализованного состояния в другое картина совершенно иная.

Итак, «подвижность», измеряемая в ситуации примесной проводимости, обычно бывает очень мала. Это приводит к тому, что длина свободного пробега становится сравнима (или меньше) размера атома, помимо того, что само представление о длине свободного пробега при этом теряет смысл. Потому общее движение электронов от одного конца образца к другому в присутствии электрического поля по своей природе представляет собой *перколяционный процесс*, или *процесс протекания*[30]. Беспорядочные изменения расстояния между соседними донорами означают, что некоторые электронные траектории являются *связными* от одного конца образца до другого, в то время как другие микроскопические области кристалла не позволяют обеспечить связность с одним или с обоими токовыми контактами. Мы вновь встретимся с явлением протекания при обсуждении аморфных полупроводников; тогда разделение на связные и несвязные области будет определяться энергией электрона.

Одним из следствий перколяционного характера тока является то, что лишь часть полного объема дает вклад в низко-

[30] Различие между диффузией и протеканием обсуждается в работе: *Shante V. K. S., Kirkpatrick S.*— Advanc. Phys., **20**, 325 (1971). Применение теории протекания к движению электронов рассмотрено среди прочих в работе: *Ziman J. M.*— J. Phys. C, **1**, 1532 (1968).

температурную прыжковую проводимость на постоянном токе, когда $N \ll N_{крит}$. Но тогда объем, дающий вклад в электропроводность, и сама электропроводность должны резко возрасти, если эксперимент проводится на переменном токе. Рис. 4.35 иллюстрирует это на примере слабо легированного кремния, здесь $\sigma_\omega \sim \omega^{0,8}$. Различие между электропроводностью на постоянном и переменном токе такого же порядка наблюдается у некоторых аморфных полупроводников и других твердых тел «с малой подвижностью», к которым мы перейдем в следующем разделе.

Проводимость твердых тел с малой подвижностью

Помимо низкотемпературной примесной проводимости в полупроводниках, в которых при более высоких температурах имеется и другой механизм проводимости, обусловленный высокоподвижными электронами и дырками, существуют несколько классов твердых тел, в которых подвижность мала при всех температурах. Предыдущие замечания, касающиеся прыжковой проводимости и протекания, приложимы ко многим из этих твердых тел при комнатной температуре.

Малые электропроводность и подвижность обусловлены разными причинами в различных классах твердых тел: в ионных кристаллах, органических полупроводниках, оксидах и сульфидах переходных металлов, аморфных полупроводниках и т. д. Непосредственно очевидные приложения материалов этих классов время от времени разрабатываются, однако лежащие в основе их свойств физические принципы привлекают гораздо меньше внимания, чем теория более «легких» твердых тел. Такое отношение теперь меняется. После десятилетий почти полного пренебрежения исследования многих из этих материалов в 70-х годах стали популярны, и список литературы по ним быстро растет [31].

Малая подвижность, наблюдаемая во многих ионных кристаллах, частично обусловлена образованием поляронов, а частично присутствием большого числа локализованных дефектов, способных захватить движущийся носитель. В оксидах переходных металлов электроны проводимости происходят от частично заполненных $3d$-состояний изолированных атомов металла, которые слишком тесно связаны с порождающими их атомами, чтобы образовать «зону» полностью делокализован-

[31] Обзор большинства классов твердых тел с малой подвижностью см. в статьях, помещенных в трудах конференции по малой подвижности: Conduction in Low-Mobility Materials, Proceedings 2nd International Conference, eds. N. Klein, D. S. Tannhauser, M. Pollak, Barnes and Noble, 1971].

ных состояний в блоховском смысле. Следовательно, электрон должен совершать прыжки от одного атома к другому. В некоторых оксидах и сульфидах переходных элементов происходит переход[32] в *металлическое* состояние при нагревании, поскольку при этом возникает перекрытие, достаточное для того, чтобы делокализация состояний привела к образованию зон, из которых $3d$-зона только частично заполнена.

Органические полупроводники

Органические полупроводники в течение многих лет исследовались главным образом в виде прессованных поликристаллических порошков (что не является идеальным для понимания их основных свойств) и лишь в отдельных случаях в виде монокристаллов. Органический полупроводник служит хорошим примером *молекулярного кристалла*, в котором каждая молекула обладает сильными ковалентными связями между составляющими ее атомами, но молекулы организованы в кристалл силами, обычно не превышающими силы Ван-дер-Ваальса. Более сильное взаимодействие молекул имеет место, когда органическое вещество подвергается полимеризации, однако, если судить по примеру полиэтилена на рис. 3.21, высокая электропроводность не является автоматическим следствием для органического кристалла с хорошей трехмерной сеткой.

К сожалению, следует признать, что большинство органических кристаллов являются очень плохими проводниками, электропроводность которых при комнатной температуре, согласно хорошо известным данным[33], составляет всего 10^{-17} $Ом^{-1} \cdot м^{-1}$. Для антрацена, который является одним из наиболее интенсивно изучаемых молекулярных твердых тел, электропроводность в направлении, в котором она максимальна, составляет 10^{-12} $Ом^{-1} \cdot м^{-1}$ и значительно меньше в других направлениях. Проводимость в типичном органическом твердом теле осуществляется посредством прыжкового меха-

---

[32] См. *Mott N. F.* Metall-Insulator Transitions, Taylor and Francis, 1974, а также: *Mott N. F., Davis E. A.* Electronic Processes in Non-Grystalline Solids, Oxford, Univ. Press, 2nd ed., 1979. [Имеется перевод 1-го издания: *Мотт Н., Дэвис Э.* Электронные процессы в некристаллических веществах — М.: Мир, 1974, с. 471.]

[33] Источником ранней информации по многим слабопроводящим органическим веществам является книга: *Gutmann F., Lyons L. E.* Organic Semiconductors, Krieger, 1983. Более поздние данные содержатся в работе: *Pope M., Swenberg C. E.* Electronic Processes in Organic Crystals, Oxford Univ. Press, 1982.

низма и (или) туннелирования, и электропроводность часто зависит от температуры в соответствии с соотношением

$$\sigma = \sigma_0 \exp(-\Delta\varepsilon/k_0 T) \qquad (4.76)$$

в широком интервале температур. Энергия активации $\Delta\varepsilon$ нередко оказывается связанной не с температурной зависимостью концентрации электронов, а с изменением подвижности.

Одну из групп органических твердых тел, для которых возможна более высокая проводимость, составляют *комплексы с переносом заряда*, в которых один тип молекул действует как донор (D), в то время как другой тип, также присутствующий в кристалле, действует как акцептор (A). Для комплекса, в котором эти ингредиенты присутствуют в отношении 1 : 1, волновая функция основного состояния представляет собой *резонанс* между ионной и неионной формами донорно-акцепторной связи:

$$\psi = \alpha\psi_0(A{:}D) + \beta\psi_1(A^-{:}D^+). \qquad (4.77)$$

В разд. 1.1 указывалось, что такой резонанс между ионной и ковалентной конфигурациями приводит к более сильной связи. Здесь резонанс между ионной конфигурацией и конфигурацией, отвечающей связи Ван-дер-Ваальса, неизбежно приводит к более сильной связи, чем может обеспечить вандерваальсовский механизм сам по себе. Он приводит также к более плотной упаковке, чем обычно наблюдаемая в молекулярных кристаллах, которая открывает возможность перекрытия и уширения уровней в зоны. Эти зоны все еще остаются узкими, с большими эффективными массами и малыми подвижностями, однако их способность к проводимости в значительной мере улучшена по сравнению с проводимостью одиночного органического соединения. Плотная упаковка, в частности, имеет место для комплексов, приводя к образованию линейной колончатой структуры, составленной из сгруппированных донорных и акцепторных элементов. Это характерная черта ряда комплексов с переносом заряда, образуемых тетрацианохинодиметаном (TCNQ) с органическими, неорганическими и металлическими донорными структурными элементами[34]. На рис. 4.36 показана структура TCNQ и то, каким образом распределяется избыточной электронный заряд. Некоторые соли TCNQ являются металлами или полуметаллами, однако неустойчивость колончатой структуры по отношению к дисторсии Пайерлса (см. стр. 347) часто приводит к полупроводниковой ситуации. То же самое происходит и в солях с переносом заряда, содержащих другие, недавно синтезированные органические акцепторные радикалы.

[34] Первоначальное сообщение: *Kepler R. G., Bierstedt P. E., Merrifield R. E.*— Phys. Rev. Lett., 5, 503 (1960).

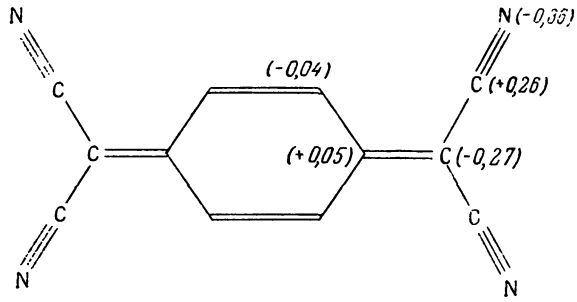


Рис. 4.36. Структура TCNQ. Числа в скобках показывают предсказываемый заряд в единицах *e* для каждого атома (или атомов в эквивалентных положениях), когда молекула находится в акцепторном зарядовом состоянии $TCNQ^-$, согласно работе *Vegter J. G. et al.* (см. примечание 31).

Спрессованные образцы из TCNQ-комплексов, часто имеют электропроводность, которая изменяется, согласно соотношению (4.76), в широком диапазоне температур; монокристаллы обладают такой же температурной зависимостью и, кроме того, обнаруживают сильную анизотропию электропроводности. Симонс и др.[35] представили для $Cs_2(TCNQ)_3$ экспериментальные данные, которые находятся в великолепном соответствии с зависимостью (4.76) для 13 порядков изменения удельного сопротивления [$\sigma_0 = 4 \cdot 10^4$ $Ом^{-1} \cdot м^{-1}$, $\Delta\varepsilon = 0{,}33$ эВ, что для электропроводности при комнатной температуре соответствует 0,1 $Ом^{-1} \cdot м^{-1}$]. Другие TCNQ-комплексы обнаруживают при комнатной температуре электропроводность, достигающую 10^4 $Ом^{-1} \cdot м^{-1}$ в благоприятном направлении и в 100 раз меньшую в направлении, перпендикулярному этому.

В органическом полупроводнике электропроводность представляет собой наиболее легко измеряемый параметр. Иногда это дополняется измерениями магнитной восприимчивости или электронного спинового резонанса (ЭПР). Попытки с помощью подходящей методики измерить подвижность обычно оказывались неудачными. Это не очень удивляет, поскольку наше понимание таких явлений, как эффект Холла, магнетосопротивление, термо-ЭДС основан на приближении кинетического уравнения Больцмана для возмущения газа подвижных электронов и на представлениях о средней длине свободного пробега и среднем времени между столкновениями. Все эти представления неприменимы для материала с малой подвижностью, в котором переходы осуществляются только посредством прыжко-

[35] *Siemons W. J., Bierstedt P. E., Kepler R. G.* — J. Chem. Phys., **39**, 3523 (1963).

вой проводимости и туннелирования. Измерения эффекта Холла, часто предпринимаемые в органических полупроводниках, были иногда полностью безуспешны или же приводили к ненадежным результатам. Так, комплексы TCNQ с *N*-метилфеноцианином дают как положительный, так и отрицательный знак напряжения Холла[36], причем в обоих случаях «холловская подвижность» по абсолютной величине оказывается меньше 10^{-6} м^2/В·с.

Другой интересный тип проводников, основанных на органической химии,— это «легированные полимеры», из которых легированный полиацетилен[37] позволил получить наиболее интересную информацию. Здесь мы имеем дело не с монокристаллом, а с полимером, степень пространственного упорядочения которого зависит от искусства выращивания $(CH)_x$-пленки. Будучи нелегированным, он обладает высоким сопротивлением, однако путем легирования щелочными металлами, галогенами и т. д. его можно получить с проводимостью *n*- и *p*-типа и высокой электропроводностью.

Аморфные полупроводники[38]

Аморфные твердые тела уже упоминались в гл. 1 в связи с их характерными структурными свойствами: отсутствием дальнего порядка даже при том, что каждый атом *может* иметь определенное число ближайших соседей на определенном расстоянии. Электронные свойства аморфных, стеклообразных и неупорядоченных систем и составляют содержание этих разделов. Было проведено множество исследований электрических характеристик аморфных твердых тел, однако часто их надежности препятствовало отсутствие контроля за условиями приготовления и получающейся структурой.

Наиболее важным аморфным полупроводником в последние годы был стеклообразный селен, активный материал для многих ксерографических копировальных устройств. Аморфные материалы с полупроводниковыми свойствами существуют как в одноатомном виде, так и в виде бинарных соединений, тройных и даже четверных сплавов, содержащих Si и Ge из группы IV, P, As, Sb и Bi из группы V и халькогенов O, S, Se и Te из группы VI периодической таблицы. Перечисленные элементы являются главными составляющими, хотя и некоторые оксиды переходных металлов могут существовать в аморфной полу-

---

[36] *Blythe A. R., Wright P. G.*— Phys. Lett., **34A**, 55 (1971).

[37] См., например, *Chien J. C. W.* Polyacetylene; Chemistry, Physics and Materials Science, Academic Press, 1984.

[38] См. книги Адлера и др., Мотта и Дэвиса и Тауца в библиографии в конце главы.

проводниковой форме, а другие элементы (как в $CdAs_2Ge$) могут объединяться с атомами перечисленных групп.

«Идеальное ковалентное стекло» определено как хаотическая сетка, не имеющая дальнего порядка, но с идеальным ближним порядком. Такое стекло не должно иметь структурных дефектов (типа пустот), и в нем должны удовлетворяться требования валентности для каждого атома, т. е. не должно быть оборванных связей. Вероятно, наилучшим приближением к такому идеалу служат аморфные пленки кремния или германия, приготовленные напылением в вакууме на холодную подложку. Данные Мосса и Грачика по дифракции электронов, приведенные на рис. 1.47, указывают на то, что каждый атом кремния имеет по четыре соседних атома кремния на том же расстоянии, что и в кристаллическом кремнии. Более того, как аморфный кремний, так и аморфный германий могут быть осаждены в виде пленок, плотность которых близка к плотности соответствующих кристаллов. В аморфной пленке Ge может быть небольшое количество пустот, так что ее плотность составляет 97 % от плотности кристалла германия.

Собственная щель ε_i (измеренная оптически) несколько меньше для высококачественной аморфной кремниевой или германиевой пленки, чем в соответствующем кристаллическом веществе. Аморфные слои с большим разупорядочением обнаруживают щель меньшего размера и менее отчетливую. Другие электронные свойства проявляют еще большее различие между кристаллом и структурами «идеального стекла». Так, на рис. 4.37 показаны четыре кривые «сканирования» плотности состояний валентной зоны германия посредством фотоэмиссии. Верхняя кривая для слоя, напыленного на подложку, находящуюся при 260 °С, обнаруживает сингулярности Ван Хова, характерные для плотности состояний валентной зоны кристаллического германия. Эти сингулярности не видны ни на одной из трех нижних кривых для аморфного германия. Такая аморфная пленка может быть кристаллизована путем последующего нагревания, и тогда сингулярности Ван Хова, присущие зоне кристаллического Ge, появятся. Следует сделать вывод, что плотность состояний $g(\varepsilon)$ для валентной зоны аморфного Ge *существует*, но она не похожа на $g(\varepsilon)$ для кристаллического Ge.

На рис. 4.38 показано изменение с обратной температурой электронной дрейфовой подвижности и электропроводности в напыленной пленке аморфного кремния толщиной 1,3 мкм. Значения $\Delta\varepsilon$, приведенные на рис. *б*, сравниваются с выражением (4.76), а энергии активации, отмеченные на кривой подвижности, аналогичным образом представляют собой значения, полученные сравнением с выражением $\mu \approx A \exp(-\varepsilon_\mu/k_0T)$. Поскольку для кристаллического кремния $\varepsilon_i = 1{,}1$ эВ, а $n_i \sim$

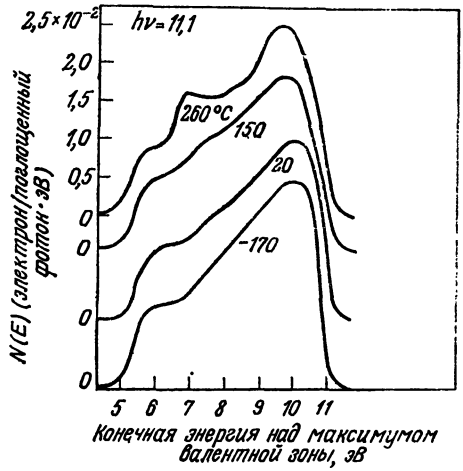


Рис. 4.37. Энергетический спектр электронов, фотоэмитированных из слоев германия, полученный с использованием фотонов с энергией 11,1 эВ для исследования плотности состояний валентной зоны. Различные кривые (с неопределенным положением нулевой ординаты) получены для разных температур подложки при напылении слоев; три нижних были аморфными, в то время как слой, образованный при 260 °С, был кристаллическим [*Ribbing C. G., Pierce D. T., Spicer W. E.*— Phys. Rev., **B4**, 4417 (1971)].

$\sim \exp(-\varepsilon_i/2k_0T)$ согласно выражению (4.17), на первый взгляд, имеется кажущееся сходство между наклонами двух частей кривой электропроводности на рис. 4.38, *б* и наклоном собственной электропроводности в упорядоченной решетке кремния. Однако как электропроводность, так и подвижность для такого аморфного кремния в 10^4 раз меньше, чем в собственном кристалле. Более того, подвижность на рис. 4.38, *а* имеет активационный характер, ненаблюдаемый в чистых кристаллах кремния.

Почему аморфный кремний должен иметь концентрацию носителей, не превышающую концентрацию собственного кремния, и малую активационную подвижность, можно схематически объяснить с помощью рис. 4.39. На рис. 4.39, *а* условно изображена плотность состояний полупроводникового кристалла. Валентные состояния показаны ниже ε_v, состояния зоны проводимости — выше ε_c. Несколько областей локализованных состояний в запрещенной зоне соответствуют примесям и собственным дефектам, возмущающим периодичность. На рис. 4.39, *б* показана ожидаемая плотность состояний в том же твердом теле, если оно является *идеальным стеклом*. Валентная зона и зона проводимости все еще отмечены, однако теперь континуум $g(\varepsilon)$ показан простирающимся вниз от ε_c и вверх

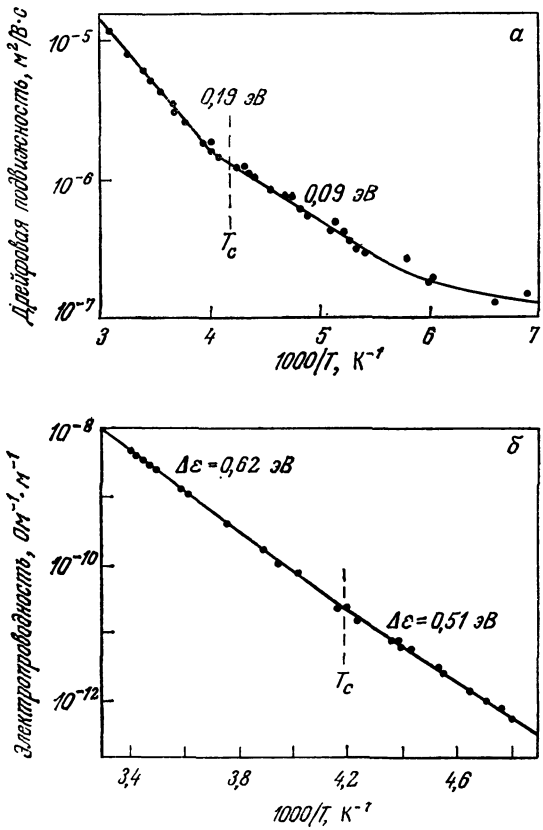


Рис. 4.38. Температурная зависимость дрейфовой подвижности электронов (*а*) и электропроводности (*б*) для аморфной кремниевой пленки [*Le Comber P. G., Spear W. E.*— Phys. Rev. Lett., 25, 509 (1970)].

от ε_v, т. е. в ту область, которая первоначально была собственной щелью. На рис. 4.39, *в* изображен еще более разупорядоченный полупроводник со многими оборванными («болтающимися») связями, обусловленными нарушениями ближнего порядка.

Локализованные состояния в запрещенной зоне, связанные с нарушениями ближнего порядка, неизбежно будут делиться поровну между *донороподобными* и *акцептороподобными* состояниями. Таким образом, разупорядочение аморфного полупроводникового элемента, соединения или сплава обладает свойством автоматической *самокомпенсации* путем образования электронных и дырочных ловушек в равных количествах.

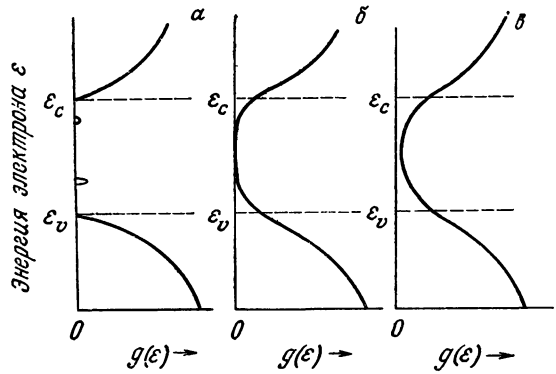


Рис. 4.39. Плотность состояний как функция энергии для кристаллического полупроводника (*а*), того же полупроводника в форме «идеального ковалентного стекла» (*б*) и того же материала в форме еще более разупорядоченного аморфного твердого тела (*в*). Такой вид плотности состояний для аморфного твердого тела основан на представлениях Мотта (см. ссылки в примечании 38) и др.

Энергия Ферми ε_F в результате этого оказывается «закреплена» в центральной части щели. Химические примеси, по-видимому, не могут передвинуть уровень Ферми далеко от средней точки щели, поскольку большинство примесных атомов не становится донорами или акцепторами, требования их валентности меняются в зависимости от локального окружения.

Сохраняет ли какой-то смысл понятие «щели», особенно если разупорядочение настолько велико, что конечная величина $g(\varepsilon)$ имеется для всех энергий? Да, очевидно, сохраняет. Все еще существует очень сильное возрастание оптического коэффициента поглощения, когда $\hbar\omega$ для падающих фотонов превосходит $(\varepsilon_c - \varepsilon_v)$ и допускает образование электронно-дырочных пар. Фотоны с меньшей энергией не создают *подвижных* пар. Дело в том, что электрон в состоянии над ε_c является подвижным, хотя мы и не можем воспользоваться функцией Блоха для описания его волновой функции. Однако состояние, для которого $\varepsilon_v < \varepsilon < \varepsilon_c$, всегда локализовано, и электрон может перемещаться только посредством прыжков или туннелируя из одного такого состояния в другое со сравнимой энергией, расположенное поблизости. Таким образом, ε_c и ε_v называют краями *щели подвижности* для аморфного полупроводника.

Результаты, подобные приведенным на рис. 4.38, могут быть интерпретированы с помощью рис. 4.39, *б*. Наблюдаемая электропроводность — это электропроводность *собственного* аморфного кремния, поскольку те немногие примесные атомы, которые были напылены вместе с кремнием, займут координа-

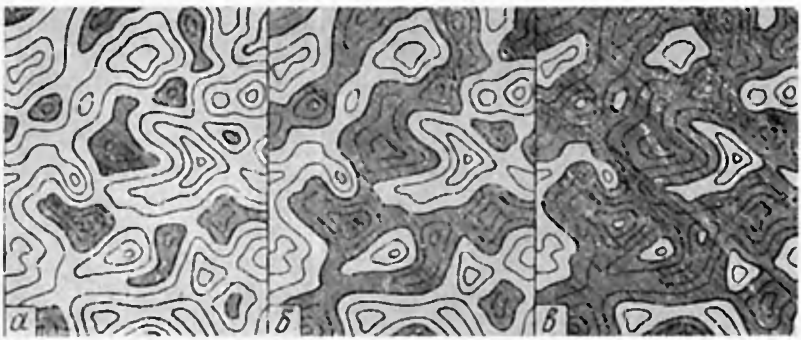


Рис. 4.40. Применение теории протекания к понятию края подвижности в аморфном полупроводнике [*Zallen R., Scher H.*— Phys. Rev., **B4**, 4471 (1971)]. Части *а*, *б* и *в* показывают области (затененные) открытых путей проводимости для трех последовательно возрастающих электронных энергий.

ционные конфигурации, совместные с числом их валентных электронов. Равенство донороподобных и акцептороподобных состояний приведет при этом к «закреплению» уровня Ферми в центре щели. Для температур ниже той, которая обозначена T_c, перенос осуществляется посредством перескоков между локализованными состояниями внутри щели. Поскольку этот процесс происходит с участием фононов оптической моды [39] с энергией примерно 0,07 эВ, можно понять, почему подвижность в этом прыжковом диапазоне должна меняться как $\exp(-0{,}09/k_0T)$. Для более высоких температур подвижность возрастает быстрее, поскольку электроны *могут* быть термически возбуждены выше ε_c, однако скорость их движения существенно ограничена частыми падениями в пустые локализованные состояния, расположенные выше ε_F, но ниже ε_c. Такие состояния служат *ловушками* для электронов. Отношение «времени жизни в свободном состоянии» к «времени на ловушке» изменяется с температурой, что и дает температурную зависимость подвижности вида $\exp(-0{,}19/k_0T)$.

На рис. 4.40 изображен двумерный аморфный полупроводник для энергий, расположенных ниже «края подвижности» ε_c, на нем и выше него. Для низкой энергии на рис. 4.40, *а* имеется

[39] Что означает и означает ли что-нибудь вообще понятие «оптический фонон» применительно к аморфному твердому телу? Ведь в этом случае нет периодической решетки и нет зоны Бриллюэна. Хотя все это и так, но тем не менее мы можем трактовать длинноволновый оптический фонон, как состояние такого движения, при котором соседние атомы движутся в противофазе друг с другом. Каждый атом в аморфном теле продолжает иметь ближайших соседей; следовательно, энергия $\hbar\omega_{\text{опт}}$ остается применимой для описания колебаний длины связи между ближайшими соседями.

несколько маленьких изолированных областей, в которых электроны могут двигаться более свободно, однако между этими областями нет связи. Ток может протекать от одного конца образца к другому только посредством электронных перескоков. Промежуточная энергия на рис. 4.40, *б* открывает некоторые связные каналы, допускающие протекание на большие расстояния без прыжков. Для еще большей энергии (рис. 4.40, *в*) остаются только немногие изолированные области, которые не дают вклада в полный ток.

Таким образом, способность аморфного полупроводника проводить ток зависит от наличия электронов в состояниях выше ε_c (или дырок в состояниях ниже ε_v). Вследствие этого почти все аморфные тела обладают электропроводностью, которая растет при нагревании. Более того, исследование вольтамперных характеристик тонких пленок аморфного материала в четверных системах Te : As : Si : Ge показывает, что сильное электрическое поле может создавать локализованные области с высокопроводящим состоянием.

Халькогенидные стекла, такие, как сплавы Fe : As : Si : Ge представляются многообещающими для использования в качестве переключателей [40] и элементов памяти в электронных системах, благодаря их вольтамперным характеристикам, полученным в тонких пленках. Для пленки толщиной в несколько микрон ток в малых полях является чувствительной функцией температуры и нелинейно зависит от приложенного напряжения. Это соответствует состоянию переключателя «выкл» с сопротивлением $\sim 10^7$ Ом при комнатной температуре. Если поле превышает 10^7 В/м, то происходит переход в проводящее состояние «вкл» с сопротивлением в несколько десятков омов. Очевидно, переключение может быть повторено много раз без значительного повреждения пленки.

Способность пленок из халькогенидного стекла к переключению иногда приписывалась тепловому пробою пленки на части ее площади, и такая возможность вполне допустима в случае более толстых пленок. Исследование [41], проведенное в импульсном режиме для очень тонких пленок, показывает, что пробой является в первую очередь электронным, а не тепловым. В этом случае представляется возможным, что один или несколько токовых шнуров (диаметром в несколько микрон, обеспечивающих плотность тока порядка 10^{10} А/м^2) устанавливают такое *неравновесное* распределение электронов, при котором все ловушки внутри щели оказываются насыщены, а ток

---

[40] *Adler D. et al.*— J. Appl. Phys., **51**, 3289 (1980).

[41] *Heinisch H. K. et al.*— J. Non-Cryst. Solids, **8—10**, 415 (1972). *Adler D. et al.*— Rev. Mod. Phys., **50**, 209 (1978).

в шнуре поддерживается плотной электронно-дырочной плазмой [42]. Каждый шнур окружен веществом, содержащим ненасыщенные ловушки, и этим саморегулируется расположение токового канала и его размер.

Отклонения от закона Ома имеют место как в аморфных полупроводниках, так и в кристаллических, что мы и рассмотрим в следующем разделе.

Горячие электроны

Из всех предыдущих рассуждений об электронном дрейфе, налагаемом на гораздо более интенсивное тепловое движение, становится ясно, что дрейфовая скорость электронов зоны проводимости в полупроводнике должна возрастать с увеличением напряженности приложенного электрического поля по крайней мере до известного предела. Разумно предположить, что предельная дрейфовая скорость будет достигнута при достаточно большом поле, и это предположение потверждается экспериментальными кривыми, подобными кривой для электронов в германии, представленной на рис. 4.41. Мы обсудим эту кривую, имея в виду, что анализ такого рода можно провести для многих других полупроводников.

На кривой рис. 4.41 отмечено пять характерных областей. Они могут быть описаны следующим образом [43]. Область A — это обычный режим слабого поля, при котором скорость прямо пропорциональна напряженности поля, так что выполняется закон Ома. В этой области слабых полей многочисленные столкновения каждого электрона с фононами делают среднюю энергию электрона не зависящей от времени и соответствующей равновесному распределению.

Мы должны, однако, помнить, что каждое электрон-фононное соударение приводит к очень малому изменению энергии электрона. По мере увеличения приложенного электрического поля электрону становится все труднее избавиться от энергии, которую он набирает в электрическом поле, и распределение электронных энергий начинает «нагреваться». Можно наде-

---

[42] Это можно сравнить с током, ограниченным пространственным зарядом, много лет изучаемым в изолирующих фотопроводниках с заполненными ловушками. См. *Rose A.* Photoconductivity of Solids, Wiley-Interscience, 1963, ch. 4. [Имеется перевод, *Роуз А.* Основы теории фотопроводимости.— М.: Мир, 1966.]

[43] Обширный обзор теории и эксперимента по проводимости германия и кремния в сильном поле сделан Ганном [*Gunn J. B.* Progress in Semiconductors, Wiley, 157, v. 2; см. также: *Conwell E. M.* High Field Transport in Semiconductors, Academic Press, 1967].

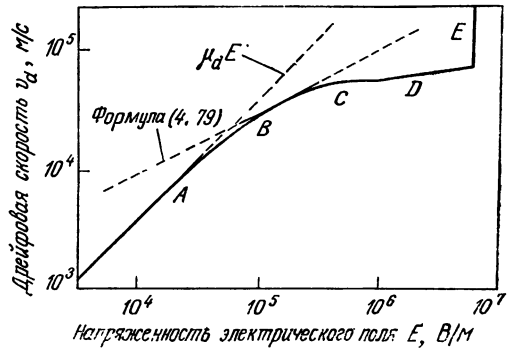


Рис. 4.41. Дрейфовая скорость электрона в германии как функция напряженности электрического поля при 300 К [*Gunn J. B.*— J. Electronics, 2, 87 (1956)]. Пять областей *A*—*E* описаны в тексте.

яться, что теплый электронный газ все еще может быть описан с помощью распределения Больцмана

$$dn \sim s^2 \exp(-m_c s^2/2k_0 T_e), \tag{4.78}$$

однако эффективная температура электрона T_e выше, чем температура решетки T, и является функцией напряженности приложенного поля. Анализ [43] стационарного баланса между скоростью поглощения энергии «теплыми» электронами и скоростью, с которой они могут потерять энергию путем рассеяния на *LA*-фононах, показывает, что дрейфовая скорость должна изменяться с полем E в соответствии с выражением

$$v_d = \frac{2\mu_d E}{[2 + \sqrt{4 + (E/E_1)^2}\,]^{1/2}}. \tag{4.79}$$

Здесь μ_d обозначает дрейфовую подвижность в слабом поле, E_1 — параметр, определяемый скоростью звука в полупроводнике. В соответствии с формулой (4.79) дрейфовая подвижность должна принимать два предельных выражения:

$$v_d = \mu_d (1 - E^2/32E_1^2), \qquad E \ll E_1, \tag{4.80}$$

$$v_d = 2\mu_d (EE_1)^{1/2}, \qquad E \gg E_1, \tag{4.81}$$

для полей, малых и больших по сравнению с параметром E_1. Эксперименты подтвердили [43], что дрейфовая скорость для ряда полупроводников действительно изменяется (от $v_d = \mu_d E$ в очень малых полях), принимая сначала вид (4.80), а затем (4.81) по мере увеличения поля.

Участок кривой B на рис. 4.41 находится в согласии с выражением (4.81), однако очевидна необходимость учета каких-то других механизмов, чтобы обеспечить дальнейшее ослабление роста v_d с приложенным полем. Область насыщения скорости (C) достигается потому, что горячие электроны теперь обладают энергией, достаточной для возбуждения оптических фононов. Вскоре после того, как электрон разогревается полем до энергии, превосходящей $\hbar\omega_{\text{опт}}$, становится вероятным возбуждение оптической моды, которая забирает почти всю его энергию. Таким образом, в области C бóльшая часть проводимости осуществляется электронами непосредственно перед тем, как они набирают от поля энергию, достаточную для того, чтобы стать чувствительными к данному механизму потери энергии. Дальнейшее увеличение поля, как видно по участку D кривой, приводит к очень умеренному росту тока, поскольку весьма малому числу электронов удается в течение какого-то времени сохранять энергию, превышающую $\omega_{\text{опт}}$.

Конечный участок кривой на рис. 4.41 обнаруживает *резкое увеличение числа носителей*, которое происходит в результате *лавинного пробоя*. При некотором критическом поле части электронов удается преодолеть энергетический рубеж $\hbar\omega_{\text{опт}}$ и достигнуть энергии, достаточной для генерации дополнительных электронно-дырочных пар в полупроводнике, и тогда ток резко возрастает при очень малом изменении приложенного напряжения (участок E). Механизм лавинного пробоя устанавливает верхний предел напряжений, допустимых в большинстве полупроводниковых приборов, использующих p—n-переходы.

В примесном полупроводнике при температурах, достаточно низких, чтобы носители из проводящих зон перешли на примесные центры, процесс лавинного пробоя может произойти в том случае, когда часть электронов, оставшихся в зоне, наберет в электрическом поле энергию, достаточную для ударной ионизации нейтральных центров. Для мелких примесных уровней этот примесный тип диэлектрического пробоя может произойти уже в поле 500 В/м. Требуемое поле, конечно, становится больше для более «глубоких» примесных состояний.

Вернемся теперь к изменению дрейфовой скорости с полем в том случае, когда ионизация примесных центров *не происходит*. Под термином «дрейфовая скорость» мы обычно понимаем величину $v_d = (E\sigma/en)$ для ансамбля n электронов, ускоряемых полем E. На рис. 4.41 показан очевидный монотонный рост v_d с E для электронов в германии, однако этот рост не обязательно во всех случаях является монотонным. Действительно, выбрав соответствующим образом ориентацию кристалла, с помощью одноосного сжатия можно заставить v_d для электронов в германии слегка уменьшаться для полей, превышающих некоторое

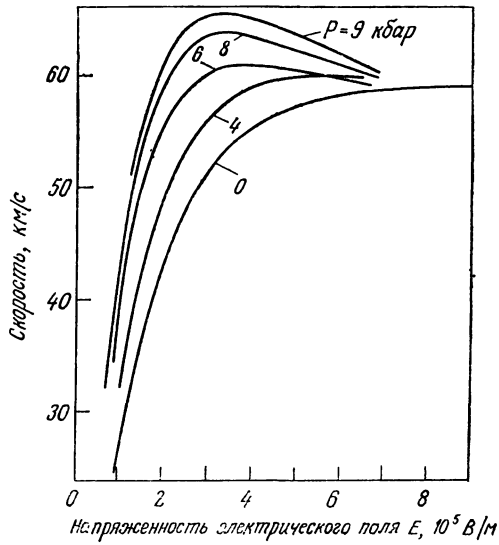


Рис. 4.42. Дрейфовая скорость электрона в германии как функция напряженности электрического поля при 300 К, подвергнутого одноосному сжатию в направлении [111] и без него. Поле приложено в направлении [11$\bar{2}$]. Кривые получены методом измерения «пролетного времени», в течение которого электроны инжектируются и дрейфуют через образец германия *p*-типа высокой чистоты [*Neukermans A., Kino G. S.*— Phys. Rev., **B7**, 2963 (1973)].

значение, как это показано на кривых рис. 4.42. Гораздо более явная *ретроградная скорость* *, или *отрицательная дифференциальная проводимость* (ОДП), показана на кривой для GaAs (рис. 4.43), причем для GaAs не требуется никаких анизотропных деформаций решетки, чтобы этот эффект был большим.

Такое несколько удивительное уменьшение эффективной электропроводности, приходящейся на один электрон, которое происходит в сильных полях в GaAs, является следствием структуры зон проводимости в этом полупроводнике. Из рис. 4.44 видно, что низший энергетический минимум зоны проводимости арсенида галлия находится в центре зоны и все электроны в тепловом равновесии размещаются в нем. Этот предпочтительный для электронов минимум, расположенный в точке $\mathbf{k}=(0, 0, 0)$, характеризуется малой эффективной массой и соответственно высокой подвижностью. Однако существуют и другие минимумы зоны проводимости (с боль-

* В отечественной литературе характер зависимости v_d от E принято называть *N*-образным.— *Прим. перев.*

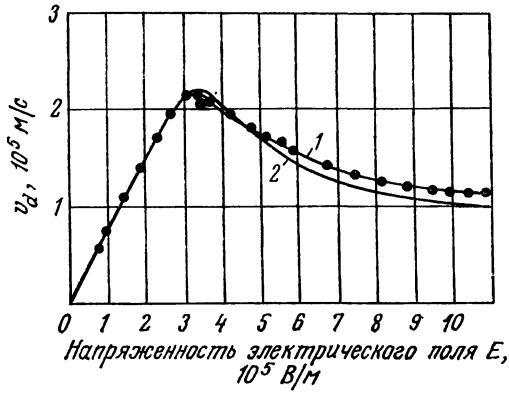


Рис. 4.43. Изменение дрейфовой скорости с электрическим полем для электронов в арсениде галлия. Кривая 2, рассчитанная Батчером и Фосеттом [*Butcher P. N., Fawcett W.*— Phys. Lett., **21**, 489 (1966)], сравнивается с экспериментальными данными Рача и Кино (кривая *1*) [*Ruch J. G., Kino G. S.*— Appl. Phys. Lett, **10**, 40 (1967)]. Из работы: *Kroemer H.* IEEE Spectrum, 5, 47 (1968).

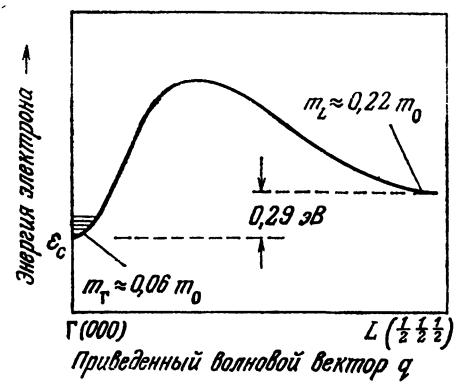


Рис. 4.44. Самая нижняя зона проводимости для арсенида галлия, обнаруживающая одиночный минимум в центре зоны, и зависимость энергии от приведенного волнового вектора при движении к одному из последующих минимумов при $L(1/2\ 1/2\ 1/2)$. Более подробную зонную схему для GaAs см. также на рис. 4.48.

шей массой и меньшей подвижностью) при более высоких энергиях; на рис. 4.44 показано, что ближайшие и наиболее важные из них расположены в восьми точках $\mathbf{k}=\left(\frac{1}{2}\frac{1}{2}\frac{1}{2}\right)$ на границе зоны.

Существование более высоко расположенных минимумов не существенно в тепловом равновесии и для слабого прило-

женного поля. Однако эти долины могут стать заселенными за счет уменьшения заполнения центрального минимума зоны в результате процессов рассеяния в сильном электрическом поле. Так, одно из следствий большой величины приложенного поля состоит в том, что некоторые электроны под действием поля переносятся из зонных состояний центра зоны (где они обладают низкой эффективной массой и высокой подвижностью) в зонные состояния этих верхних «долин», где их эффективная масса гораздо больше, а подвижность меньше. Конечно, они будут стремиться вернуться в низшую зону, и в стационарном состоянии будет существовать динамическое равновесие между числом переходов вверх, обусловленных полем, и числом спонтанных переходов вниз. Поскольку в GaAs для полей, превышающих 300 кВ/м, способность многих электронов к переносу тока оказывается уменьшенной таким вынужденным перемещением электронов, наблюдаемая кривая для $v_d = (E\sigma/ne)$ падает по мере того, как E возрастает. Для *достаточно* большого поля величина v_d должна вновь начать расти.

Зону проводимости германия мы будем обсуждать в разд. 4.3. и увидим, что в этом случае самая нижняя зона имеет четыре эквивалентные «долины». Небольшая ОДП, показанная для германия на рис. 4.42, возникает из-за вызванного полем перераспределения электронов между этими долинами. ОДП была обнаружена для целого ряда полупроводников, и было предсказано[44], что ряд четверных интерметаллических твердых растворов окажется в проявлении ОДП более эффективным, чем GaAs.

Что происходит, когда *среднее* электрическое поле в полупроводниковом образце соответствует области ретроградной зависимости скорости электрона от поля? Ганн[45] экспериментально показал, что соответствующее напряжение, приложенное к образцу GaAs, приводит к образованию в материале последовательности областей слабого и сильного поля — *доменов* — таким образом, что производная (dv_d/dE) в обоих видах доменов положительна. Домен сильного поля образуется на одном конце образца, движется вдоль него (со скоростью, приблизительно равной v_d для домена слабого поля) и исчезает на другом конце. Ток, переносимый в полупроводниковом образце, осциллирует с частотой, которая обратна пролетному

---

[44] См., например, *Shaw M. P., Grubin H. L., Solomon P. R.* The Gunn-Hilsum Effect, Academic Press, 1979.

[45] *Gunn J. B.*— Solid State Comm., **1**, 88 (1963); IBM J. Res. Devel., 8, 141 (1964).

времени домена, и это представляет основную форму проявления *эффекта Ганна*, который имеет множество применений [46].

Поскольку скорость домена в GaAs порядка 10^5 м/с, то в образце этого полупроводника при длине токового пути 5 мкм будут генерироваться колебания с частотой 20 ГГц; эта величина близка к максимально достижимой для приборов, основанных на эффекте Ганна с движущимися доменами. Несколько бóльшая частота возможна для колебаний в режиме ограниченного накопления объемного заряда (ОНОЗ), при котором не происходит полного распространения домена.

Краткое описание эффекта Ганна лишний раз напоминает нам о том, что серьезное обсуждение многих явлений в полупроводниках невозможно без учета особенностей зонной структуры реальных полупроводников. Это и составляет содержание следующего раздела.

4.3. Зонная структура реальных полупроводников

Идеально простая зонная структура полупроводника

При обсуждении *концентраций* свободных и связанных электронов и дырок и соответствующей им энергии Ферми безразлично, действительно ли тензоры эффективной массы для зон сводятся к скалярным величинам или же в плотность состояний просто подставлены их «скалярные эквиваленты». Однако многие явления в полупроводниковых материалах радикально зависят от фактической формы поверхностей постоянной энергии в трехмерном **k**-пространстве. Некоторые указания на это уже следовали из рассмотрения процессов переноса, а в данном разделе мы рассмотрим явления, демонстрирующие простой или сложный характер зонной структуры в гораздо большей степени.

Исключительно для удобства рассмотрения предположим, что полупроводник с идеально простой зонной структурой образует кристалл с *трансляционной* симметрией г. ц. к.-решетки (гранецентрированной кубической). (Такое предположение позволяет нам рассматривать любое твердое тело, кристаллизующееся как в г. ц. к.-решетке, так и в решетке каменной соли, алмаза или цинковой обманки.) Тогда энергетические зоны

---

[46] Различные виды приборов с «переносом электронов» обсуждаются в книге Шоу и др. (примечание 44), а также Цзе и Майлнсом (см. общую библиографию на с. 12).

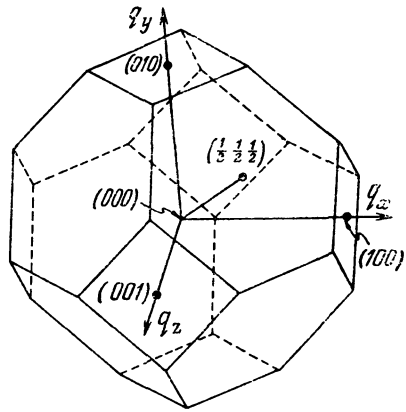


Рис. 4.45. Приведенная зона Бриллюэна для твердого тела, кристаллизующегося в решетке с г. ц. к.-структурой алмаза и цинковой обманки (см. рис. 1.31). Чтобы получить истинную величину компоненты k, каждую координату следует умножить на величину $(2\pi/a)$, где a — длина ребра элементарного куба. Для любой точки, находящейся между (000) и (100), существует еще пять точек, эквивалентных вследствие симметрии. Аналогичным образом имеется восемь эквивалентных точек для любого направления, описываемого как (*ttt*). Для точки зоны, которая не находится на одной из главных осей симметрии, существует 47 других эквивалентных точек, которые должны обладать таким же набором собственных значений энергий.

могут быть представлены набором поверхностей постоянной энергии внутри приведенной зоны Бриллюэна, изображенной на рис. 4.45. Ниже в этом разделе мы покажем, какой вид имеют непростые зоны в зоне Бриллюэна (ЗБ) такой же формы.

«Идеально простая» зонная структура изображена на рис. 4.46. Предполагается, что единственный минимум зоны проводимости находится при $\mathbf{k}=(000)$ и что для любой энергии, слегка превышающей ε_c, поверхность постоянной энергии может быть изображена в виде сферы с центром в точке минимума. Это означает, что эффективная масса электрона m_c представляет собой скаляр. Если потребовать, чтобы величина $(\varepsilon-\varepsilon_c)$ возрастала пропорционально k^2, то эта масса также не зависит от энергии. На рис. 4.46 показано, что энергия монотонно возрастает вплоть до самой границы ЗБ как вдоль ребра куба, так и вдоль его диагонали. Предельные энергии на границе ЗБ для этих двух направлений в обоих случаях настолько превышают ε_c, что точная форма зоны для свободных электронов не имеет значения.

Очевидно, должны существовать другие совершенно незаполненные зоны при энергиях, расположенных выше, чем самая нижняя зона проводимости, однако предполагается, что они

находятся настолько высоко, что не нарушают постулированную простоту изображенной зоны. Аналогично полупроводник должен иметь несколько полностью заполненных зон, однако на рис. 4.46 показана только самая верхняя из них. Предполагается, что самая верхняя валентная зона имеет форму, подобную форме зоны проводимости (только перевернутую по отношению к последней), с максимумом в точке (000) и что состояния для энергий, расположенных слегка ниже максимума, характеризуются скалярной и не зависящей от энергии массой m_v.

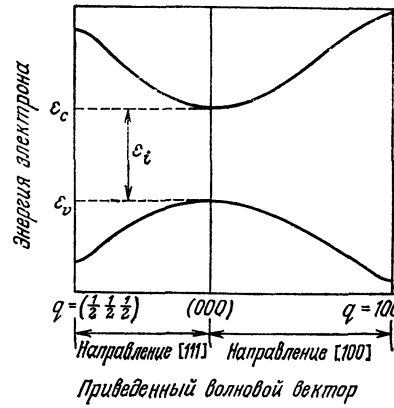


Рис. 4.46. Идеально простая зонная структура для полупроводника.

Именно такой картиной мы и пользовались в двух последних разделах, за исключением отдельных случаев, касающихся особенностей явлений переноса для *реальных* зон. Теперь следует признать, что до сих пор еще не было найдено такого полупроводника, в котором вблизи собственной щели зонная структура оказалась бы (после исчерпывающего изучения) настолько простой, как это показано на рис. 4.46.

Примеры зон для реальных твердых тел

Констатировав, что полупроводники, по-видимому, всегда обладают сложной зонной структурой, мы приведем несколько примеров, сопровождая их пояснениями о тех методах, с помощью которых различные виды особенностей зонной структуры могут быть исследованы. Те картины зон, которые приводятся на нескольких следующих страницах, не следует считать количественно точными, поскольку в данной области непрерывно поступающая новая информация постоянно заставляет пересматривать наши представления о величине энергетических щелей и т. д.

По-видимому, наиболее близка к нашей «идеализованной» структуре рис. 4.46 зонная структура сульфида кадмия, изображенная на рис. 4.47. Это твердое тело обычно кристаллизуется в решетке вюрцита (см. рис. 1.32), которая обладает единственным неэквивалентным направлением. Как можно видеть, в соответствии с этим имеется единственная самая верхняя валентная зона, однако эффективная масса в этой зоне сильно

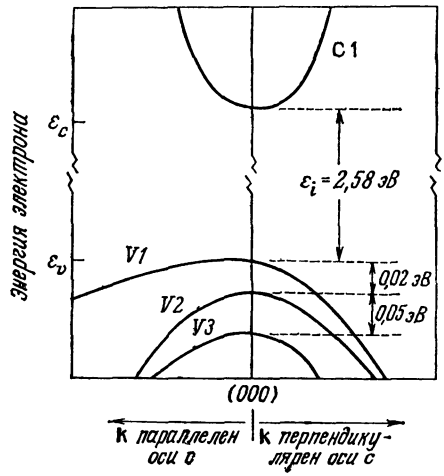


Рис. 4.47. Три самые верхние валентные зоны и самая нижняя зона проводимости для CdS — твердого тела, кристаллизующегося в структуре вюрцита, изображенной на рис. 1.32. Отметим, в частности, сильную анизотропию массы для самой верхней валентной зоны.

зависит от направления k. Кроме того, имеются две другие валентные зоны, отделенные от самой верхней небольшими энергетическими зазорами, так что свободные дырки в CdS при высоких температурах должны быть распределены в верхних частях всех трех валентных зон. В результате свободные дырки в CdS имеют очень малую подвижность. Однако поведение зоны проводимости оказывается достаточно простым и анизотропия эффективной массы в этой зоне мала.

Несколько более сложная зонная структура, однако, такая, при которой экстремумы зоны проводимости и валентной зоны все еще расположены в центре зоны Бриллюэна, показана на рис. 4.48. Для этих зон арсенида галлия валентные зоны $V1$ и $V2$ являются *вырожденными* [47] в точке их верхнего экстремума. Конечно, любая теория валентной зоны в таком материале должна требовать вырождения всех трех валентных зон в центре зоны Бриллюэна до тех пор, пока не учитывается *спин-орбитальное*

[47] Здесь термин «вырожденный» использован в спектроскопическом смысле, т. е. означает, что «два или более состояний обладают одной и той же энергией». В примечании 16 гл. 3 мы уже говорили, что термин «*вырожденный*» используется как в этом случае, так и для описания конденсированного электронного газа Ферми — Дирака с «вырожденной» электронной теплоемкостью.

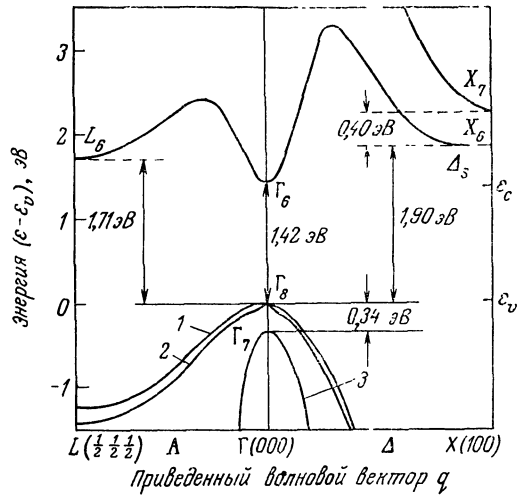


Рис. 4.48. Зоны, находящиеся непосредственно над собственной щелью арсенида галлия и непосредственно под ней ($T \approx 300$ К). Этот материал кристаллизуется в структуре цинковой обманки, изображенной на рис. 1.31, *б*. Зона Бриллюэна представляет собой многогранник, изображенный на рис. 4.45. *1* — зона тяжелых дерок *V1*; *2* — зона легких дырок *V2*; *3* — отщепленная зона *V3*.

взаимодействие[48]. Валентную зону *V3* на рис. 4.48 называют *отщепленной*, а энергетический зазор, составляющий, как отмечено, 0,36 эВ,— спин-орбитальным расщеплением.

Зона *V3* в арсениде галлия расположена достаточно низко, чтобы оставаться незаполненной свободными дырками, распределяющимися по верхним состояниям зон *V1* и *V2*. Зону *V1*, обладающую большей плотностью состояний (т. е. большой эффективной массой), обычно называют *зоной тяжелых дырок*, а зону *V2* — *зоной легких дырок*. Из-за сходства кривизны зависимостей $\varepsilon(\mathbf{k})$ для левой и правой частей зонной структуры на рис. 4.48 может показаться, что тяжелые и легкие дырки в соответствующих зонах по крайней мере обладают скалярными эффективными массами. Однако более детальный анализ показывает, что такое впечатление ошибочно. Поверхность постоянной энергии отклоняется от сферической формы вдоль

[48] Спин-орбитальные взаимодействия происходят как в твердом теле, так и в изолированных атомах. Явление состоит во взаимодействии магнитного поля, обусловленного электронным спином, с магнитным полем, вызванным орбитальным движением того же самого электрона. Роль этого взаимодействия в расщеплении энергетических зон, которые без него были бы вырожденными, показана Эллиотом [*Elliot R. J.*— Phys. Rev., **96**, 266 (1954)].

направлений [100] для тяжелых дырок и [111] для легких дырок. Кроме того, имеются теоретические предпосылки полагать, что зона $V1$ достигает максимальной энергии в нескольких эквивалентных точках k-пространства, расположенных на малом расстоянии от $\mathbf{k}=(000)$.

Зоны проводимости GaAs менее сложны, чем валентные зоны, однако и они не соответствуют «идеализированной» ситуации на рис. 4.46. Самый нижний минимум зоны проводимости, обозначенный Γ_6, обладает скалярной эффективной массой; однако энергия над ε_c растет медленнее, чем k^2, так что величина эффективной массы m_c должна увеличиваться с ростом энергии. В тепловом равновесии верхние долины зоны проводимости, обозначенные L_6, не содержат свободных электронов, однако электроны могут быть возбуждены из Γ_6 в L_6 в сильном электрическом поле, как говорилось в разд. 4.2. Если в зоне Бриллюэна арсенида галлия изобразить поверхности постоянной энергии для L_6-долин, то энергии непосредственно над дном такой зоны будет соответствовать восемь вытянутых полуэллипсоидов, напоминающих изображенные для Ge на рис. 4.50.

Конечно, арсенид галлия и другие соединения $A_{III}B_V$, являющиеся его гомологами, имеют различные зоны проводимости. На рис. 4.48 показаны только две нижние из них. Более высокие зоны не играют роли при исследовании квазиравновесных состояний полупроводника, однако они могут повлиять на оптические и фотоэмиссионные его характеристики в видимой и ультрафиолетовой частях спектра. Когда мы рассматриваем зонную структуру твердого тела, кристаллографически подобного GaAs, однако с другим представлением Фурье для периодического потенциала, порядок расположения зон вполне может оказаться другим. В качестве примера на рис. 4.49 показаны зоны, лежащие выше и ниже собственной щели, для элементарных полупроводников кремния и германия. Алмазная решетка каждого из этих элементов представляет собой структуру цинковой обманки GaAs, взятую с двумя идентичными атомами в качестве базиса.

Первое, что можно увидеть на рис. 4.49,— это то, что порядок следования валентных зон совершенно такой же, как для валентных зон GaAs на рис. 4.48. Энергия спин-орбитального расщепления очень мала как для германия, так и для кремния (0,28 эВ в германии и только 0,04 эВ в кремнии). Однако во всех трех случаях имеется «зона тяжелых дырок» с экстремумом, расположенным в точке (000), и поверхностями энергии, несколько отклоняющимися от сферической формы для более низких энергий; эта зона определяет бóльшую часть свойств кристаллов p-типа.

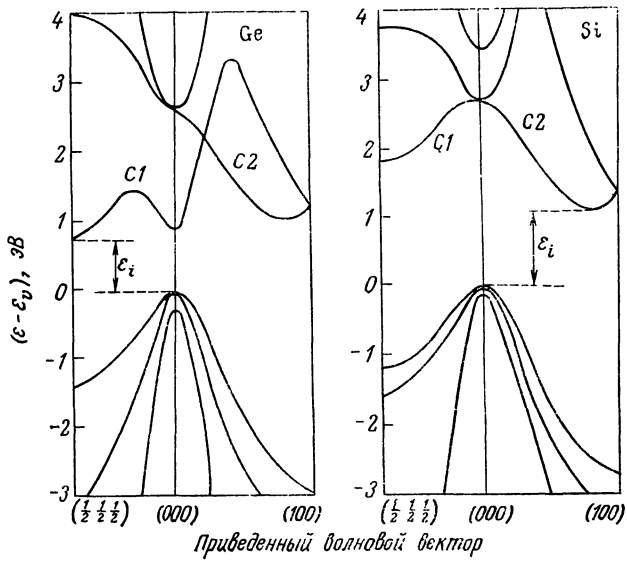


Рис. 4.49. Полуколичественный вид зависимости энергии электрона от приведенного волнового вектора для трех верхних валентных зон и нескольких первых зон проводимости германия и кремния. Зоны изображены в соответствии с представлениями Германа и др. [*Herman F. et al.* Quantum Theory of Atoms, Molecules and the Solid State, ed. P. O. Löwdin, Academic Press, 1966, p. 381].

Важное различие между структурой, изображенной на рис. 4.48, и двумя структурами на рис. 4.49 заключается в положении точки (точек) k-пространства, отвечающих минимально возможной энергии в зоне проводимости. В случае германия минимальное значение достигается не в центре ЗБ (как в GaAs), а в восьми точках на границе зоны при $q=[^1/_2{}^1/_2{}^1/_2]$; электроны зоны проводимости будут преимущественно занимать состояния в этих восьми областях ЗБ. В кремнии наинизшее значение энергии зоны проводимости достигается в шести точках, для которых $q=[^3/_4 00]$.

Для зон такого типа, как зона проводимости кремния и германия, *плотность состояний* $g(\varepsilon)$ и *эффективная масса плотности состояний* m_c должны учитывать вклад от всего набора эллипсоидов, а не только от одного минимума энергии. На рис. 3.50 мы уже показывали набор из шести эллипсоидов, позволяющий описать энергию, расположенную над ε_c в кремнии. Предположим, что минимум зон проводимости располагается в $\mathfrak{N}$ эквивалентных точках *внутри* приведенной зоны. Предположим также, что мы можем нарисовать набор эллипсоидных

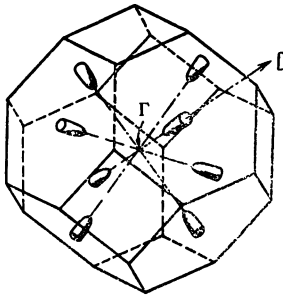


Рис. 4.50. Поверхности постоянной энергии в ЗБ для энергии, расположенной над ε_c в германии [*Long D.* Energy Bands in Semiconductors, Wiley-Interscience, 1968].

поверхностей для любой энергии чуть выше минимальной и что величина $[\hbar^2(\partial^2\varepsilon/\partial k^2)^{-1}]$ равна m_l вдоль продольной оси эллипсоида и равна m_t вдоль каждой из его поперечных осей. Тогда суммарный вклад от $\mathfrak{N}$ эллипсоидов приводит к плотности состояний

$$g(\varepsilon)=\frac{\mathfrak{N}}{2\pi^2}\left(\frac{2}{\hbar^2}\right)^{3/2}(\varepsilon-\varepsilon_c)^{1/2}\left(m_l m_t^2\right)^{1/2}, \tag{4.82}$$

что дает для массы плотности состояний в формулах (4.1) и далее величину

$$m_c=\mathfrak{N}^{2/3}\left(m_l m_t^2\right)^{1/3}. \tag{4.83}$$

Сделанные выше замечания полностью применимы к зоне проводимости кремния, где $\mathfrak{N}$ минимумов располагаются *внутри* приведенной зоны. Для зоны проводимости германия минимумы энергии расположены в восьми точках на границе зоны Бриллюэна, и энергия, слегка превышающая минимальную, может быть представлена восемью *полуэллипсоидами* в приведенной зоне (рис. 4.50). Для каждой пары экстремумов, находящихся на расстоянии единичного вектора обратной решетки, два полуэллипсоида вместе дают один целый эллипсоид, т. е. в формулы (4.82) и (4.83) в данном конкретном случае следует подставлять $\mathfrak{N}=4$. В задаче 4.16 рассматриваются два других простых примера, использующие такой же принцип.

Прямые и непрямые собственные переходы

Библиография в конце этой главы содержит ссылки на источники, детально анализирующие зонную структуру большого числа полупроводников, и пространное повторение здесь этих сведений едва ли целесообразно. Из приведенных выше примеров мы можем видеть, что максимум валентной зоны ε_v может соответствовать, а может и не соответствовать положению в **k**-пространстве минимума энергии зоны проводимости ε_c. Когда *происходит* такое совпадение (в центре ЗБ или в другой точке **k**-пространства), то говорят что полупроводник имеет *прямую собственную щель* *. Так, и CdS, и GaAs имеют прямые щели,

* В полном соответствии с этим определением в отечественной литературе такие полупроводники называют прямозонными. — *Прим. перев.*

как и многие другие полупроводники, в том числе и те, для которых ε_v и ε_c расположены в одних и тех же точках **k**-пространства, но не совпадающих с центром ЗБ [49]. О полупроводнике, для которого максимальная энергия валентной зоны достигается *не в тех* точках **k**-пространства, где расположен минимум энергии зоны проводимости, говорят, что он обладает *непрямой* собственной щелью. Из рис. 4.49 ясно, что как кремний, так и германий, являются непрямозонными материалами; к этой категории относится и множество других твердых тел.

Сначала мы обсудим свойства прямозонных материалов, а позднее снова вернемся к непрямозонным. Смысл термина «прямая щель» состоит в том, что фотон с энергией $\hbar\omega = \varepsilon_i = (\varepsilon_c - \varepsilon_v)$ может возбудить электрон из верхней части заполненной зоны *прямо* в одно из состояний у самого дна зоны проводимости посредством перехода, который в **k**-пространстве будет вертикальным, как показано на рис. 4.51, *а*. Любой фотон с энергией, несколько превышающей ε_i, может вызвать прямой (вертикальный) переход, подобный тем, которые показаны штриховыми стрелками на рис. 4.51, *а*; свободные электрон и дырка, полученные таким образом, делят между собой избыточную энергию $(\hbar\omega - \varepsilon_i)$ так, чтобы электрон и дырка обладали одним и тем же значением **k**. Как можно легко убедиться (задача 4.17), для простых зон как электрон, так и дырка будут обладать волновым вектором

$$\mathbf{k} = (1/\hbar)\,[2m_r(\hbar\omega - \varepsilon_i)]^{1/2} \qquad (4.84)$$

относительно положения экстремума. Величину $m_r \equiv [m_c m_v (m_c + m_v)^{-1}]$ в выражении (4.84) называют *эффективной массой комбинированной плотности состояний.*

Предположим, что фотоны с энергией $\hbar\omega \geqslant \varepsilon_i$ падают на полупроводник с прямой запрещенной зоной. Величина поглощения этих фотонов зависит от *комбинированной плотности состояний* для состояний зоны проводимости и валентной зоны, разделенных по вертикали энергией $\hbar\omega$, и зависит также от степени взаимосвязи состояний этих зон. Эта последняя величина описывается квантовомеханически с помощью *межзонного матричного элемента.*

---

[49] Прямой (вертикальной) щелью обладают, например, соединения свинца PbS, PbSe и PbFe, хотя они и не имеют зонных максимумов и минимумов в центре ЗБ. Все эти твердые тела кристаллизуются в структуре каменной соли (для которой ЗБ тоже имеет форму, показанную на рис. 4.45), и результаты ШдГ-эффекта и циклотронного резонанса показывают, что максимумы валентной зоны и минимумы зоны проводимости совпадают в восьми точках $q = [1/2\,1/2\,1/2]$. Следовательно, набор поверхностей постоянной энергии для энергии, несколько превышающей ε_c или лежащей чуть ниже ε_v, в каждом из этих твердых тел будет выглядеть, как комбинация полуэллипсоидов, изображенная на рис. 4.50.

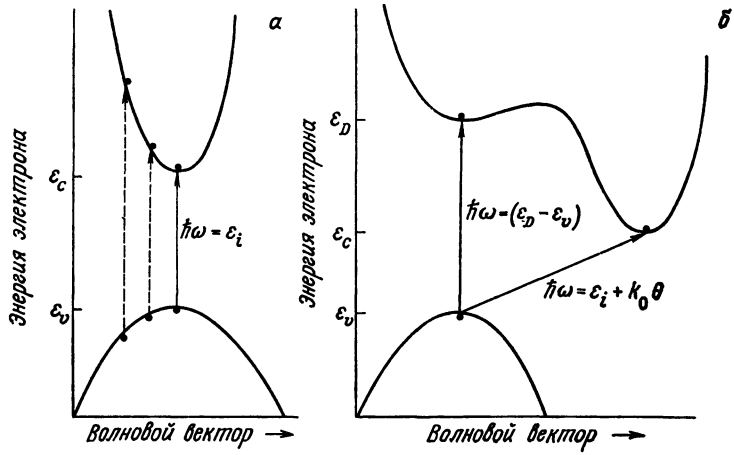


Рис. 4 51. Зависимость энергии электронов от волнового вектора для прямозонного (*а*) и непрямозонного (*б*) полупроводников. Прямой переход на такой диаграмме изображен вертикальной линией, что соответствует переходу, вызванному только фотоном (волновой вектор которого обычно составляет около 10^{-4} ширины зоны Бриллюэна). Вертикальный переход, обусловленный фотоном с энергией, большей ε_i, начинается от состояний, лежащих ниже ε_v, и идет вверх по вертикали до состояний с тем же $\mathbf{k}$, но выше ε_c, как показано штриховыми линиями в части *а* рисунка. Непрямые (невертикальные) переходы части *б* идут с участием фононов, обеспечивающих сохранение квазиимпульса. Любой прямой переход в таком твердом теле требует фотонов с более высокой энергией.

Если теперь переход из ε_v в ε_c не нарушает какое-либо спектроскопическое правило отбора [50], то его называют *разрешенным переходом*, и матричный элемент для энергии фотона $\hbar\omega_i = \varepsilon_i$ не равен нулю. Для такого случая естественно ожидать, что матричный элемент не будет зависеть от энергии перехода (по крайней мере для энергий фотона, не слишком превышающих ε_i) и что спектральная зависимость оптического поглоще-

[50] Оптические правила отбора для твердых тел являются менее строгими, чем для изолированных атомов, поскольку в случае твердого тела состояние редко бывает «чистым» *s*-, *p*- или *d*-состоянием. Однако некоторые переходы из самой верхней части заполненной зоны на самое дно пустой зоны формально оказывается запрещенными. Матричный элемент такого перехода становится исчезающе малым. Но даже в этом случае переход между такой парой зон обычно становится все «менее запрещенным» по мере того, как мы рассматриваем вертикально разделенные пары состояний, все более удаленные от точки $\mathbf{k}_0$, соответствующей положению экстремума. Во многих таких случаях матричный элемент, связывающий нижнее и верхнее состояния, пропорционален $|\mathbf{k}-\mathbf{k}_0|$ или $|\mathbf{k}-\mathbf{k}_0|^2$.

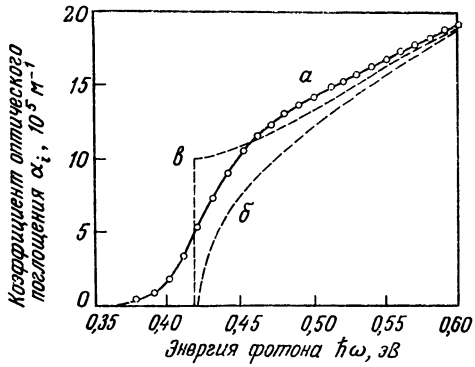


Рис. 4.52. Спектральная зависимость коэффициента оптического поглощения для полупроводника с прямой запрещенной зоной. Кривая *а* показывает экспериментальные данные для пленки сульфида свинца толщиной 1,33 мкм (осажденной эпитаксиально на монокристалл NaCl), полученные при комнатной температуре Пракашем [*Prakash. V.* Thesis, Harvard University, 1967]. Эти данные сравниваются с поведением, описываемым кривой *б* [уравнение (4.85)] для случая, когда полное экранирование исключает какое-либо взаимодействие между свободным электроном и свободной дыркой, созданными фотоном. Сравнение может быть сделано и с кривой *в*, которая соответствует зависимости (4.86) для модели оптического поглощения с электронно-дырочным взаимодействием Эллиота [52].

ния α_i будет определяться только спектральной зависимостью комбинированной плотности состояний:

$$\alpha_i = A(\hbar\omega - \varepsilon_i)^{1/2}, \quad \hbar\omega \geqslant \varepsilon_i. \tag{4.85}$$

Кривая *б* на рис. 4.52 построена в соответствии с формулой (4.85), и можно видеть, что это выражение находится хотя и не в идеальном, но удовлетворительном согласии с экспериментальными результатами кривой *а* на этом рисунке.

Одна из причин нарушения этого согласия состоит в том, что при выводе формулы (4.85) не учитывалось электронно-дырочное взаимодействие, которое может оказаться существенным, когда величина $(\hbar\omega - \varepsilon_i)$ не слишком велика. Поскольку электрон и дырка, образующие фотовозбужденную пару, создаются в одной и той же области реального пространства, они должны испытывать со стороны друг друга действие кулоновской силы [51], в какой-то мере экранированной другими зарядами, находящимися поблизости (что учитывается соответст-

[51] Кулоновское притяжение между электроном и дыркой может также привести к ряду *связанных* водородоподобных конфигураций, в которых электрон и дырка вращаются вокруг общего центра масс с квантованным моментом количества движения. Такое образование представляет собой *экситон* — твердотельный аналог «атома» позитрония. Экситоны, существование которых было в 1931 г. постулировано Френкелем, могут быть созданы при

вующей диэлектрической проницаемостью). Влияние электронно-дырочного взаимодействия перестает сказываться, когда систематическое движение и процессы рассеяния разделяют электрон и дырку, однако с его учетом спектральная зависимость поглощения, полученная Эллиотом [52], принимает вид

$$\alpha_i = B\{1 - \exp[-C(\hbar\omega - \varepsilon_i)^{-1/2}]\}^{-1}, \qquad \hbar\omega \geqslant \varepsilon_i, \qquad (4.86)$$

где B и C — постоянные.

Согласно формуле (4.86), α_i резко скачком возрастает при $\hbar\omega = \varepsilon_i$ и в конце концов сливается с ходом зависимости (4.85) для более высоких энергий, как показывает кривая *в* на рис. 4.52. Мы можем рассматривать выражение (4.85) как предельный случай выражения (4.86), соответствующий идеальному электростатическому экранированию кулоновского взаимодействия между электроном и дыркой.

Сравнение экспериментальных данных для PbS на рис. 4.52 с теоретическими зависимостями (4.85) и (4.86) показывает, что электронно-дырочное кулоновское взаимодействие в этом твердом теле частично экранировано большой диэлектрической проницаемостью. Кривая *а* на рис. 4.52 указывает также на особенность, которая обнаруживается у края поглощения любого прямозонного диэлектрика или полупроводника: постепенное увеличение α_i с увеличением энергии фотона в той области, где $\hbar\omega$ меньше, чем ширина щели, а не скачкообразное поглощение при $\hbar\omega = \varepsilon_i$. Как впервые было отмечено Мозером и Ур-

---

поглощении света. Фотоны, подходящие для этой цели, должны обладать энергией, несколько меньшей, чем ε_i. Насколько меньшей — исследуется в задаче 4.18. Теория экситонов обсуждается в небольшой книге: *Dexter D. L., Knox R. S.* Exitons, Wiley, Interscience, 1965.

[52] Модель оптического поглощения света, учитывающая неэкранированное взаимодействие электронов и дырок, впервые была описана Эллиотом [*Elliott R. J.*— Phys. Rev., **108**, 1384 (1957)]. Обзор моделей для «разрешенных» и «запрещенных» типов переходов с учетом и без учета поправки на электронно-дырочное взаимодействие дан Мак-Лином [*McLean T. P.* Progress in Semiconductors, v. 5, Wiley, 1960.].

[53] Мозер и Урбах [*Mosez F., Urbach F.*— Phys. Rev., **102**, 1519 (1956)] сообщили об экспоненциальной зависимости края поглощения ниже собственного порога для прямозонных диэлектриков AgCl и AgBr. Название *правило Урбаха* было сохранено для аналогичных экспоненциальных краев поглощения в других прямозонных твердых телах. Для некоторых твердых тел (включая те, которые первоначально исследовались Мозером и Урбахом) параметр, характеризующий крутизну спада, обратно пропорционален температуре. Для других твердых тел температура мало влияет на крутизну спада, указывая на то, что экспоненциальная зависимость края поглощения не всегда имеет одно и то же происхождение. Большие локальные электрические поля в диэлектрике или полупроводнике, обусловленные беспорядочным распределением доноров и акцепторов, в качестве причины экспоненциального края поглощения предложили Афромович М А, Редфильд Д. [Труды IX Межд конф. по физике полупроводников — Л : Наука, 1969, т. 1, с. 103.]

бахом[53] для галоидов серебра, первый участок кривой поглощения в прямозонном твердом теле возрастает по закону

$$\alpha_i \sim \exp[\beta(\hbar\omega - \varepsilon_i)], \qquad \hbar\omega < \varepsilon_i \tag{4.87}$$

даже если эта кривая согласуется с моделью Эллиота, описываемой формулой (4.86) для бо́льших энергий. Такому поведению, подчиняющемуся *правилу Урбаха* для энергий фотонов ниже той, которая должна была бы быть пороговой, было предложено множество объяснений, и, по-видимому, множество механизмов в различных классах полупроводников и диэлектриков в различной степени дают свой вклад в этот эффект.

Такой полупроводник, как германий или кремний, обладает *непрямой собственной щелью*, поскольку минимумы зоны проводимости расположены в областях k-пространства, удаленных от области максимума валентной зоны. Прямые оптические переходы показаны в правой части рис. 4.51 как возможный случай, когда используются фотоны с энергией, превосходящей (ε_D—ε_v), однако электронно-дырочные пары с минимальной энергией образования не могут быть получены в одноступенчатом (прямом) оптическом процессе. Однако, как было предложено Бардином и др.[54], *непрямой*, или *невертикальный*, переход может в таких случаях происходить через промежуточное виртуальное состояние таким образом, что поглощение фотона сопровождается либо рождением, либо уничтожением одного фотона из спектра колебаний решетки. Эта модель требует минимальной энергии фотона ($\varepsilon_i + k_0\theta$) для непрямого перехода, создающего электронно-дырочную пару и фотон с энергией $k_0\theta$ и квазиимпульсом $\hbar(k_c - k_v)$, причем этот процесс может происходить при любой температуре. Другой процесс, в котором при образовании пары поглощается фотон и *также* поглощается фонон с квазиимпульсом $\hbar(k_c - k_v)$, может происходить при меньшей пороговой энергии ($\varepsilon_i - k_0\theta$), однако этот механизм будет ослабляться при низких температурах, когда фононов, необходимых для поглощения, в кристалле в достаточном количестве уже не будет. Если для определения заселенности фононного состояния воспользоваться статистикой Бозе — Эйнштейна, то коэффициент оптического поглощения для непрямого перехода должен иметь вид

$$\alpha_i = \frac{A}{\hbar\omega}\left[\frac{(\hbar\omega - \varepsilon_i - k_0\theta)^2}{1 - \exp(-\theta/T)} + \frac{(\hbar\omega - \varepsilon_i + k_0\theta)^2}{\exp(\theta/T) - 1}\right], \tag{4.88}$$

где учтена только одна ветвь спектра колебаний. В общем случае фононы из нескольких ветвей спектра колебаний могут

[54] *Bardeen J., Blatt F. J., Hall L. H.* Photoconductivity Conference, Wiley, 1956. Обзор проблемы прямых и непрямых краев поглощения дан Мак-Лином (примечание 52).

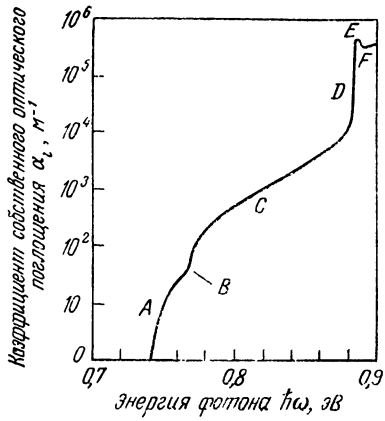


Рис. 4.53. Край собственного оптического поглощения для германия при $T=20$ К [*McLean T. P.* Progress in Semiconductors, v. 5, Wiley, 1960]. В области *A* возможны только непрямые межзонные переходы, в которых происходит одновременное поглощение фотона и фонона. Точка *B* соответствует начальной энергии $(\varepsilon_i+k_0\theta)$, с которой становится возможным непрямой процесс поглощения фотона, сопровождаемый образованием фотона, и этот процесс доминирует в области *C*. В точке *D* энергия фотона становится достаточной для прямых переходов в точке $\mathbf{k}=(000)$; показан экситонный пик (*E*), за которым следует область образования электронно-дырочных пар (*F*).

обеспечить необходимую величину $(k_c—k_v)$, причем каждый из них будет иметь свое характерное значение θ. Таким образом, коэффициент поглощения представляет собой сумму пар членов, соответствующих процессам рождения и аннигиляции фонона. При низких температурах каждый из членов, определяемых аннигиляцией фонона, резко падает по интенсивности, в то время как члены, определяемые рождением фонона, остаются.

Относительные интенсивности прямых и непрямых процессов оптического поглощения в полупроводнике проиллюстрированы экспериментальными данными для германия на рис. 4.53. Поскольку (непрямая) собственная энергетическая щель не намного меньше порога прямых оптических переходов в данном твердом теле, из этого рисунка легко видеть, что прямые процессы превосходят по интенсивности непрямые по крайней мере в 100 раз.

Когда первый край оптического поглощения в полупроводнике или диэлектрике соответствует *прямым* переходам, нам достаточно знать положение только одного минимума зоны проводимости: 1) чтобы определить положение других минимумов по их симметрии и 2) чтобы убедиться в том, что максимумы валентной зоны находятся в тех же точках ЗБ. С другой сто-

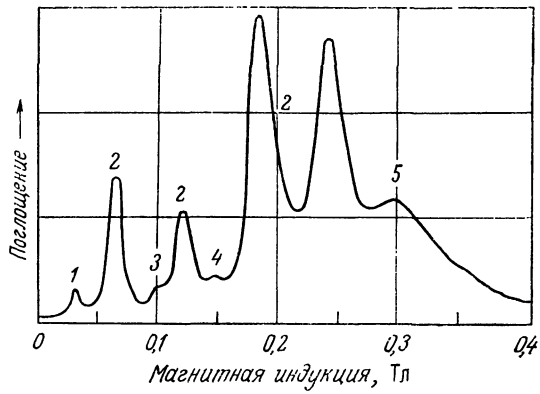


Рис. 4.54. Типичная кривая циклотронного резонанса для полупроводникового образца. Эта кривая была получена Декстером и др. [*Dexter R. N., Zeiger H. J., Lax B.* — Phys. Rev., **104**, 637 (1956)] для германия при 4 К в высокочастотном поле с частотой 23 ГГц. При этой температуре слабое освещение возбуждает и электроны, и дырки и сканирование приложением магнитного поля позволяет наблюдать различные возможные моды резонанса. *1* — легкие дырки; *2* — электроны, *3* — 3-я гармоника; *4* — 2-я гармоника; *5* — тяжелые дырки.

лика по сравнению с диаметром циклотронной орбиты, и сильные осцилляторные эффекты для геометрии Азбеля — Канера в них не возникают. На рис. 4.54 изображен типичный ход кривой циклотронного резонанса для образца германия с резонансами для фотовозбужденных электронов и фотовозбужденных дырок и со слабым указанием на осцилляторные эффекты. Измерения, сделанные для разных ориентаций в германии, свидетельствуют об эллипсоидальной форме поверхностей постоянной энергии электронов и об отклонении от сферичности поверхностей энергии для тяжелых и легких дырок.

4. Осцилляторное поведение магнетосопротивления, любая анизотропия которого указывает на анизотропию поверхностей постоянной энергии в **k**-пространстве. В разд. 3.5 был описан эффект осцилляций магнетосопротивления холодного вырожденного электронного газа в металле или сильнолегированном полупроводнике, называемый *эффектом Шубникова — де Гааза* (ШдГ-эффект). Магнитное поле перестраивает плотность состояний в серию магнитных подзон (как это описано при объяснении рис. 3.59—3.62) с зазором между уровнями Ландау $\hbar\omega_c$. Если мы хотим наблюдать осцилляции ШдГ, величина n_0 должна быть достаточно велика, а T достаточно мало, чтобы поверхность Ферми была четко выделена, так что $k_0T < \hbar\omega_c$. Тогда магнетосопротивление испытывает осцилляцию всякий

раз, когда изменение магнитного поля заставляет уровень Ландау проходить через энергию Ферми. В качестве примера на рис. 3.66 и в задаче 3.35 была выбрана простейшая ситуация сферической поверхности Ферми с центром при $\mathbf{k}=(000)$. Для такой ситуации период ШдГ-осцилляций P (который обратно пропорционален сечению поверхности Ферми S_F) не зависит от ориентации. Однако если электроны (или дырки) сгруппированы в ряде эллипсоидальных или других несферических поверхностей внутри зоны типа показанных на рис. 3.50 и 4.50, то анизотропия периода ШдГ будет давать информацию о степени вытянутости каждого эллипсоида и угловом размещении эллипсоидов.

5. Осцилляции магнетосопротивления могут наблюдаться в полупроводнике в совершенно других условиях. Это явление *магнетофононного резонанса,* предсказанное Гуревичем и Фирсовым (1961)* и экспериментально в значительной мере впервые наблюдавшееся в работе Стредлинга [56]. Магнетофононный резонанс наблюдается в слаболегированном полупроводнике, для которого газ электронов проводимости невырожден и требует присутствия длинноволновых продольных оптических фононов с энергией $\hbar\omega_{LO}$ и малым волновым вектором. Последнее требование означает, что температура не должна быть настолько низкой, чтобы полностью исключалось наличие оптических фононов, однако температура не должна быть и настолько высока, чтобы плотность состояний, изображенная на рис. 3.62, оказалась термически размытой. Уровни Ландау, изображенные на рис. 3.62, будут сохраняться, если $\omega_c\tau_m \gg 1$. Роль *фононов* в магнетофононном резонансе состоит в рассеянии электронов с самого нижнего уровня Ландау на более высокие, происходящем для набора магнитных полей, удовлетворяющих условию

$$\hbar\omega_{LO} = N\hbar\omega_c = (N\hbar eB/m^*), \qquad N = 1,2,3 \ldots . \tag{4.89}$$

Таким образом, магнетосопротивление имеет компоненту, периодичную по обратному полю [57]. Стредлинг указывает на то, что осцилляционный член может быть легко обнаружен путем

---

* *Гуревич В. Л., Фирсов Ю. А.* — ЖЭТФ, 1961, т. 40, с. 199. — *Прим. ред.*

[56] *Stradling R. A., Wood R. A.*— J. Phys. C, 1, 1711 (1968); 3, L94 (1970). Подробный обзор этого вопроса дан Харпером [*Harper P. G., Hodby J. W., Stradling R. A.*— Repts. Progr. Phys., 36, 1 (1973)].

[57] Периодическая зависимость магнетосопротивления от $(1/B)$ выглядит как ШдГ-резонанс, хотя явление, лежащее в основе магнетофононного резонанса, совершенно другое. Магнетофононный резонанс более сходен с циклотронным резонансом, в котором источником возбуждения служат не высокочастотные фотоны, а фононы оптической моды.

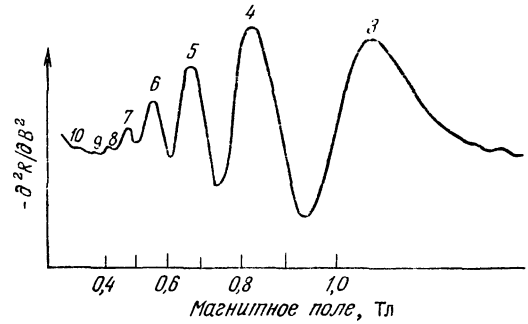


Рис. 4.55. Вторая производная сопротивления (d^2R/dB^2) как функция индукции магнитного поля B для электронов в зоне проводимости InSb при $T=77$ К. Осцилляции указывают на магнетофононный резонанс, в данном случае для электронов в изотропной зоне [*Stradling R. A.*— Phys., E5, 736 (1972)]. Стредлинг утверждает, что эти результаты были получены выпускником Оксфордского университета на последнем году лабораторного курса. Анализ результатов предлагается в задаче 4.20.

построения второй производной сопротивления образца (d^2R/dB^2) как функции B или ($1/B$). Именно так обработаны экспериментальные результаты на рис. 4.55. Они представляют собой экспериментальные данные для изотропной зоны проводимости InSb (их предлагается проанализировать в задаче 4.20). Однако можно обнаружить и проанализировать более сложные и зависящие от ориентации осцилляции[58], когда исследуемая зона имеет сложные анизотропные поверхности энергии.

6. Эффект де Гааза — ван Альфена (дГвА), состоящий в осцилляциях магнитной восприимчивости, может быть использован в полупроводнике, как и в металле, для определения размера и формы поверхности Ферми для ситуации, соответствующей вырождению при низкой температуре. Для большинства полупроводников этот эффект довольно слаб, и поэтому он с бо́льшим успехом используется в металлах и *полуметаллах* — материалах, в которых очень слабое перекрытие зон приводит к образованию областей с равными количествами электронов и дырок. В магнии такое перекрытие (см. рис. 3.40) настолько велико, что этот элемент является ярко выраженным металлом. Гораздо меньшее перекрытие наблюдается для элементов VB группы периодической системы. Число электронов (и дырок) на атом изменяется в пределах от $4{,}6 \cdot 10^{-3}$ в мышьяке до всего лишь $1{,}1 \cdot 10^{-5}$ в висмуте. Хотя это и го-

[58] См., например, *Stradling R. A. et al.* Proc. X Intl. Semiconductor Conf., A. E. C., 1970, p. 369.

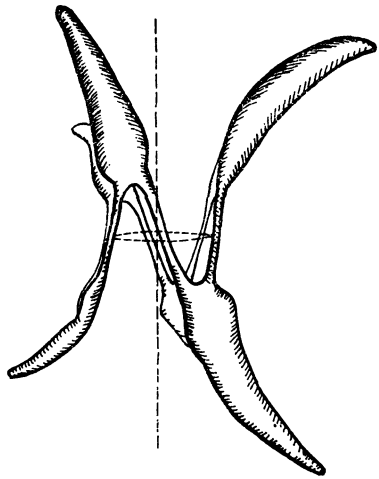


Рис. 4.56. Изображение в перспективе «коронообразной» поверхности Ферми для дырок в мышьяке, основанное на расчетах методом псевдопотенциала Лина и Фаликова [59], использующим экспериментальные данные измерений эффекта дГвА. Мышьяк обладает кристаллической структурой, которая в основе своей является гранецентрированной кубической, деформированной вдоль диагонали куба таким образом, чтобы получилась тригональная решетка. Изображенная фигура имеет выделенную ось, представляющую собой вращательную ось симметрии третьего порядка [*Lomer W. M., Gardner W. E.* Progress in Materials Science, **14**, 143 (1969)].

раздо меньше, чем в нормальном металле, концентрации носителей еще достаточно велики, чтобы вызвать эффект дГвА, и с успехом использовать для исследования размера и анизотропии поверхностей Ферми электронов и дырок в полуметаллах. На рис. 4.56 показана поверхность Ферми, предсказываемая для дырок в мышьяке зонным расчетом, основанным на методе псевдопотенциала [59] с подгоночными параметрами, удовлетворяющими экспериментальным данным эффекта дГвА. Предполагается, что для сурьмы эти тонкие связывающие области исчезают, оставляя шесть изолированных областей, которые еще сильно отклоняются от эллипсоидальной формы.

Исследования полупроводника, которые дают комбинированные свойства двух зон,— это, конечно, анализ оптического поглощения, подробно описанный в последнем разделе, и ряд других методов, основанных на взаимодействии фотонов с твердым телом. Поскольку полупроводник непрозрачен для фотонов с энергиями большими, чем его первая прямая щель, переходы из заполненных зон в пустые состояния со значительно бо́льшими энергиями нельзя наблюдать по оптическому

[59] *Lin P. J., Falicov L. M.*— Phys. Rev., **142**, 441 (1966).

пропусканию (если не использовать очень тонкие образцы). Спектральный диапазон более высоких энергий часто изучают с помощью *коэффициента оптического отражения*, производная которого может сильно изменяться для энергии, соответствующей сингулярности Ван Хова в комбинированной плотности состояний [60]. В разд. 3.3 и 3.5 мы уже указывали, что фотоэмиссия представляет собой еще один ценный аналитический метод для исследования таких переходов, обусловленных фотонами больших энергий, для которых конечное состояние электрона выше, чем поверхностный барьер работы выхода [61].

Комбинация деталей тонкой структуры состояний нижней зоны проводимости и верхней валентной зоны может быть получена путем измерения *магнитопоглощения*, т. е. изменения коэффициента краевого оптического поглощения в магнитном поле. Такие измерения были с большим искусством выполнены Лэксом с сотрудниками [62].

Процесс электронно-дырочной рекомбинации (о котором мы раскажем более подробно в следующем разделе) может также быть достаточно информативным в отношении того, является ли прямым или непрямым собственный переход. Эффективная *люминесценция* (излучательная рекомбинация) совершенно невозможна, если переход не является прямым. Процесс, известный как *оже-рекомбинация*, или *рекомбинация с участием третьего тела*, также очень маловероятен, если экстремумы зоны проводимости и валентной зоны не совпадают.

4.4. Явления, обусловленные избыточными носителями

Многие интересные явления происходят в полупроводнике в том случае, когда распределение электронов по разрешенным состояниям существенно отличается от равновесного. С такими

---

[60] Имеется много современных статей на эту тему, однако для начального знакомства с вопросом можно рекомендовать статью Филиппа и Эренрайха [*Philipp H. R., Ehrenreich H.*— Phys. Rev., **129**, 1550 (1963)], которая содержит данные по коэффициету отражения для ряда полупроводников и в которой обсуждается анализ информации о вещественной и мнимой частях диэлектрической проницаемости методами Крамерса — Кронига. Оптические свойства могут быть связаны с определенными заполненными и определенными пустыми зонами.

[61] Обзор фотоэмиссионных методов, применяемых для неметаллических твердых сел см. в работе: *Spicer W. E., Eden R. C.* Proceedings 9th International Semiconductor Conference, Moscow, Academy of Sciences, 1968, p. 65

[62] См, например, *Lax B.* Proceedings 9th International Semiconductor Conference, Moscow, Academy of Sciences, 1968, p. 253.

отклонениями от термодинамического равновесия мы сталкиваемся в любом биполярном полупроводниковом приборе, а во многих случаях и в однородных полупроводниковых образцах.

Описание неравновесных явлений, которое дано в этом разделе, весьма кратко. Гораздо более обширные сведения могут быть получены из литературы, приведенной в конце главы, включающей монографии Рывкина, Бьюба и автора этих строк.

Детальное равновесие

Рассмотрим полупроводник, изображенный на рис. 4.57, содержащий p дырок/м^3 в верхних областях валентной зоны и n электронов/м^3 в нижней части зоны проводимости. Размещение электронов по состояниям с различной энергией является не-

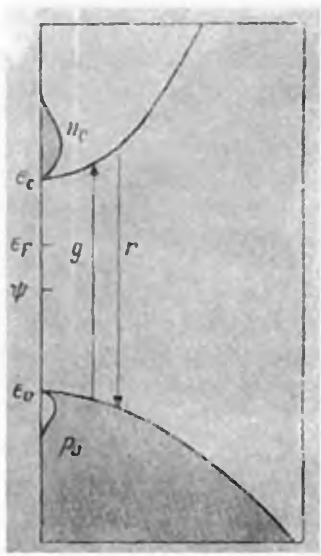


Рис. 4.57. В условиях термодинамического равновесия концентрации электронов и дырок n_0 и p_0 могут быть выражены через единственный уровень Ферми ε_F с помощью уравнений (4.22) или (4.90) и (4.91). Равновесные значения n_0 и p_0 обусловлены процессами генерации, однако совершенно не зависят от равных по величине, но противоположных по направлению скоростей g и r. Также ни одна из величин n_0, p_0 и ε_F не зависит от *природы* генерационно-рекомбинационного процесса. Природа процесса рождение — аннигиляция и его интенсивность приобретают большое значение при нарушении термодинамического равновесия.

прерывным и динамическим, однако мы знаем, что в *тепловом равновесии*

$$n = n_0 = N_c \exp\left(\frac{\varepsilon_F - \varepsilon_c}{k_0 T}\right), \tag{4.90}$$

$$p = p_0 = N_v \exp\left(\frac{\varepsilon_v - \varepsilon_F}{k_0 T}\right). \tag{4.91}$$

независимо от природы и эффективности процессов превращения энергии. Так может быть только в случае, если тепловое равновесие представляет собой такую ситуацию, для которой скорость генерации в точности равна скорости рекомбинации для любого процесса превращения энергии и для любого энергетического интервала. Тогда *полная* скорость генерации g равна полной скорости рекомбинации r. Суть нашего утверждения состоит в том, что для каждого процесса существует обратный ему процесс, обладающий в равновесных условиях *такой же* вероятностью. Это одна из формулировок *принципа детального равновесия.*

Так, в качестве конкретного примера фотоны излучения черного тела, окружающего наш равновесный полупроводник, поглощаются с определенной скоростью как в *целом*, так и *детально* при образовании электронно-дырочных пар. Эта скорость равна той скорости, с которой полупроводник в равновесии порождает фотоны в излучательном рекомбинационном процессе аннигиляции пары электрон — дырка. Более того, излучательная рекомбинация должна генерировать фотоны с таким же спектром, как у поглощенных фотонов, если выполняется второй закон термодинамики.

Неравновесная статистика и квазиуровни Ферми

Предположим, что концентрации свободных носителей на рис. 4.57 изменены внешним воздействием таким образом, что $np \neq n_0 p_0$. Тогда, стремясь восстановить равновесную ситуацию, генерация будет преобладать над рекомбинацией, если $np < n_0 p_0$. Аналогичным образом рекомбинация будет преобладать над генерацией, если $np > n_0 p_0$, что является более распространенным видом отклонения от равновесия.

В любом случае *полные* концентрации n и p можно все еще описывать выражениями

$$n = N_c \exp\left(\frac{\varphi_n - \varepsilon_c}{k_0 T}\right), \tag{4.92}$$

$$p = N_v \exp\left(\frac{\varepsilon_v - \varphi_p}{k_0 T}\right), \tag{4.93}$$

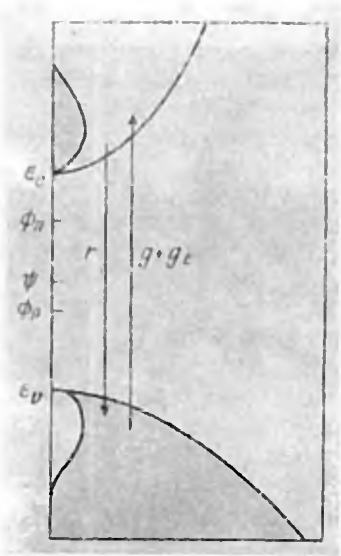


Рис 4.58. Естественные скорости генерации и рекомбинации перестают уравновешивать друг друга как в целом, так и детально, когда тепловое равновесие нарушается генерацией, вызванной внешним воздействием, или инжекцией носителей из области другого состава. Полная концентрация электронов проводимости может теперь быть описана уравнением (4.92) или (4.94), которые переходят в (4.22) и (4.90), когда восстанавливается равновесие.

где φ_n и φ_p являются параметрами нормировки, имеющими размерность энергии, которые (следуя Шокли) мы будем называть *квазиуровнями Ферми* для электронов и дырок соответственно. Величины φ_n и φ_p сливаются в энергию Ферми ε_F при термодинамическом равновесии, но $\varphi_n > \varepsilon_F > \varphi_p$, если $np > n_0 p_0$, как показано на рис. 4.58.

Следует заметить, что выражения (4.92) и (4.93) описывают *полные* концентрации электронов и дырок в неравновесной ситуации, однако ничего не говорят нам о распределении n электронов или p дырок по имеющимся состояниям. Очень важно знать об этом предостерегающем замечании, поскольку распределение электронов или дырок в пространстве скоростей может быть очень далеким от вида Максвелла — Больцмана (3.4) и (3.5), если большое число избыточных носителей попадает в зону в течение короткого времени с совершенно другим распределением скоростей. Фононы и дефекты решетки способствуют термализации избыточных носителей, поэтому квазиуровень Ферми может характеризовать и полное число носителей *и* распределение всех носителей в зоне по скоростям

при условии, что *время жизни носителей* велико по сравнению со средним временем свободного пробега между двумя актами рассеяния τ_m. Время жизни — это среднее время пребывания избыточного носителя в зоне, и мы кратко остановимся на том, насколько велико или мало оно может быть для конкретных процессов превращения энергии.

С использованием формулы (4.16) концентрации электронов и дырок вдали от равновесия могут быть записаны через собственную энергию Ферми ψ и собственную концентрацию n_i:

$$n = n_i \exp\left(\frac{\varphi_n - \psi}{k_0 T}\right); \quad (\varphi_n - \psi) = k_0 T \ln(n/n_i), \tag{4.94}$$

$$p = n_i \exp\left(\frac{\psi - \varphi_p}{k_0 T}\right); \quad (\psi - \varphi_p) = k_0 T \ln(p/n_i). \tag{4.95}$$

Эти выражения являются обобщением формул (4.22) для неравновесных условий. Таким образом,

$$np = n_i^2 \exp\left(\frac{\varphi_n - \varphi_p}{k_0 T}\right), \tag{4.96}$$

что означает

$$e^{-1}(\varphi_n - \varphi_p) = (k_0 T/e) \ln(np/n_i^2) \tag{4.97}$$

Величина в левой части выражения (4.97) может быть истолкована как электродвижущая сила, приводящая к восстановлению равновесия. Ниже мы увидим, что величины $\nabla\varphi_n$ и $\nabla\varphi_p$ связаны с полной плотностью тока в полупроводнике, создаваемой соответственно электронами и дырками и обусловленной двумя видами тока: дрейфовым движением под действием поля и диффузионным движением.

Рассмотрим теперь, как отклонение от равновесия смещает величины φ_n и φ_p по отношению к положению собственного уровня Ферми ψ и равновесному положению ε_F. Сначала предположим, что образец в равновесии является собственным, так что $\varepsilon_F = \psi$. При интенсивном освещении или каком-то другом неравновесном воздействии, создающем избыточные электронно-дырочные пары, возможна ситуация, при которой p и n будут фактически продолжать оставаться равными между собой [63]. Тогда φ_n будет настолько выше ψ, насколько φ_p опустится ниже этого уровня, являющегося индикатором собственного состояния: $(\varphi_n - \psi) = (\psi - \varphi_p) = k_0 T \ln(n/n_i)$.

[63] Такая ситуация возможна, но не обязательна. Большое количество уровней дефектов, распределенных в запрещенной зоне, может быть таким, что $n_0 = p_0 = n_i$ в равновесии, хотя эти дефекты захватывают электроны с большей вероятностью, чем дырки. При сильном освещении эти дефекты быстро захватят больше избыточных электронов, чем дырок, так что величина $n_e = (n - n_0)$ станет меньше $p_e = (p - p_0)$.

В качестве противоположного случая рассмотрим теперь образец сильно выраженного n-типа, в котором $n_0 \gg p_0$, так что ε_F лежит намного k_0T выше, чем ψ. При освещении образца *относительное* изменение оказывается гораздо более значительным для *неосновных носителей* (дырок), чем для *основных носителей* (электронов). Поэтому φ_n будет перемещаться очень слабо, пока избыточная концентрация электронов не приблизится к величине n_0. Однако φ_p будет двигаться быстро для любой избыточной концентрации дырок, располагаясь ниже ψ, как только $p = (p_0 + p_e)$ станет больше n_i, причем φ_p окажется ниже ψ настолько же, насколько ψ_n превышает ψ в тот момент, когда концентрация избыточных пар составляет n_0. Задача 4.21 рассматривает простой пример этого принципа, который имеет прямое отношение к величине тока через смещенный p—n-переход в полупроводниковом диоде или транзисторе.

Когда концентрации электронов и дырок превышают свои равновесные значения, естественные процессы рекомбинации ускоряются, стремясь восстановить равновесную ситуацию. Они ускоряются именно потому, что в кристалле оказывается больше электронов и дырок, способных к рекомбинации. Предположим теперь, что в нашем распоряжении имеется некое средство, позволяющее сделать n_0 и p_0 *меньше*, чем в равновесии. В этих условиях *эксклюзии*, или *обеднения носителями*[64], рекомбинация замедляется, однако процесс генерации идет в таком же темпе, как обычно, стремясь восстановить нормальные концентрации носителей. В полупроводниковом образце, в котором достигнуто обеднение носителями, квазиуровень Ферми для дырок φ_p расположен *выше*, чем квазиуровень Ферми для электронов, как этого требуют соотношения (4.92)—(4.95).

Любой объемный заряд, который устанавливается при обеднении носителями, будет сопровождаться возникновением электростатического потенциала (см. задачу 4.22), как это следует из уравнения Пуассона. Поэтому уменьшение концентрации дырок и концентрации электронов должно быть одинаковым,

---

[64] Об этом процессе впервые сообщил Гибсон [*Gibson A. F.*— Physica, **20**, 1058 (1954)], назвав его «экстракцией носителей», однако в последующих публикациях предпочтение было отдано терминам «эксклюзия» и «обеднение». Теория обеднения носителями обсуждается Ридом [*Read W. T.*— Bell. Syst. Tech. J., **35**, 1239 (1956)]. Обеднение в объемном полупроводниковом образце требует таких контактов, которые экстрагируют, а не инжектируют носители. Тогда импульс напряжения приводит к экстракции, которая охватывает всю длину образца, подобно ударной волне.

[Четкое разграничение понятий «экстракции» и «эксклюзии» — двух различных по физическому смыслу видов обеднения носителями — содержится в монографии: *Бонч-Бруевич В. Л., Калашников С. Г.* Физика полупроводников. — М.: Наука, 1977, с. 271—273. — *Прим. перев.*].

если одновременно не происходит перестройка заряда, связанного на примесных центрах. Это означает прежде всего, что условия обеднения легче всего реализовать на материале, близком к собственному.

Центральная область p—n-перехода испытывает глубокое обеднение, если переход смещен в обратном (т. е. запорном, или высокоомном) направлении.

Уравнения непрерывности и времена жизни носителей

Предположим, что электроны в зоне проводимости образуются со скоростью g_E в результате какого-то внешнего воздействия. Кроме того, существуют скорости естественной генерации g и рекомбинации r, обусловленные различными механизмами превращения энергии. Тогда, если мы для определенности рассмотрим концентрацию электронов в зоне проводимости и ее зависимость от времени, можно записать *уравнение непрерывности* для электронов в виде

$$\frac{\partial n}{\partial t} = (g - r) + g_E + e^{-1}\nabla \cdot \mathbf{I}_n. \tag{4.98}$$

Здесь последний член отражает тенденцию избыточных электронов куда-нибудь перемещаться, если их распределение не однородно.

Далее разумно предположить, что величина $(r - g)$ в грубом приближении должна быть пропорциональна концентрации избыточных электронов $n_e \equiv (n - n_0)$, если концентрация избыточных носителей не слишком велика. Удобно определить *время жизни электронов* τ_n с помощью соотношения

$$\tau_n = \frac{n_e}{(r - g)}, \tag{4.99}$$

отдавая себе при этом отчет, что время жизни, определенное таким образом, по-видимому, не будет полностью независимым от n_e. Следовательно, для полупроводника, в котором зависимостями от координат можно пренебречь, уравнение непрерывности сводится к простому виду

$$\frac{dn_e}{dt} = \frac{dn}{dt} = g_E - n_e/\tau_n, \tag{4.100}$$

а для случая стационарного возбуждения время жизни избыточных электронов есть просто (n_e/g_E). Во время переходного процесса затухания, который следует за периодом генерации, стимулированной внешним воздействием, постоянная времени

затухания равна $\tau_n = -n_e(dt/dn)$ при условии, что первоначальная генерация могла считаться пространственно однородной [65].

Уравнение непрерывности для дырок может быть составлено аналогичным образом:

$$\frac{\partial p}{\partial t} = (g' - r') + g'_E - e^{-1}\nabla \cdot \mathbf{I}_p. \tag{4.101}$$

В пространственно однородной ситуации это уравнение становится обыкновенным дифференциальным уравнением

$$\frac{dp_e}{dt} = \frac{dp}{dt} = g'_E - p_e/\tau_p, \tag{4.102}$$

записанным с помощью времени жизни дырок $\tau_p = p_e(r' - g')^{-1}$. Электроны и дырки имеют равные времена жизни, если как генерация, так и рекомбинация обусловлены исключительно процессами зона — зона: $g = g'$, $r = r'$, $n_e = p_e$. Времена жизни дырок и электронов различны, если только один тип носителей выводится из состояния равновесия, как это происходит, если полупроводник освещается фотонами с малой энергией, которые возбуждают электроны из примесных состояний. Времена жизни электронов и дырок, вообще говоря, различны и в том случае, когда генерация электронов и дырок идет по схеме зона — зона, но рекомбинация происходит главным образом по механизму Шокли — Рида, который состоит в последовательном захвате свободных электронов и свободных дырок локализованными примесными центрами.

Динамика примесной генерации — рекомбинации [66]

Прежде чем обсуждать процессы, идущие по схеме зона — зона, рассмотрим неравновесную динамику концентрации носителей одного типа, например избыточной концентрации n_e в зоне

[65] Конечно, представление о полупроводнике, в котором избыточная генерация совершенно однородна, есть идеализация, к которой можно приблизиться с большим или меньшим успехом. Для полупроводникового образца конечных размеров неизбежно будет существовать рекомбинация не только в объеме, но и на поверхности, и в этом случае пространственное распределение избыточных электронов может быть описано суммой фурье-компонент, каждая из которых в различной степени испытывает на себе воздействие поверхности и имеет отличное от других эффективное время жизни. Когда генерация, обусловленная внешним воздействием, пространственно неоднородна (как это в определенной степени и должно быть для любого вида ионизующего излучения или инжектирующего контакта), становятся существенными компоненты Фурье более высоких порядков с короткими временами жизни, что автоматически приводит к *уменьшению* неоднородности распределения избыточных электронов. Литература, освещающая анализ разложений Фурье для пространственно неоднородных ситуаций, обсуждается в гл. 10 «Статистики полупроводников»

[66] Краткое обсуждение этого вопроса приводится здесь потому, что с ним может быть связано решение нескольких полезных задач (4.23—4.25)

проводимости. Модель, которая должна быть использована, основана на рис. 4.18 для частично компенсированного полупроводника n-типа с донорными центрами, обладающими энергией ионизации ε_d. Поскольку состояния зоны проводимости слабо заполнены, скорость рекомбинации должна быть пропорциональна полной концентрации свободных электронов $n = (n_0 + n_e)$ и числу $\beta^{-1}(n + N_a)$ имеющихся состояний на ионизованных донорах. Мы можем записать эту скорость в виде

$$r = (\bar{v}\bar{\sigma}/\beta)\, n\,(N_a + n), \tag{4.103}$$

где $\bar{v}$ есть средняя скорость свободного электрона (усредненная по тепловому распределению Больцмана), а $\bar{\sigma}$ — аналогичным образом усредненное сечение захвата, которым обладает пустой донор для любого свободного электрона.

Скорость противоположного процесса генерации определяется числом электронов на донорах $N_{dn} = (N_d - N_a - n)$ и тепловым окружением. При условии что зона проводимости не вырождена, эта скорость не будет зависеть от незначительной степени заполнения зоны. Таким образом, мы можем записать

$$g = A\,(N_d - N_a - n)\exp(-\varepsilon_d/k_0T), \tag{4.104}$$

поскольку влияние теплового окружения должно определяться отношением ε_d к k_0T. Величина A будет зависеть от плотности состояний зоны проводимости, но *не* от N_d, N_a или n.

Поскольку в равновесии $g = r$, из выражений (4.103) и (4.104) следует, что

$$\frac{n_0\,(n_0 + N_a)}{N_d - N_a - n_0} = (A\beta/\bar{v}\bar{\sigma})\exp(-\varepsilon_d/k_0T) = n_1. \tag{4.105}$$

Так соображения динамики генерации и рекомбинации вновь приводят нас к уравнению, полученному ранее под номером (4.36), которому мы тогда дали название *уравнения действующих масс*. Теперь мы можем видеть происхождение такого названия. Величина $(A/\bar{v}\bar{\sigma})$ точно равна N_c. Поскольку обе части равенства (4.105) имеют размерность (объем)$^{-1}$, удобно записать правую часть как «концентрацию действующих масс» n_1. Выраженная через n_1 скорость генерации вдали от равновесия есть

$$g = (v\sigma/\beta)\, n_1\,(N_d - N_a - n). \tag{4.106}$$

Подстановка (4.103) и (4.106) в (4.99) [с помощью (4.105)] приводит (см. задачу 4.23) к выражению для времени жизни электрона

$$\tau_n = (\beta/\bar{v}\bar{\sigma})\,[N_a + 2n_0 + n_1 + n_e]^{-1}, \tag{4.107}$$

умеренной сложности. Этот раздел можно опустить без серьезного ущерба для связности изложения.

которое следует подставлять в уравнение непрерывности (4.100).

Наиболее сильно зависящая от температуры величина в выражении (4.107) — это обычно n_1, хотя и имеется зависимость вида $T^{1/2}$ у средней скорости $\bar{v}$, а для некоторых типов центров и сечение захвата $\bar{\sigma}$ может быть чувствительной функцией температуры.

Изменяется ли величина $\bar{\sigma}$ с температурой и насколько сильно, зависит от физической природы процесса захвата электрона. Как выделяется энергия при рекомбинации? Существуют три возможности и комбинации из них:

а. Энергия может освобождаться в виде фотона, по-видимому, это возможно и с участием фононов.
б. Энергия может быть передана как дополнительная кинетическая энергия другому свободному электрону.
в. Может быть образовано несколько фононов как одновременно, так и в определенной временной последовательности.

Примесная люминесценция представляет собой процесс, при котором энергия рекомбинации выделяется в виде фотона, возможно, с одновременным образованием одного или двух фононов. Эффективность этого процесса обычно не зависит от температуры. Для большинства примесных центров эффективность такой *излучательной рекомбинации* очень мала и доминирующим является процесс в. Существует очень большой практический интерес к таким полупроводниковым системам с примесными центрами, в которых примесная люминесценция *может* быть эффективной. В некоторых таких интенсивно изучаемых системах (некоторые соединения III—V и II — VI) происходят люминесцентные переходы из зоны проводимости на состояния примесных центров, расположенные гораздо ниже середины запрещенной зоны. Для светящихся люминесцентных дисплеев (таких, как в цифровых электроизмерительных приборах, электронных калькуляторах и т. п.) могут быть использованы излучательные переходы как зона — зона, так и зона — центр.

Процесс, при котором энергия рекомбинации проявляется как дополнительная кинетическая энергия другого электрона, представляет собой примесный вариант *оже-рекомбинации*, названной так по аналогичному процессу в атомной физике. Оже-рекомбинация — это процесс, обратный ударной ионизации, при которой быстрый электрон теряет большую часть своей энергии на ионизацию донора. В полупроводнике, в котором оже-рекомбинация действительно преобладает, выраже-

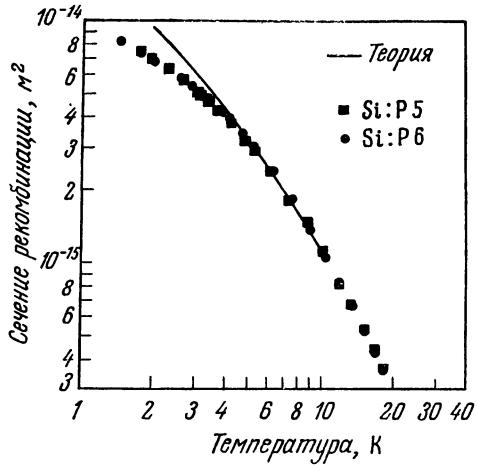


Рис. 4.59. Сечение захвата для ионизованных доноров фосфора в кремнии в зависимости от температуры. Предполагается, что захват сопровождается эмиссией каскада фононов в основных чертах в соответствии с моделью Лэкса [67] [*Norton P., Braggins T., Levinstein H.*— Phys. Rev. Lett., **30**, 489 (1973)].

ние для скорости рекомбинации r в формуле (4.103) должно быть заменено на $r=Bn^2(n+N_a)$, поскольку «сечение захвата» само пропорционально концентрации электронов. В задаче 4.24 рассматриваются вытекающие отсюда изменения в выражении для времени жизни. Поскольку оже-скорость изменяется как n^2, этим процессом в полупроводнике обычно пренебрегают, если величина n мала.

Во многих процессах захвата центрами возмущение решетки, создаваемое присутствием в ней центра, обеспечивает легкость взаимодействия с колебаниями решетки. Иногда *многофононная эмиссия* сразу забирает всю энергию рекомбинации. В других случаях происходит *фононный каскад* [67], при котором фононы испускаются один за другим по мере того, как электрон движется по спектру связанных возбужденных состояний центра, в конце концов достигая основного состояния. Этот процесс имеет сильную температурную зависимость, поскольку при высоких температурах электрон обладает большой вероятностью покинуть мелкое возбужденное состояние вместо того, чтобы продолжать двигаться вниз по энергии. На рис. 4.59 представлена температурная зависимость σ для донора

[67] Процесс фононного каскада был впервые проанализирован Лэксом [*Lax M.*— Phys. Rev., **119**, 1502 (1960)]. Имеется много последующих усовершенствований этой модели.

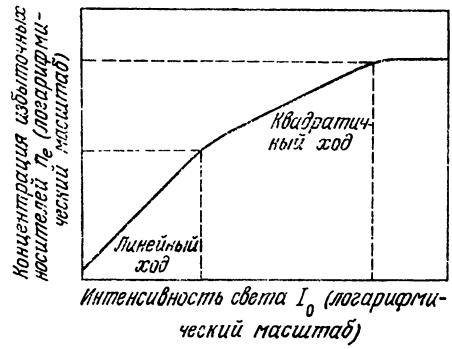


Рис. 4.60. Изменение стационарной избыточной концентрации электронов в зависимости от интенсивности освещения для примесного полупроводника, в котором преобладает либо фононная рекомбинация, либо спонтанная излучательная рекомбинация. Ситуация становится несколько более сложной в случае оже-рекомбинации. В задаче 4.23 предлагается получить аналитические выражения для различных участков кривой, а также для интенсивностей света, соответствующих концу линейной области и началу насыщения.

фосфора в кремнии; заметим здесь, что сечение при низких температурах огромно по сравнению с размерами атома. Захвату, показанному на рис. 4.59, способствует кулоновское притяжение ионизованным донором электрона проводимости; гораздо меньшие сечения наблюдаются для захвата электронов и дырок нейтральными центрами.

Динамика генерации и рекомбинации для переходов с участием фононов такая же, как с участием фотонов. Так (в отличие от оже-рекомбинации), зависимость τ_n от n_e определяется выражением (4.107).

Предположим, что мы используем стационарное освещение, чтобы создать стационарное неравновесное состояние в примесном полупроводнике n-типа. Пусть имеется постоянный поток I_0 фотонов на 1 м^2 в 1 с, и пусть излучательная генерация определена потоком I_0 и сечением фотоионизации σ_I каждого нейтрального донора. Для простоты будем считать скорость генерации однородной в пространстве, что справедливо, если толщина образца мала по сравнению с величиной $(N_d - N_a - n)\sigma_I$. В этих условиях существует монотонная зависимость между I_0 и результирующим значением n_e для стационарного состояния, как это изображено на рис. 4.60. Вторая часть задачи 4.23 касается идентификации областей, в которых n_e изменяется с I_0 как I_0, затем как $I_0^{1/2}$ и в конце выходит на насыщение. Форма кривой, показанная на рис. 4.60, соответствует слабо компенсированному образцу при низкой температуре при условии, что n_e удовлетворяет соотношению (4.107) (т. е. преоб-

ладает излучательная или фононная рекомбинация). Область квадратичной зависимости, показанная на рисунке, сужается и исчезает либо при высоких температурах, либо при большей компенсации. В случае когда излучательная рекомбинация является преобладающим механизмом, поведение при больших I_0 и больших n_e осложняется процессами вынужденной эмиссии; этот вопрос будет обсуждаться нами ниже в связи с излучательными переходами зона — зона.

Как исследуется в задаче 4.24, зависимость n_e от I_0 становится более сложной, чем это показано на рис. 4.60, если рекомбинация преимущественно следует по оже-механизму.

Избыточные электроны в зоне проводимости полупроводника, обусловленные поглощением света, приводят к дополнительной проводимости полупроводника — *примесной фотопроводимости.* Примесные фотопроводники широко используются, как очень чувствительные детекторы фотонов в инфракрасной области спектра [68]. Нестационарный фотоответ n_e на нестационарное освещение обычно используется в этих практических приложениях. Так, предположим, что избыточная генерация, обусловленная светом, представляет собой некоторую произвольную функцию от времени $g_E(t)$. Каков будет отклик $n_e(t)$? Ответ на этот вопрос может быть сложным, если τ_n является функцией n_e [как в формуле (4.107)], но предположим, что отклонение от равновесия достаточно мало, так что $\tau_n = \tau_0$ представляет собой константу. В этом случае связать $n_e(t)$ с g_E может простой интеграл по всем временам, предшествующим t.

Итак, запишем g_E для всех моментов времени как последовательность мгновенных процессов генерации, каждый из которых описывается дельта-функцией. Так, например, активность генерации в момент времени t_0 равна $N\delta(t-t_0)$. Тогда уравнение (4.100) требует, чтобы из N электронов, созданных в момент времени t_0,

$$n_e(t) = N \exp[(t_0 - t)/\tau_0] \tag{4.108}$$

еще сохранились к моменту времени t_0. Поскольку для предполагаемого нами постоянного времени жизни уравнение (4.100) линейно, решение для $n_e(t)$ в том случае, когда g_E представляет собой непрерывную или разрывную функцию времени, как заранее ясно, есть сумма членов типа (4.108), т. е.

$$n_e(t) = \int_{-\infty}^{t} g_E(t_0) \exp[(t_0 - t)/\tau_0]\, dt_0. \tag{4.109}$$

---

[68] Книги Бьюба и Рывкина, приведенные в библиографии к гл. 4, содержат подробное обсуждение практических аспектов использования примесных фотопроводников.

Таким образом, когда время жизни *не зависит* от n_e и справедлива формула (4.109), становится достаточно простой задача определения временно́й зависимости n_e для некоторой функции генерации, которая может быть изолированным импульсом (прямоугольным, треугольным, дельта-функцией и т. п.) или серией импульсов. Задача 4.25 касается отклика на прямоугольные импульсы и на синусоидально изменяющуюся генерацию. Далее в задаче рассматривается, что происходит, когда $(r—g)$ изменяется как n_e^2 для большого отклонения от равновесия [как это имеет место при выполнении соотношения (4.107)]. Затухание в этом случае для больших возмущений является гиперболическим и приобретает экспоненциальный вид, когда величина n_e становится достаточно малой, чтобы обеспечить постоянство τ_n.

Динамика генерационно-рекомбинационного процесса обсуждалась нами главным образом для примесной ситуации не потому, что такая ситуация является более важной, чем динамика генерации-рекомбинации для процессов зона — зона, а потому, что такая ситуация служит упрощенным введением при переходе к более сложным представлениям обсуждаемого вопроса.

Генерация и рекомбинация электронно-дырочных пар

Рискуя повториться, отметим опять, что соответствие законам термодинамики и принципу детального равновесия требует, чтобы каждый механизм генерации носителей имел механизм рекомбинации, в точности ему противоположный. Сейчас мы рассмотрим такие пары обратных друг другу процессов, которые создают или уничтожают электронно-дырочные па́ры. Скорости таких взаимно обратных процессов в равновесии точно равны друг другу для процесса превращения энергии каждого вида. Однако, когда $n \neq n_0$ и $p \neq p_0$, каждый вид процесса превращения энергии должен вносить некоторый дисбаланс в соотношение скоростей генерации и рекомбинации.

Предположим, например, что избыточные электронно-дырочные па́ры создаются в полупроводниковом образце освещением. Тогда мы имеем избыточные электроны и дырки, которые были генерированы излучением. Это вызывает увеличение скорости рекомбинации каждого типа, однако будет или нет *излучательная* рекомбинация играть при этом главную роль зависит от типа полупроводника, его температуры, распределения в нем примесных центров. Важно отметить, что тот механизм рекомбинации, который доминирует в возвращении величины $(r—g)$ к равновесию, зависит от самого *полупровод-*

ника, а не от того, что послужило причиной первоначального отклонения от равновесия.

Автором этой книги были в свое время подробно рассмотрены различные виды механизмов рекомбинации, физически возможные в полупроводнике («Статистика полупроводников»). В этой книге читатель может найти детальное описание и связанные с ним математические выводы для процессов, перечисленных ниже.

Излучательная рекомбинация

Излучательная рекомбинация, идущая из зоны в зону, есть процесс, обратный собственному оптическому поглощению. Для твердого тела с непрямой запрещенной зоной это очень слабый процесс, так как в таких твердых телах слабым является оптическое поглощение. Однако излучательная рекомбинация может быть очень эффективным процессом в прямозонном полупроводнике, и *люминесценция* (испускание рекомбинационного излучения) для некоторых прямозонных полупроводников может быть достаточно сильной. Когда n и p в результате какого-то внешнего воздействия становятся очень велики, спонтанная излучательная рекомбинация значительно усиливается *вынужденной* (стимулированной) излучательной рекомбинацией. В соответствующих условиях вынужденная рекомбинация может становиться настолько доминирующей, что создает возможность для *лазерного* эффекта.

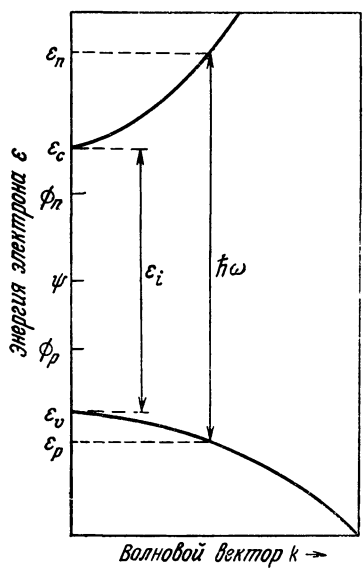


Рис. 4.61. Модель прямозонного полупроводника для обсуждения излучательных переходов между валентными состояниями с энергией ε_p и состояниями зоны проводимости, отвечающими тому же волновому вектору и энергии ε_n

Обсуждая в предыдущем разделе примесные генерационно-рекомбинационные процессы, мы касались механизмов превращения энергии, однако нас не беспокоил вопрос сохранения количества движения. Этот закон сохранения можно было не учитывать, так как большая масса примесного центра позволяет ему легко поглотить изменение количества движения электрона. Теперь, при рассмотрении излуча-

тельных переходов между состояниями валентной зоны и зоны проводимости мы должны требовать сохранения количества движения так же, как и энергии. Поэтому возможны переходы между состояниями с энергией ε_p валентной зоны (рис. 4.61) и теми состояниями с энергией ε_n зоны проводимости, которые обладают таким же волновым вектором. Полупроводник на рис. 4.61 показан отклоненным от равновесия, так как изображены два разных квазиуровня Ферми φ_n и φ_p. Относительное заполнение состояний с энергией ε_n есть

$$f_n = \{1 + \exp[(\varepsilon_n - \varphi_n)/k_0T]\}^{-1}, \tag{4.110}$$

в то время как доля состояний с энергией ε_p, *не занятых* электронами, составляет

$$(1 - f_p) = \{1 + \exp[(\varphi_p - \varepsilon_p)/k_0T]\}^{-1}. \tag{4.111}$$

Таким образом, результирующая скорость излучательных переходов между ε_p и ε_n есть алгебраическая сумма трех членов:

$$dr = dr_{sp} + dr_{st} - dg_{st}, \tag{4.112}$$

где dr_{sp} — скорость, с которой происходит спонтанная электронно-дырочная рекомбинация, при которой энергия освобождается в виде фотонов с энергией $\hbar\omega = (\varepsilon_n - \varepsilon_p)$. Третий член dg_{st} есть скорость, с которой процессы генерации *стимулируются* присутствием в полупроводнике фотонов с соответствующей энергией, а dr_{st} — скорость, с которой *обратные* переходы стимулируются присутствием фотонного поля. При записи этих трех скоростей может быть использован закон действующих масс в следующем виде:

$$\begin{aligned} dr_{sp} &= Af_n(1 - f_p), \\ dr_{st} &= Bf_n(1 - f_p)N_\omega, \\ dg_{st} &= Bf_p(1 - f_n)N_\omega, \end{aligned} \tag{4.113}$$

где N_ω — число фотонов на единицу частоты с энергией $\hbar\omega$ [которое должно свестись к величине $\langle n \rangle$ в формуле (2.54) в случае теплового равновесия], а величины A и B содержат информацию о плотности состояний при энергиях ε_n и ε_p и матричном элементе, связывающем эти состояния. Эйнштейн[69] показал, что вынужденные переходы вверх и вниз должны иметь один и тот же коэффициент B, а также продемонстрировал, что

$$A = B. \tag{4.114}$$

Это значительное упрощение величин в формулах (4.113) можно проверить (задача 4.26), потребовав, чтобы скорость dr

[69] *Einstein A.*— Z. Physik, **18**, 121 (1917).

в (4.112) обращалась в нуль в тепловом равновесии, т. е. в ситуации, когда $N_\omega \to \langle n \rangle$ из формулы (2.54), а величины f_n и f_p выражаются через одну и ту же энергию Ферми.

Вследствие равенства коэффициентов Эйнштейна A и B результирующая скорость излучательной рекомбинации в отсутствие равновесия есть

$$dr = A\,[f_n(1-f_p)-(f_p-f_n)\,N_\omega]. \tag{4.115}$$

Для умеренного отклонения от равновесия величина (f_p-f_n) почти равна единице, а N_ω незначительно изменено по сравнению с его равновесным значением $\langle n \rangle$. Тогда

$$dr \approx n_e(n_0+p_0+n_e)\,(A/N_cN_v)\exp\left[-(\hbar\omega-\varepsilon_i)/k_0T\right]. \tag{4.116}$$

Если проинтегрировать эту величину по всем комбинациям состояний валентной зоны и зоны проводимости, то получим время жизни для спонтанной излучательной рекомбинации.

$$\tau_R \equiv (n_e/r) = \tau_{R0}\left[1+\left(\frac{n_e}{n_0+p_0}\right)\right]^{-1}, \tag{4.117}$$

которое для малой глубины модуляции имеет значение

$$\tau_{R0} = [N_cN_v/Ak_0T\,(n_0+p_0)]. \tag{4.118}$$

Поэтому, когда в полупроводнике преобладают излучательные переходы, время жизни электронов и дырок при малой модуляции имеет характерную зависимость от концентрации $(n_0+p_0)^{-1}$. Время жизни будет очень быстро меняться с температурой в случае собственного полупроводника, однако оно очень слабо зависит от температуры в диапазоне, в котором примесная концентрация основных носителей не изменяется. Из выражения (4.117) можно видеть, что время жизни, не зависящее от n_e, может наблюдаться, если величина n_e мала по сравнению с n_0 *или* p_0, однако рекомбинация квадратична по n_e для больших отклонений от равновесия, как это было найдено для излучательного времени жизни в примесной ситуации [формула (4.107)].

Когда концентрации избыточных носителей велики, вынужденная рекомбинация становится важным фактором, а *инверсная заселенность* $f_p > f_n$ приводит к тому, что вынужденная рекомбинация преобладает над вынужденной эмиссией. Это можно видеть, по-другому записав выражение (4.115). Скорость спонтанной рекомбинации составляет $dr_{sp} = Af_n(1-f_p)$, и чистая скорость переходов вниз за вычетом излучательных переходов вверх может быть выражена как

$$dr = dr_{sp}\left[1-N_\omega\left\{\exp\left[\frac{\hbar\omega-(\varphi_n-\varphi_p)}{k_0T}\right]-1\right\}\right]. \tag{4.119}$$

Таким образом, если $(\varphi_n-\varphi_p)$ больше, чем энергетический зазор между состояниями (это равносильно утверждению, что f_n

меньше, чем f_p), то член в скобках в формуле (4.119) *увеличивает* скорость рекомбинации, причем это увеличение пропорционально N_ω. Это является основой лазерного эффекта в полупроводнике. Достаточно большое число электронно-дырочных пар, которое должно быть создано, чтобы сделать $(\varphi_n - \varphi_p)$ больше, чем собственная щель, является необходимым, но не достаточным условием лазерного эффекта. Работающий лазер можно получить, только если выполнены также условия оптического усиления [70].

Необходимая инверсная заселенность может быть достигнута в некоторых полупроводниках путем инжекции в переходную область p—n-перехода. Для полупроводников, в которых нельзя осуществить p—n-переходы, лазерный эффект может быть возбужден кратковременной генерацией электронно-дырочных пар с помощью электронных или фотонных пучков [71].

Лазерное усиление в полупроводнике чрезвычайно маловероятно, если полупроводник не является прямозонным, с сильной связью между валентной зоной и зоной проводимости, возможной у прямой щели. Для непрямозонного вещества излучательная рекомбинация гораздо слабее, поскольку в каждом переходе должны участвовать фононы. Так, GaAs дает усиление на длине волны порядка 0,9 мкм, однако GaP (с непрямой щелью) не может давать усиления в видимой части спектра. Люминесцентные устройства в видимой области представляют огромный интерес для систем с оптическими дисплеями [72], и эффективность спонтанной излучательной рекомбинации для этих целей часто может быть достаточна. Другие применения люминесценции неосуществимы без мощного усиления лазерной системы.

Излучательная рекомбинация может быть двухступенчатым процессом, в котором первая ступень состоит в образовании электронно-дырочного экситона [51] с последующей излучательной рекомбинацией этого экситона. Келдыш (1968) предсказал и Покровский с сотр. [73] подтвердили экспериментально, что

---

[70] Эти условия, при которых вынужденная излучательная рекомбинация может стать определяющим фактором в полупроводнике, обсуждается в работе: *Lasher G., Stern F.*— Phys. Rev., **133**, A553 (1964).

[71] *Kressel H.*, Lasers, v. 3, eds. A. K. Levine, A. J. DeMaria, Dekker, 1971. В этой работе дано подробное обсуждение и библиография по полупроводниковым лазерам.

[72] *Loebner E. E.*— Proc. IEEE, **61**, 837 (1973). В работе дан обзор прошлого и вероятного будущего использования электролюминесцентных твердых тел для дисплеев в области видимого спектра.

[73] *Pokrovsky Y., Kaminsky A., Svistunova K.* Proc. 10th Int. Conf. on Physics of Semiconductors, A. E. C., 1970, p. 504. См. также: *Benoit C., á la Guillaume et al.*— Phys. Rev., **B5**, 3079 (1972); **B7**, 1723 (1973).

высокая концентрация экситонов в полупроводнике, создавшаяся в результате интенсивного образования пар, может при определенных условиях конденсироваться в дискретные сферические капли вырожденной электронно-дырочной плазмы диаметром в несколько микрон. Излучательное время жизни таких капель в непрямозонном полупроводнике (таком, как Ge или Si) при низких температурах достаточно велико, так что эти капли можно наблюдать с помощью оптических методов, таких, как рэлеевское рассеяние.

Оже-рекомбинация

В прямозонном полупроводнике (в котором излучательная рекомбинация может быть очень эффективным процессом) с заметной вероятностью может протекать и рекомбинация типа *Оже*, или рекомбинация *с участием третьего тела*. Это процесс, обратный ударной ионизации; электрон и дырка рекомбинируют и передают энергию аннигиляции третьей частице, электрону или дырке. Такой процесс должен происходить с сохранением как энергии, так и импульса, и вероятность таких процессов сильно возрастает, если отношение (m_c/m_v) значительно меньше или значительно больше единицы (см. задачу 4.27). Мы не будем здесь входить в подробности оже-рекомбинации зона — зона, отсылая заинтересованного читателя к нашей монографии «Статистика полупроводников».

Для прямозонного полупроводника время жизни избыточных носителей, обусловленное комбинацией излучательной и оже-рекомбинаций, оказывается весьма малым (в диапазоне 10^{-10}—10^{-7} с в зависимости от равновесных концентраций электронов и дырок). Однако одновременно удовлетворить сохранению энергии и импульса, как того требует оже-рекомбинация, в непрямозонном полупроводнике значительно труднее. Аналогично излучательная рекомбинация в непрямозонном твердом теле протекает очень медленно, поскольку излучательный процесс является обратным довольно малоэффективному процессу поглощения с непрямыми переходами. Поэтому время жизни избыточных носителей, рассчитанное для непрямозонного полупроводника, может лежать в пределах от 10^{-2} до 10^{6} с, если оно зависит только от рекомбинационных процессов зона — зона.

Рекомбинация Холла—Шокли—Рида (ХШР)

Несмотря на сделанные выше оценки, дающие очень большие времена жизни для непрямозонного полупроводника, в большинстве хорошо известных непрямозонных твердых тел

для измеряемых значений времен жизни носителей гораздо более характерны значения, лежащие в диапазоне 10^{-7}—10^{-3} с. Рекомбинация в таких материалах полностью определяется двухступенчатыми процессами, включающими случайные (а иногда сознательно введенные) локализованные примесные центры. По-видимому, в любом прямозонном кристалле эти процессы, идущие с участием примесных центров, также дают некоторый вклад в рекомбинацию, однако они с меньшей вероятностью могут доминировать, поскольку прямые процессы очень эффективны.

Холл [74] и Шокли и Рид [75] независимо пришли к выводу, что примесный центр может действовать как центр захвата, содействуя электронно-дырочной рекомбинации, и так называемая модель ХШР оказалась справедливой при самых разных условиях. Когда центр захватывает электрон, а затем дырку, он возвращается к своему исходному зарядовому состоянию и может начать весь процесс снова. Модель ХШР может быть достаточно сложной даже в том случае, когда рекомбинация происходит с помощью одного типа центров, имеющих одну и ту же энергию в запрещенной зоне, как это подробно обсуждается в «Статистике полупроводников». Рассмотрение центров, которые могут захватить несколько дырок (или электронов) одну за другой, прежде чем начнут захватывать носители противоположного знака, приводит к еще большей сложности анализа [76]. После первоначального рассмотрения Калашникова [77] формальный анализ полупроводника с несколькими различными типами центров захвата серьезно выполнен не был, хотя имеется достаточно глубокое *качественное* понимание того, какие типы центров захвата будут играть главную роль в широкозонном полупроводнике или диэлектрике со спектром состояний дефектов, распределенных по всей запрещенной зоне [78].

Примесные центры, которые обеспечивают электронно-дырочную рекомбинацию,— это обычно не те центры, которые дают главный вклад в концентрацию носителей в состоянии равновесия. Примесные состояния наиболее эффективны с точки зрения рекомбинации, когда они расположены вблизи середины запрещенной зоны. Дефект, который действует как рекомбинационный центр, обычно имеет сильно различающиеся сечения захвата для электронов и дырок, и это часто приводит к проблемам прилипания, т. е. к ситуации, когда $n_e \neq p_e$.

[74] *Hall R. N.*— Phys. Rev., 83, 228 (1951); Phys. Rev., 87, 387 (1952).
[75] *Shockley W., Read W. T.*— Phys. Rev., **87**, 835 (1952).
[76] *Sah C. T., Shockley W.*— Phys. Rev., **109**, 1103 (1958).
[77] *Калашников С. Г.*— ЖТФ, 1956, т. 26, с. 241.
[78] *Rose A.* Concepts in Photoconductivity and Allied Problems, Interscience, 1963; *Simmons J. G., Taylor G. W.*— Phys. Rev., **B4**, 502 (1971).

Здесь не будет дано вывода выражений для времен жизни в модели ХШР ни с прилипанием, ни без него. Укажем просто, что если концентрация центров достаточно мала, чтобы n_e и p_e были одинаковы, время жизни дырок τ_p равно времени жизни электронов τ_n. Это время жизни имеет вид

$$\tau_p = \tau_n = \frac{\tau_0 (n_0 + p_0) + \tau_\infty n_e}{n_0 + p_e + n_e}, \tag{4.120}$$

где τ_0 и τ_∞ суть предельные значения времени жизни для очень малых и очень больших нарушений равновесия (и каждое из них обратно пропорционально концентрации рекомбинационных центров).

Хотя примесные состояния, находящиеся при различных энергиях, и могут играть какую-то небольшую роль в рекомбинационном процессе, какой-то один тип дефекта [как это предполагалось при выводе выражения (4.120)] часто оказывается гораздо более важным, чем все остальные вместе взятые. Поэтому так существенно избегать малейших следов загрязнения медью, если германий должен быть получен с большим временем жизни, как это требуется для сплавных германиевых *pnp*-транзисторов. Напротив, если изготовитель транзисторов хочет получить кремниевые переключающие транзисторы с малым временем переключения, он сознательно вводит в свой кремний атомы золота — очень эффективные рекомбинационные центры.

Время жизни в сочетании со способностью носителей к дрейфу и диффузии определяет, насколько далеко *неосновные носители* могут продвинуться в глубь полупроводника, прежде чем они исчезнут путем рекомбинации. Далее мы должны рассмотреть, что заставляет электроны и дырки в полупроводнике двигаться.

Дрейфовые и диффузионные токи

Имея дело с металлическими проводниками, мы привыкли считать, что ток возникает исключительно как результат действия приложенного поля. Диффузионные токи, обусловленные пространственным распределением концентрации свободных носителей, в металле пренебрежимо малы. Не так обстоит дело в полупроводнике, где должны учитываться оба типа токов. Рассмотрим ток, переносимый электронами в полупроводниковом кристалле.

Дрейфовый ток электронов зависит от электрического поля, концентрации электронов и их подвижности:

$$\mathbf{J}_f = -e\,\mathbf{E}\,n\mu_n = n\mu_n \nabla \varepsilon_c. \tag{4.121}$$

Его можно описать выражением, представленным в правой части этого равенства, поскольку электрическое поле создает у зон как функций координат такой наклон, чтобы выполнялось соотношение $\nabla_c \varepsilon_c = -e\mathbf{E}$.

Диффузионный ток, который следует прибавлять к $\mathbf{J}_f$, зависит от градиента электронной концентрации и коэффициента диффузии электронов

$$\mathbf{J}_d = eD_n \nabla_n. \tag{4.122}$$

Как показал Эйнштейн[79] в своем классическом анализе броуновского движения, коэффициент диффузии и подвижность связаны соотношением

$$D = (k_0 T \mu / e), \tag{4.123}$$

а это означает, что

$$\mathbf{J}_d = k_0 T \mu_n \nabla_n. \tag{4.124}$$

Таким образом, полная плотность тока, обусловленная всеми электронами, имеет вид

$$\mathbf{J}_n = (\mathbf{J}_f + \mathbf{J}_d) = \mu_n (k_0 T \nabla n + n \nabla \varepsilon_c). \tag{4.125}$$

Если теперь мы запишем полную концентрацию электронов в виде (4.92), то ∇n можно будет выразить через $\nabla \varphi_n$. Согласно (4.92),

$$\nabla n = (n/k_0 T)\,[\nabla \varphi_n - \nabla \varepsilon_c], \tag{4.126}$$

и из соотношений (4.125) и (4.126) мы получаем очень простое выражение для плотности тока

$$\mathbf{J}_n = n \mu_n \nabla \varphi_n. \tag{4.127}$$

Таким образом, величина $e^{-1}\nabla \varphi_n$ служит *эффективным полем*, производящим нескомпенсированный электронный ток. В том случае, когда величина $\nabla \varphi_n$ всюду равна нулю (как должно быть в тепловом равновесии), не существует *результирующего* потока электронов из одного места в другое в неоднородно легированном кристалле, хотя и могут существовать равные и противоположно направленные диффузионные и дрейфовые потоки. Как только тепловое равновесие оказывается нарушенным, возможность для диффузионного и дрейфового токов точно компенсировать друг друга утрачивается и возникает результирующий поток электронов. Аналогичные рассуждения применимы к дрейфу и диффузии дырок.

Предположим, что избыточные электроны размещены в зоне проводимости полупроводника при таких условиях, которые делают их движение обусловленным диффузией, а не электри-

[79] *Einstein A.*— Z. Physik, **17**, 549 (1905).

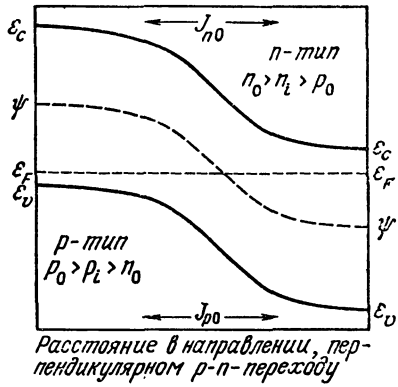


Рис. 4.62. Идеализированная одномерная картина изгиба зон в направлении p—n-перехода в тепловом равновесии. Поскольку величина $\nabla\mathcal{E}_c$ отлична от нуля, существует большое статическое электрическое поле в области p—n-перехода, вероятно, порядка 10^6 В/м или более. Оно обусловливает дрейфовое движение электронов вправо и дырок влево. В каждой точке $\mathbf{J}_n = \mathbf{J}_p = 0$, поскольку диффузия электронов влево и дырок вправо компенсирует дрейфовые токи.

ческим полем. В конце концов они все испытают рекомбинацию, и доля электронов, которая до рекомбинации пройдет при беспорядочном движении расстояние L, есть $\sim\exp(-L/L_n)$. Здесь характерную длину L_n называют *диффузионной длиной электронов*:

$$L_n = (\tau_n D_n)^{1/2}. \tag{4.128}$$

Эффективная длина, характеризующая диффузию дырок, имеет аналогичный вид $L_p = (\tau_p D_p)^{1/2}$.

Длины L_n и L_p чрезвычайно важны для работы полупроводниковых приборов, действие которых основано на инжекции неосновных носителей в часть полупроводникового кристалла. Основным прибором такого типа является p—n-переход, для которого Шокли[80] сформулировал принцип действия, общепринятый и сегодня. На рис. 4.62 изображен p—n-переход в состоянии теплового равновесия с таким изгибом зон, который требуется, чтобы уровень Ферми был повсюду одинаковым. Две равные и противоположно направленные плотности токов (J_{p0}) обозначают диффузию и дрейф дырок и аналогично две взаимно компенсирующие друг друга плотности (J_{n0}) — дрейф и диффузию электронов.

[80] *Shockley W.*— Bell Syst. Tech. J., 28, 435 (1949). Обсуждение этого вопроса дано также в его монографии: *Shockley W.* Electrons and Holes in Semiconductors, Van Nostrand, 1950. [Имеется перевод: *Шокли В.* Теория электронных полупроводников.— М.: ИЛ, 1953.]

Примерный порядок величин J_{n0} и J_{p0} мы можем оценить следующим образом. Дрейфовый электронный поток обусловлен теми электронами, которые оказались генерированными с p-стороны перехода на расстоянии от него, не превышающем диффузионной длины электрона, и поэтому смогли продиффундировать к переходу и скатиться по наклону. Поэтому можно ожидать

$$J_{n0} \approx (en_i^2 L_n/p_0\tau_n) = (en_i^2/p_0)(D_n/\tau_n)^{1/2},$$
$$J_{p0} \approx (en_i^2 L_p/n_0\tau_p) = (en_i^2/n_0)(D_p/\tau_p)^{1/2}. \qquad (4.129)$$

Таковы эти величины в равновесии. Если к переходу приложено напряжение, такое, что p-область становится заряженной отрицательно по отношению к n-области, увеличивая тем самым высоту барьера, то диффузионные токи оказываются заблокированными. Результирующая плотность тока составляет $(J_{n0}+J_{p0})$ и обусловлена генерацией электронов и дырок в пределах примерно одной диффузионной длины с каждой стороны барьера. Существенная особенность этого тока состоит в том, что при $V>k_0T/e$ он не зависит от напряжения, и ток в хороших p—n-переходах в высокоомном (или запорном) направлении очень мал.

Напряжение другого знака понижает потенциальный барьер и позволяет диффузии переносить большое количество электронов в p-область и дырок в n-область. В этом состоит явление инжекции носителей, которое приводит к биполярному транзисторному действию, если инжектированные носители могут быть захвачены еще одной по-иному легированной областью кристалла. Соотношение между результирующим полным током и приложенным напряжением должно, таким образом, иметь вид

$$J = (J_{n0} + J_{p0})[\exp(eV/k_0T) - 1]. \qquad (4.130)$$

Любой реальный p—n-переход обладает характеристикой, отличающейся от (4.130) по множеству осложняющих причин. Заинтересованный читатель найдет обширную литературу, касающуюся принципов работы полупроводниковых транзисторов, интегральных схем и многочисленных других приборов. Библиография в конце этой главы должна послужить отправной точкой в знакомстве с этой литературой.

Задачи

4.1. В антимониде индия n-типа, содержащем мелкие донорные примеси, перекрытие волновых функций соседних примесных центров приводит к тому, что концентрация n_0 не зависит от температуры. Электроны прово-

димости в этом полупроводнике имеют эффективную массу $m_c=0{,}014\,m$. Покажите, что при абсолютном нуле в образце, в котором $n_0=10^{22}$ м$^{-3}$, уровень Ферми расположен на 0,012 эВ выше ε_c. Рассчитайте, как понижается ε_F по мере нагревания образца до комнатной температуры, и постройте график его положения относительно ε_c в зависимости от T. (При желании можете использовать для интеграла Ферми выражения, данные в примечании 6, которые перекрывают неудобную область частичного вырождения. У вас должно получиться, что $\varepsilon_F=\varepsilon_c$ при $T\approx 140$ К.) Теперь опять представьте себе, что образец находится при $T=0$, и найдите период P в Тл$^{-1}$ между соседними максимумами для осцилляций ШдГ.

4.2. Определите наклон и пересечение прямой с осью ординат для результатов измерений собственной проводимости в германии по рис. 4.7, *б*. Эти числа следует проинтерпретировать с помощью соотношения (4.20). Считайте, что $m_c=0{,}56m$ есть скалярная *эффективная масса плотности состояний* для зоны проводимости Ge и что соответствующая величина для валентной зоны составляет $m_v=0{,}37m$. У вас должно получиться, что согласие измеренного вами отрезка оси ординат с равенством (4.20) требует температурной зависимости с $\alpha=(-d\varepsilon_i/dT)=3{,}7\cdot 10^{-4}$ эВ/К. Сравните выведенную вами зависимость $\varepsilon_i=(\varepsilon_{i0}-\alpha T)$ с кривой на рис. 4.8, полученной из оптических данных.

4.3. С помощью различных видов измерений считается установленным, что InSb при 300 К имеет ширину запрещенной зоны $\varepsilon_i=0{,}18$ эВ и что $(d\varepsilon_i/dT)=-2{,}8\cdot 10^{-4}$ эВ/К. Рассчитайте величину щели в InSb из данных для n_i на рис. 4.9. Пусть величина щели удовлетворяет формуле (4.17) для четырех или пяти температур на этом рисунке. Положите $m_c=0{,}014m$, а $m_v=0{,}4m$. Считайте также, что m_c и m_v не зависят от температуры (что для m_c не является хорошим допущением), а формула (4.17) справедлива при всех температурах. (Вырождение электронов при высоких температурах нарушает это последнее предположение.) Предположения, которые вам предложено было сделать, приведут вас к неверному результату. Пример этого анализа служит предупреждением об осторожности, необходимой при выборе аппроксимаций для экспериментальных результатов, полученных на новом полупроводнике.

4.4. Ширина собственной зоны арсенида индия была измерена Диксоном и Эллисом [*Dixon J. R., Ellis J. M.*— Phys. Rev., **123**, 1560 (1961)] по энергии фотонов, при которой начинается интенсивное оптическое поглощение. Как было показано Варшни (примечание 9) эта оптическая щель должна изменяться с температурой согласно соотношению

$$\varepsilon_i=0{,}426-3{,}16\cdot 10^{-4}\,T^2(93+T)^{-1}\ \text{эВ}.$$

Положив $m_c=0{,}022m$, $m_v=0{,}4m$, постройте ε_c и ψ как функции температуры в интервале температур 250—500 К, используя ε_v как начало отсчета энергии. Вы должны для различных температур измерить по рис. 4.11 величины отношения (p_0/n_i) с точностью, достаточной для того, чтобы показать ход ε_F для каждого из этих двух образцов InAs при температурах ниже перехода от собственной проводимости к примесной.

4.5. Рассмотрим *водородоподобный донор*, для которого кулоновское притяжение обращающегося связанного электрона к примесному центру предполагается обратно пропорциональным диэлектрической проницаемости $\varkappa$ на всех расстояниях. Покажите, что основное состояние (состояние с моментом количества движения, равным $\hbar$) модели Бора имеет в этом случае радиус, определяемый формулой (4.25). Вы должны получить $\varepsilon_d=0{,}0133$ эВ и $a_d=51$ Å для водородоподобных доноров в теллуриде кадмия, так как этот

полупроводник имеет зону проводимости с одним минимумом в центре этой зоны с массой $m_c=0{,}11m$ и диэлектрическую проницаемость $\varkappa=10{,}6$.

4.6. Решите уравнение (4.36) численно для нескольких температур в интервале от 10 до 100 К для двух гипотетических образцов CdTe ($m_c=0{,}11m$), содержащих водородоподобные доноры ($\beta=1/2$, $\varepsilon_d=0{,}0133$ эВ) того типа, который рассматривался в задаче 4.5. Для каждого из этих образцов $(N_d-N_a)=10^{22}$ м$^{-3}$, однако в одном из них содержится $N_a=10^{21}$ м$^{-3}$ компенсирующих акцепторов, в то время как в другом — только $N_a=10^{19}$ м$^{-3}$. Вы должны получить, что полулогарифмическая зависимость n_0 от $(1/T)$ ведет себя подобно зависимости рис. 4.19. Теперь с помощью равенства (4.41) определите, как ε_F изменяется с температурой в каждом из образцов; в общих чертах результат будет напоминать рис. 4.20. (Эта задача, естественно, предназначена для решения с помощью компьютера.)

4.7. Спектроскопические исследования, выполненные в инфракрасной области Аггарвалом и Рамдасом [*Aggarwal R. L., Ramdas A. K.*— Phys. Rev., 137, A602 (1965)], показывают, что доноры мышьяка в кремнии имеют расщепленное $1s$-состояние, подобное тому, которое изображено для доноров фосфора на рис. 4.16. В случае мышьяка они нашли $\Delta_c=0{,}022$ эВ. Помня об этом, воспользуйтесь равенством (4.41) для построения кривых зависимости энергии Ферми от температуры для легированных мышьяком образцов кремния *а* и *б* на рис. 4.19. Находите ли вы, что для какого-либо из образцов максимальное значение ε_F расположено достаточно высоко, чтобы вызвать заметную заселенность верхнего $1s$-состояния в некотором интервале температур? Должно ли это заметно испортить подгонку кривой к уравнению (4.36), которым мы воспользовались для построения сплошных кривых на рис. 4.19?

4.8. Дан полупроводник, в котором $m_c=0{,}25m$. Рассчитайте энергию, квазиимпульс, волновой вектор и длину волны де Бройля для электрона зоны проводимости, движущегося при $T=300$К со средней тепловой скоростью. Найдите те же величины для LA-фонона, движущегося со скоростью 5000 м/с и имеющего длину волны, вдвое большую, чем электрон. Покажите, что законы сохранения энергии и квазиимпульса позволяют такому фонону рассеять электрон на угол 29°, причем изменение энергии электрона при этом составит менее 2 %.

4.9. Электроны в зоне проводимости полупроводника, описанного в задаче 4.8, рассеиваются LA-фононами. Эффективность рассеяния может быть описана с помощью постоянной потенциала деформации $\varepsilon_1=5$ эВ. Если относительная плотность материала равна 5, покажите, что среднее время свободного пробега между столкновениями равно $\tau_m\sim 5\cdot 10^{-13}$ с при комнатной температуре, а решеточная подвижность есть $\mu_L\approx 2000\ T^{-3/2}$ м^2/(В·с). При какой температуре она окажется равной подвижности для рассеяния на заряженных центрах μ_I, вычисленной по формуле (4.62) для 10^{20} м$^{-3}$ однозарядных ионизованных центров, если диэлектрическая проницаемость $\varkappa=10$? Покажите, как подвижность, обусловленная комбинацией рассеяния на решетке и на заряженных центрах, изменяется с температурой, указав температуру, ниже которой логарифмический множитель в μ_I влияет на температурную зависимость подвижности.

4.10. Ток с плотностью J_x переносится электронами в полупроводнике, к которому приложено магнитное поле B_z. Следовательно, существует результирующее электрическое поле, которое можно описать компонентами E_x и E_y. Если все электроны имеют одинаковую дрейфовую скорость, покажите, что $E_y=E_x\mathrm{tg}(B_z\mu_H)$. Что произойдет, если длина образца не превосходит в несколько раз его ширину?

4.11. Рассмотрите собственный полупроводник, в котором все электроны имеют одну дрейфовую скорость, а все дырки — другую, так что их подвижности μ_n и μ_p различны. Покажите, что в случае протекания тока в поперечном магнитном поле поле Холла должно возникнуть в направлении $\mathbf{J}\times\mathbf{B}$, а его величина будет в стационарном состоянии уравновешивать сумму сил Лоренца, приложенных к электронам и к дыркам, так что

$$R_H = \frac{-1}{n_i e}\,\frac{(\mu_n - \mu_p)}{(\mu_n + \mu_p)}.$$

4.12. Постарайтесь проанализировать поведение некоторых из образцов антимонида индия, для которых результаты измерений электропроводности и эффекта Холла представлены на рис. 4.31 и 4.32. Считайте, что $r=1$ и концентрация ионизованных центров (n_0-p_0) или (p_0-n_0) не изменяется в исследуемом температурном интервале. Если хотите, можете воспользоваться температурной зависимостью n_i для InSb на рис. 4.9, хотя это не должно понадобиться. Покажите, что если отношение $b\equiv(\mu_n/\mu_p)$ не зависит от температуры, то отрицательный знак постоянной Холла (для образца, имеющего p-тип проводимости при низких температурах) появляется при температуре, определяемой условием $n_0\mu_n = p_0\mu_p$. Попытайтесь описать поведение образца V, одного из образцов n-типа с меньшим удельным сопротивлением и одного из образцов p-типа, показанных на рис. 4.32.

4.13. Пусть донорные центры в полупроводнике распределены регулярным образом так, что расстояние между донорами оказывается *максимальным* для данного N_d. Какие два вида распределений могут это обеспечить? Докажите, что расстояние между донорами составит в этом случае $(2/N_d^2)^{1/6}$, как указано в примечании [27]. Теперь предположите другую ситуацию, что доноры распределены в пространстве *хаотически*. Покажите, почему в этом случае следует ожидать, что в пределах расстояния $(3/4\pi N_d)^{1/}$ возле каждого донора окажется один или более других доноров.

4.14. Используйте водородоподобную модель для доноров мышьяка в германии ($\varkappa=16$), чтобы получить радиус a_d волновой функции основного состояния для таких доноров (Используйте величину ε_{d0}, показанную для таких доноров на рис. 4.33, соответствующую их бесконечному удалению друг от друга.) По величине N_{di}, при которой ε_d падает до нуля, покажите, что полупроводник ведет себя как примесный металл, если *среднее расстояние* между донорами меньше примерно $6a_d$ или если расстояние до *ближайшего* донора меньше примерно $3{,}2a_d$.

4.15. Эта довольно трудная задача рассматривает возможность создания теплового пробоя в аморфном полупроводнике. Для упрощения задачи ее геометрия выбрана так, что эквипотенциальные поверхности являются и изотермическими. Тепло течет наружу из центра симметрии вдоль линий тока. Представьте две маленькие полусферы из аморфного полупроводника с электропроводностью σ и теплопроводностью $\varkappa$. Поместите омические электроды на их плоских гранях и прижмите друг к другу кривые поверхности. При прохождении тока от одного электрода к другому область вблизи центральной точки нагревается, а сопротивление представляет собой *сопротивление растекания* с обеих сторон от этого контакта. При комнатной температуре T_0 предположим, что максимальная температура в точке контакта равна T_m. Если напряжение между электродами составляет V_0, покажите, что в этом случае

$$\int_{T_0}^{T_m} (\varkappa/\sigma)\,dT = \left(V_0^2/8\right).$$

Если $\varkappa=(\varkappa_0/T)$, а $\sigma=\sigma_0\exp(-\Delta\varepsilon/k_0T)$, то этот интеграл может быть представлен в виде разности *интегральных показательных функций* $\overline{\mathrm{Ei}}(x)$ [табулированная функция] для $x=(\Delta\varepsilon/k_0T_0)$ и $x=(\Delta\varepsilon/k_0T_m)$. Если $\varkappa_0=10^4$ Вт/м, $\sigma_0=10^4$ Ом$^{-1}$ м$^{-1}$, $\Delta\varepsilon=0{,}3$ эВ и $T_0=300$ К, покажите, что приложенное напряжение $V_0=250$ В приведет к повышению T_m до ~ 350 К, а увеличение напряжения до 290 В вызовет бесконечно большой подъем температуры центрального контакта.

4.16. Предположим, что твердое тело кристаллизуется в простой кубической решетке. Нарисуйте зону Бриллюэна для этого твердого тела. Представьте, что минимум зоны проводимости расположен на границе зоны Бриллюэна в направлении а) [100], б) [110] и в) [111]. Каким числом полных замкнутых поверхностей можно описать энергию непосредственно над ε_c и какие виды отклонения от сферической формы возможны в этих случаях? Теперь предположите, что другое твердое тело имеет гексагональную решетку (для которой 3Б также гексагональна), и считайте, что минимум энергии зоны проводимости расположен в вершине этой зоны. Сколько полных замкнутых поверхностей может существовать для энергии, слегка превышающей ε_c в такой зоне?

4.17. Покажите, что при поглощении фотона в прямозонном полупроводнике избыточная энергия $(\hbar\omega-\varepsilon_i)$ делится между электроном и дыркой таким образом, что каждый член пары приобретает волновой вектор, определяемый формулой (4.84). Если InSb обладает такой же зонной структурой, как структура GaAs, изображенная на рис. 4.48, с $m_c=0{,}014\ m$, $m_v=0{,}4\ m$ и $\varepsilon_i=0{,}18$ эВ при $T=300$ К для InSb, рассчитайте, какая часть энергии фотона, равной 0,50 эВ, превращается в кинетическую энергию свободного электрона, а какая — в энергию свободной дырки. Если мы включили импульс самого фотона в уравнение сохранения импульса, то на сколько процентов будут различаться импульсы электрона и дырки?

4.18. Связанное состояние простого прямозонного экситона (в рамках теории Бора) заключается в совместном орбитальном движении электрона с массой m_c и дырки с массой m_v в среде с диэлектрической проницаемостью $\varkappa$, таком, что полный момент количества движения относительно центра масс составляет целое кратное число $\hbar$. Получите выражение для энергии связи основного состояния такого электрона и для семейства возбужденных состояний. Удобно выразить эти величины с помощью эффективной массы комбинированной плотности состояний m_r. Рассчитайте энергии связи основного состояния и первого возбужденного состояния для прямого возбуждения в арсениде индия ($m_c=0{,}022\ m$, $m_v=0{,}4\ m$, $\varkappa=11{,}6$) и покажите, что при комнатной температуре энергия фотона, необходимая для образования экситона в основном состоянии, составляет 99,40 % энергии, необходимой для образования свободного электрона и свободной дырки в состоянии покоя (см. задачу 4.4, рассматривающую ширину запрещенной зоны в InAs). Будет ли экситон в этом полупроводнике устойчив при комнатной температуре?

4.19. Заметьте по табл. 2.1, что продольные звуковые волны низкой частоты распространяются в кристалле германия со скоростью 5400 м/с. Рассчитайте угловую частоту LA-мод, соответствующих зонной границе $\left[\frac{1}{2}\frac{1}{2}\frac{1}{2}\right]$, поскольку, как можно предположить, именно эти моды могут создаваться или исчезать при непрямых оптических переходах в Ge. Рассчитайте, какова ожидаемая разница $2k_0\theta$ между пороговыми энергиями для переходов, идущих с поглощением и с образованием фононов, если бы это была единственная ветвь фононного спектра, и сравните ваши результаты с тем, что вы видите на рис. 4.53. Используя формулу (2.33), помните, что Ge имеет двухатомный базис и что оба атома базиса обладают одинаковой массой.

4.20. По данным комбинационного рассеяния известно, что длинноволновые LO-фононы в InSb имеют частоту $5{,}85 \cdot 10^{12}$ Гц. В InSb эти фононы взаимодействуют с электронами, приводя к магнетофононному резонансу, показанному на рис. 4.55. Отложите величину $(1/B)$ для каждого максимума этой экспериментальной кривой как функцию от напечатанного над ним целого числа и убедитесь, что зависимость линейна, как это следует из формулы (4.89). Из наклона этой зависимости, используя значение частоты LO-фононов, приведенное выше, определите эффективную массу m_c для электронов проводимости в InSb. Сравните полученную вами величину (m_c/m_0) со значением, использованным вами раньше в задаче 4.17.

4.21. Рассмотрите образец 1 на рис. 4.11. Если вы решали задачу 4.4, то уже знаете, как изменяется в этом материале с температурой равновесная энергия Ферми ε_F. Изобразите теперь примерный ход квазиуровней Ферми φ_n и φ_p в этом образце при 300 и 400 К для концентраций избыточных пар 10^{20}, 10^{21} и 10^{22} м$^{-3}$. Предположите теперь, что в части образца с помощью освещения при 300 К поддерживается стационарная концентрация избыточных пар 10^{22} м$^{-3}$. Область перехода к неосвещенному материалу составляет по толщине 100 мкм. Обсудите, чем будут обусловлены токи электронов и дырок в области перехода.

4.22. Пусть в образце, имеющем форму куба с ребром 1 мм, происходит экскльзия носителей. Собственная концентрация составляет $n_i = 10^{19}$ м$^{-3}$, а $N_r = (n_0 - p_0)$ [см. формулы (4.23) и (4.24)] составляет 10^{18} м$^{-3}$. Таким образом полное обеднение должно привести к $n = N_r$, $p = 0$. Рассчитайте электростатический потенциал в центре образца, если концентрация электронов фактически уменьшена до значения $n = 9{,}95 \cdot 10^{17}$ м$^{-3}$.

4.23. Убедитесь в том, что выражение для времени жизни (4.107) получается в результате подстановки скоростей генерации и рекомбинации (4.103) и (4.106) в (4.99). Чтобы взаимно уничтожились нежелательные члены, вам понадобится использовать соотношение (4.105). Предположите теперь, что имеется стационарный поток фотонов I_0 на 1 м^2 в 1 с и способность к отклику нейтральных доноров выражается сечением фотоионизации σ_I. Определите вид зависимости результирующей стационарной величины n_e от I_0 для ситуации, в которой $n_0 \ll N_d$ и $N_a \ll N_d$. Обоснуйте форму кривой на рис. 4.60 и определите значения I_0 и n_e, при которых кривая в двойном логарифмическом масштабе изменяет наклон.

4.24. Представьте себе примесный полупроводник, в котором скорость обратного захвата электронов на примесные уровни определяется оже-рекомбинацией. Проследите переход от формулы (4.103) к (4.107), потребовав теперь, чтобы скорость рекомбинации изменялась как $n^2(n + N_a)$, и покажите, что время жизни должно изменяться как $[(n_0 + n_e)(N_a + n_e)]^{-1}$, если температура достаточно низка, чтобы удовлетворялись неравенства $n_0 \ll N_a < N_d$. Пусть теперь этот полупроводник возбуждается фотонами, так что генерация происходит излучательным путем, а в рекомбинации все еще доминирует оже-процесс. Для потока фотонов I_0 покажите, что n_e будет изменяться линейно с I_0, пока поток мал, затем последуют области зависимости $I_0^{1/2}$, $I_0^{1/3}$ и, наконец, насыщение.

4.25. Рассмотрите полупроводник, в котором время жизни τ_0 не зависит от n_e. Обсудите зависимость n_e от времени, когда дополнительная генерация g_E представляет собой периодическую прямоугольную функцию времени. Рассмотрите случаи большого и малого отношения времени, когда генерация включена, ко времени, когда она выключена. Каков будет отклик на функцию генерации $g_E = A[1 + B\sin(\omega t)]$ и для какой частоты сдвиг фаз между g_E и n_e составляет 45°? Предположите теперь, что время жизни изменяется как $\tau_n = \tau_0[1 + n_e/C]^{-1}$ [именно такой вид имеет функциональная зависимость вре-

мени жизни от n_e в формуле (4.107)]. Поставленная задача требует использования при больших n_e формулы (4.109). Однако эта формула вам не понадобится, если вы будете выяснять, как n_e меняется со временем в ответ на освещение в виде большого прямоугольного импульса. Каков будет спад n_e в ответ на очень мощный импульс освещения, имеющий вид дельта-функции?

4.26. Покажите, что коэффициенты Эйнштейна A и B равны, как это предполагается в формуле (4.114), потребовав, чтобы полная излучательная скорость dr в равновесии обращалась в нуль, и подставив равновесные значения в величины dr_{sp}, dr_{st} и dg_{st} в выражении (4.113). При этом покажите, что, когда $N_\omega \approx <n>$ и квазиуровни Ферми расположены достаточно глубоко внутри запрещенной зоны, скорость dr для неравновесного случая дается выражением (4.116).

4.27. Рассмотрите два вида межзонной оже-рекомбинации, которые могут происходить между зоной проводимости с массой m_c и валентной зоной с массой m_v. Оба экстремума находятся в центре зоны. Покажите, как можно одновременно удовлетворить условиям сохранения энергии и количества движения, возбуждая либо дырку с высокой энергией, либо высокоэнергетический электрон одновременно с аннигиляцией дырочно-электронной пары. Почему вероятность такого вида рекомбинации увеличивается, если массы m_c и m_v заметно отличаются друг от друга?

Литература

Полупроводники: общие вопросы

Progress in Semiconductors, ed. A. F. Gibson, Wiley, v. 1, 1956.— v. 9, 1965.

Киреев П. С. Физика полупроводников.— М.: Высшая школа, 1975.

McKelvey J. P. Solid State and Semiconductor Physics, Krieger, 1982.

Handbook on Semiconductors, ed. T. S. Moss, North-Holland, 1980—82, 4 vols.

Ridley B. K. Quantum Processes in Semiconductors, Oxford Univ. Press, 1982. [Имеется перевод: *Ридли Б.* Квантовые процессы в полупроводниках.— М.: Мир, 1986.]

Seeger K. Semiconductor Physics, Springer, 2nd ed., 1982. [Имеется перевод: *Зеегер К.* Физика полупроводников.— М.: Мир, 1977.]

Shockley W. Electrons and Holes in Semiconductors, Van Nostrand, 1950. [Имеется перевод: *Шокли В.* Теория электронных полупроводников.— М.: ИЛ, 1953.]

Smith R. A. Semiconductors, Cambridge Univ. Press, 2nd ed., 1978. [Имеется перевод: *Смит Р.* Полупроводники.— М.: Мир, 1981.]

Semiconductors and Semimetals, eds. R. K. Willardson, A. C. Beer, Academic Press, v. 1, 1966 — v. 21, 1984.

* *Бонч-Бруевич В. Л., Калашников С. Г.* Физика полупроводников.— М.: Наука, 1977.

Равновесное распределение электронов

Blakemore J. S. Semiconductor Statistics, Pergamon Press, 1962. [Имеется перевод: *Блейкмор Дж.* Статистика электронов в полупроводниках.— М.: ИЛ, 1964.]

Spenke E. Electronic Semiconductors, McGraw-Hill, 1958. (См. также цитированные выше книги Киреева, Мак-Келви, Шокли и Бонч-Бруевича).

Кристаллические полупроводниковые материалы

Semiconductor Silicon 1981, ed. H. R. Huff, Electrochemical Soc., 1981.

Lovett D. R. Semimetals and Narrow Bandgap Semiconductors, Pion, 1977.

Madelung O. Physics of III—V Compounds, Wiley, 1964. [Имеется перевод: *Маделунг О.* Физика полупроводниковых соединений элементов III и V групп.— М.: Мир, 1967.]

GaInAsP Alloy Semiconductors, ed. T. P. Pearsall, Wiley, 1982.
Pope M., Swenberg C. E. Electronic Processes in Organic Cristals, Oxford Univ. Press, 1982.
Ray B. II — VI Compounds, Pergamon Press, 1969.
Shay J. L., Wernick J. H. Ternary Chalcopyrite Semiconductors, Pergamon Press, 1975.

Примесные эффекты в полупроводниках

Blakemore J. S., Rahimi S. Models for Midgap Centers in Gallium Arsenide.— In: Semiconductors and Semimetals, v. 20, Academic Press, 1984.
Bourgoin J., Lannoo M. Point Defects in Semiconductors II: Experimental Aspects, Springer, 1983. [Имеется перевод: *Бургуен Ж., Ланоо М.* Точечные дефекты в полупроводниках II. Экспериментальные аспекты.—М.: Мир, 1985.]
Defects in Silicon, eds. W. M. Bullis, L. C. Kimmerling, Electrochemical Society, 1983.
Фистуль В. И. Сильно легированные полупроводники.— М.: Физматгиз, 1965.
Jaros M. Deep Levels in Semiconductors, Hilger, 1982.
Lannoo M., Bourgoin J. Point Defect in Semiconductors I: Theoretical Aspects, Springer, 1981. [Имеется перевод: *Ланоо М., Бургуен Ж.* Точечные дефекты в полупроводниках. Теория.— М.: Мир, 1984.]
* *Шкловский Б. И., Эфрос А. Л.* Электронные свойства легированных полупроводников.— М.: Наука, 1979.
Larrabee R. D. Neutron Transmutation of Semiconductor Materials, Plenum Press, 1984.
Defects in Semiconductors II: eds. S. Mahajan, J. W. Corbett, North-Holland, 1983.
Milnes A. G. Deep Impurities in Semiconductors, Wiley-Interscience, 1973. [Имеется перевод: *Милнс А.* Примеси с глубокими уровнями в полупроводниках.— М.: Мир, 1977.]
Stoneham A. M. Theory of Defects in Solids, Oxford Univ. Press, 1975. [Имеется перевод: *Стоунхэм А. М.* Теория дефектов в твердых телах.— М.: Мир, 1978.]
Townsend P. D., Kelly J. C. Colour Centres and Imperfections in Insulators and Semiconductors, Crane, Russak, 1973.

Явления переноса в кристаллических полупроводниках

Beer A. C. Galvanomagnetic Effects in Semiconductors, Academic Press, 1963.
Conwell E. M. High Field Transport in Semiconductors, Academic Press, 1967. [Имеется перевод: *Конуэлл Э.* Кинетические свойства полупроводников в сильных электрических полях.— М.: Мир, 1970.]
Nag B. R. Electron Transport in Compound Semiconductors, Springer, 1980.
Putley E. H. The Hall Effect and Related Phenomena, Butterworth, 1960.

Явления переноса в диэлектриках и некристаллических материалах

Physical Properties of Amorphous Materials, eds. D. Adler, B. B. Schwartz, M. C. Steele, Plenum Press, 1985.
Amorphous Semiconductors, ed. M. H. Brodsky, Springer, 1979. [Имеется перевод: Аморфные полупроводники/Под ред. М. Бродски.— М.: Мир, 1982.]
Cutler M. Liquid Semiconductors, Academic Press, 1977.
Lampert M. A., Mark P. Current Injection in Solids, Academic Press, 1970. [Имеется перевод: *Ламперт М., Марк П.* Инжекционные токи в твердых телах.— М.: Мир, 1973.]

Mott N. F. Metal-Insulator Transitions, Taylor and Francis, 1974. [Имеется перевод: *Мотт Н. Ф.* Переходы металл — изолятор.— М.: Наука, 1979.]
Mott N. F., Devis E. A. Electronic Processes in Non-Crystalline Materials, Oxford Univ. Press, 2nd ed., 1979. [Имеется перевод: *Мотт Н., Дэвис Э.* Электронные процессы в некристаллических веществах. Т. 1, 2.— М.: Мир, 1982.]
O'Dwyer J. J. The Theory of Electrical Conduction and Breakdown in Solid Dielectrics, Oxford Univ. Press, 1973.
Таис Дж. Amorphous and Liquid Semiconductors, Plenum Press, 1974.

Энергетические зоны в полупроводниках

Cardona M. Modulation Spectroscopy, Academic Press, 1969. [Имеется перевод: *Кардона М.* Модуляционная спектроскопия.— М.: Мир, 1972.]
Long D. Energy Bands in Semiconductors, Wiley-Interscience, 1968.
Phillips J. C. Bands and Bonds in Semiconductors, Academic Press, 1973.
* *Цидильковский И. М.* Электроны и дырки в полупроводниках.— М.: Наука, 1972.
(См. также раздел «Зонная теория» в литературе к гл. 3).

Оптические и неравновесные явления

Basov N. G. Optical Properties of Semiconductors, Plenum Press, 1976.
Blakemore J. S. Semiconductor Statistics, Pergamon Press, 1962. [Имеется перевод: *Блейкмор Д.* Статистика электронов в полупроводниках.—М.: Мир, 1964.]
Bube R. H. Photoconductivity of Solids, Wiley, 1960. [Имеется перевод: *Бьюб Р.* Фотопроводимость твердых тел.— М. ИЛ, 1962.]
Electron-Hole Droplets in Semiconductors, eds. C. D. Jeffries, L. V. Keldysh, North-Holland, 1983.
Pankove J. I. Optical Processes in Semiconductors, Dover, 1975.
Rose A. Concepts in Photoconductivity and Allied Problems, Wiley, 1963. [Имеется перевод: *Роуз А.* Основы теории фотопроводимости.— М.: Мир, 1966.]
Рывкин С. М. Фотоэлектрические явления в полупроводниках.— М.: Физматгиз, 1963.

Процессы в приборных полупроводниковых структурах

(См. также раздел «Электронные свойства твердых тел» на с. 12).
Brodie I., Murray J. J. The Physics of Microfabrication, Plenum Press, 1982.
Casey H. C., Panish M. B. Heterostructure Lasers, Academic Press, 1978, 2 vols.
Fraser D. A. Physics of Semiconductor Devices, Oxford Univ. Press, 3rd ed., 1983.
The Physics of Submicron Structures, eds. H. L. Grubin, K. Hess, G. J. Lafrate, Plenum Press, 1984.
Henish H. R. Semiconductor Contacts, Oxford Univ. Press, 1984.
Devices for Optical Communication, ed. H. Kressel, Springer, 2nd ed., 1982.
Milnes A. G., Feucht D. Heterojunctions and Metal-Semiconductor Junctions, Academic Press, 1972. [Имеется перевод: *Милнс А., Фойхт Д.* Гетеропереходы и переходы металл—полупроводник.— М.: Мир, 1975.]
Solymar L., Walsh D. Lectures on the Electrical Properties of Solids, Oxford Univ. Press, 3rd ed., 1984.
Thompson G. B. Physics of Semiconductor Laser Devices, Wiley, 1980.
Till W. C., Luxon J. T. Integrated Circuits: Materials, Devices, and Fabrication, Addison-Wesley, 1982.
Williams E. W., Hall R. Luminescence and the Light Emitting Diode, Pergamon Press, 1978.

5

Диэлектрические и магнитные свойства твердых тел

В этой главе на атомном уровне описываются явления, которые определяют макроскопические диэлектрические и магнитные свойства твердых тел. Обсуждение некоторых вопросов дается здесь кратко по двум причинам. Первая связана с существенным отличием рассматриваемых свойств от тех, которым посвящены предыдущие четыре главы. Ранее мы практически постоянно имели дело со следствиями, вытекающими из периодичности в реальном и **k**-пространствах. Дальний порядок в кристаллической решетке существен и при обсуждении сегнетоэлектрических, ферромагнитных или антиферромагнитных свойств. Однако основная часть материала настоящей главы в меньшей степени связана с периодичностью в **k**-пространстве.

Автор ограничился кратким изложением многих аспектов диэлектрических и магнитных явлений еще и для того, чтобы сохранить объем книги в разумных пределах, для изучения ее в течение одного семестра. Соответственно некоторые вопросы излагаются чисто описательно, при этом на случай необходимости их более глубокого изучения в конце главы приведена соответствующая библиография.

При изложении предполагается, что читатель уже обладает основными знаниями по электромагнитной теории, т. е. по электростатике, магнитостатике и электродинамике сплошных сред, и поэтому мы начнем рассмотрение с уравнений Максвелла, а затем перейдем к локальным полям в твердом теле. Это дает возможность изложить на атомном уровне явления поляризации, возникающие при приложении внешнего поля. В первом разделе данной главы рассматривается реакция диэлектриков на внешнее электрическое поле, которая проявляется в виде дипольной ориентации и ионной и электронной поляризуемости. Далее кратко упоминается о пьезоэлектрических, сегнетоэлектрических и антисегнетоэлектрических явлениях.

Второй раздел посвящен магнитным свойствам; при этом вначале рассматриваются присущий всем веществам слабый диамагнетизм и парамагнетизм невзаимодействующих магнит-

ных диполей. Ферромагнетизм обсуждается в рамках представлений теории молекулярного поля и обменных сил, причем последние используются для описания качественной картины антиферромагнетизма и ферримагнетизма. В третьем и последнем разделе этой главы, добавленном в новом издании, излагаются основы ядерного магнитного резонанса (ЯМР) и электронного спинового резонанса (ЭСР). Последнее явление также хорошо известно как ЭПР (электронный парамагнитный резонанс) и эти два названия эквивалентны *.

5.1. Диэлектрические свойства

Основные законы электромагнетизма

В системе единиц СИ[1] вектор электрического поля **E** и вектор электрического смещения (электрической индукции) **D** для изотропной диэлектрической среды связаны соотношением

$$\mathbf{D} = \varepsilon_0 \mathbf{E} + \mathbf{P} = \varepsilon_0 (1 + \chi_e) \mathbf{E} = \varepsilon_0 \varkappa \mathbf{E}. \tag{5.1}$$

Здесь **P** — поляризация на единицу объема, χ_e — диэлектрическая восприимчивость, $\varkappa$ — относительная диэлектрическая проницаемость (диэлектрическая постоянная), $\varepsilon_0 = (10^7/4\pi c^2) = 8{,}8542 \cdot 10^{-12}$ Ф/м — электрическая постоянная вакуума. Уравнение (5.1) записано для изотропной среды. В случае анизотропной среды **P** и **E** могут быть непараллельны, а восприимчивость и диэлектрическая постоянная являются тензорами.

Аналогичное соотношение между напряженностью магнитного поля **H** и соответствующей магнитной индукцией **B** записывается в виде

$$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}) = \mu_0 (1 + \chi_m) \mathbf{H} = \mu_0 \mu_m \mathbf{H} = \mu \mathbf{H}. \tag{5.2}$$

Здесь **M** — намагниченность, χ_m — магнитная восприимчивость, μ_m — относительная, а μ — абсолютная магнитные проницаемости, $\mu_0 = 4\pi \cdot 10^{-7}$ Гн/м — магнитная проницаемость вакуума. В разд. 5.2 для обозначения величины магнитного дипольного момента нам понадобится символ μ. Свойства магнитных

---

* В нашей отечественной литературе чаще встречается аббревиатура ЭПР. Поэтому в настоящем переводе книги используется именно это сокращение. — *Прим. перев.*

[1] Большая часть литературы по результатам экспериментальных исследований диэлектрических явлений и в особенности магнитных явлений использует гауссову систему единиц. Изложение теории электрических и магнитных явлений в системе СИ можно найти, например, в монографии: *Cheston W. B.* Elementary Theory of Electric and Magnetic Fields, Wiley, 1964.

материалов мы будем описывать с использованием либо намагниченности **M**, либо магнитной восприимчивости χ_m.

Уравнения (5.1) и (5.2) согласуются с записью уравнений Максвелла в системе СИ (рационализованная система МКСА)

$$\nabla \cdot \mathbf{D} = \rho, \tag{5.3а}$$

$$\nabla \times \mathbf{E} = -(\partial \mathbf{B}/\partial t), \tag{5.3б}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{5.3в}$$

$$\nabla \times \mathbf{H} = \mathbf{J} + (\partial \mathbf{D}/\partial t), \tag{5.3г}$$

где **J** — плотность электрического тока, ρ — плотность электрического заряда. Магнитный аспект уравнений (5.2) и (5.3) рассмотрим в разд. 5.2. Здесь же мы сосредоточим внимание на диэлектрических свойствах, предполагая, что имеем дело с некоторым твердым телом, магнитная восприимчивость которого $\chi_m \ll 1$.

Макроскопическая и микроскопическая теории диэлектрических свойств

Естественно, уравнение (5.3а) полностью согласуется с законом Кулона

$$\mathbf{E} = (q\mathbf{r}/4\pi\varepsilon_0 r^3), \tag{5.4}$$

описывающим электрическое поле, создаваемое в вакууме на расстоянии **r**, изолированным зарядом q. [В однородной диэлектрической среде это поле приводится к виду $\mathbf{E} = (q\mathbf{r}/4\pi\varkappa\varepsilon_0 r^3)$.] Диэлектрическая восприимчивость чаще связана с диполями, а не с изолированными зарядами. Для диполя, изображенного на рис. 5.1, определим *дипольный момент* **p** как

$$\mathbf{p} = q\,\mathbf{R}. \tag{5.5}$$

Решение общеизвестной несложной электростатической задачи дает следующее выражение для поля **E**, создаваемого диполем в вакууме:

$$\mathbf{E}(\mathbf{r}) = \frac{3(\mathbf{p}\cdot\mathbf{r})\,\mathbf{r} - r^2\,\mathbf{p}}{4\pi\varepsilon_0 r^5}. \tag{5.6}$$

В задаче 5.1 в качестве простейшего упражнения в векторной геометрии мы предлагаем читателю получить это выражение.

Любая асимметричная молекула, состоящая из атомов с заметно различными электроотрицательностями, обладает постоянным электрическим дипольным моментом. На рис. 5.2 приведены два простых примера. Для сложных и асимметричных молекул могут быть введены моменты более высоких порядков, однако в настоящий момент нас интересуют дипольные мо-

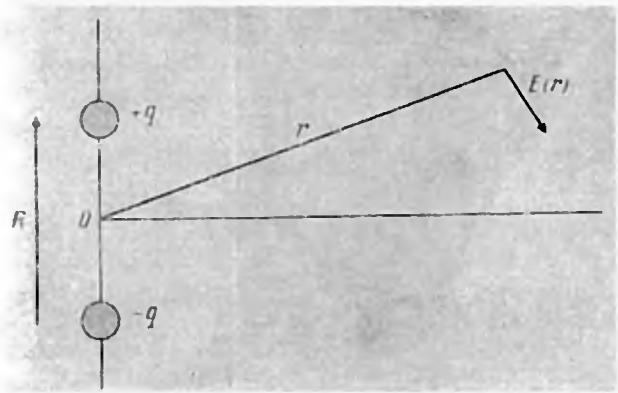


Рис. 5.1. Электрический диполь, состоящий из равных, но противоположных по знаку зарядов $+q$ и $-q$, разделенных в пространстве. Если расстояние между зарядами равно R, то дипольный момент $\mathbf{p}=q\mathbf{R}$. Электрическое поле $\mathbf{E}(\mathbf{r})$, создаваемое таким диполем в некоторой точке, положение которой задается радиус-вектором $\mathbf{r}$, проведенным от центра диполя, определяется путем векторного сложения кулоновских полей каждого из зарядов. Можно проверить, что результатом такого сложения является соотношение (5.6) (см. задачу 5.1).

менты. Диэлектрические свойства среды также могут быть связаны с возникновением *индуцированных* электрических диполей. Введенная в уравнении (5.1) поляризация $\mathbf{P}$ представляет собой не что иное, как полный дипольный момент единицы объема, т. е.

$$\mathbf{P}=\sum_n \mathbf{p}_i=\sum_n q_i\,\mathbf{R}_i. \tag{5.7}$$

Эта интерпретация отражает существование диполей атомных размеров, которое приводит к макроскопическому различию между электрическим смещением и электрическим полем.

В простейшем виде это показано на рис. 5.3 на примере плоского конденсатора. Предположим, что на его пластинах имеются заряды $+Q$ и $-Q$ и что пластина с отрицательным зарядом заземлена. Рис. 5.3, *а* иллюстрирует случай, когда пространство между пластинами — вакуум. Тогда электрометр фиксирует потенциал $V_0=(Qd/\varepsilon_0 A)$ при электрическом поле между пластинами $E_0=(Q/\varepsilon_0 A)$. (Предполагается, что можно не учитывать краевые эффекты, поскольку существует охранное кольцо, которое не показано на рисунке.) Если между пластинами конденсатора поместить диэлектрик (рис. 5.3, *б*), это приведет к уменьшению величины потенциала и среднего поля в $\varkappa$ раз. Причина этого эффекта становится более очевидной при

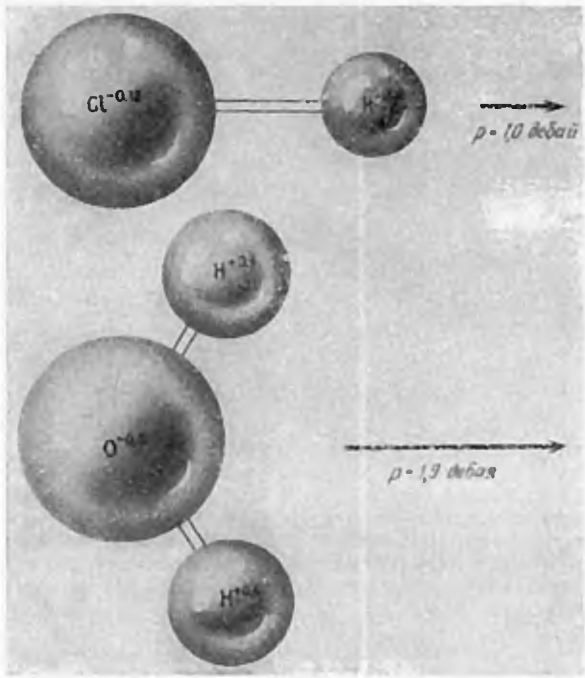


Рис. 5.2. Простые примеры молекул с постоянным электрическим дипольным моментом. В случае молекулы H_2O дипольный момент направлен от иона кислорода к средней точке прямой, соединяющей два иона водорода. Эффективный ионный заряд, полученный, исходя из дипольного момента и линейных размеров, согласуется с разностью электроотрицательностей и относительной ионностью по Полингу (рис. 1.11 и 1.12). Ионность HCl, полученная по дипольному моменту, значительно меньше, чем можно ожидать, исходя из разности электроотрицательностей для водорода и хлора. Заметим, что дипольные моменты этих молекул выражены в единицах системы СГС — *дебаях* (1 дебай $=10^{-18}$ ед. заряда СГСЭ·см $=3{,}33\times$ $\times 10^{-30}$ Кл·м). Эта единица связана с разностью атомных электроотрицательностей, что уже отмечалось при обсуждении равенства (1.21).

рассмотрении рис. 5.3, *в*, на котором показаны только остаточные *индуцированные на поверхности* заряды, возникающие как следствие объемной поляризации. Дипольный момент единицы объема **P** приводит к поверхностной плотности заряда

$$q_s = -\mathbf{P}\cdot\mathbf{n}. \tag{5.8}$$

Здесь **n** — единичный вектор, нормальный к поверхности образца. Теперь мы используем этот результат для того, чтобы определить величину поля в центре полости, окруженной поляризованным диэлектриком.

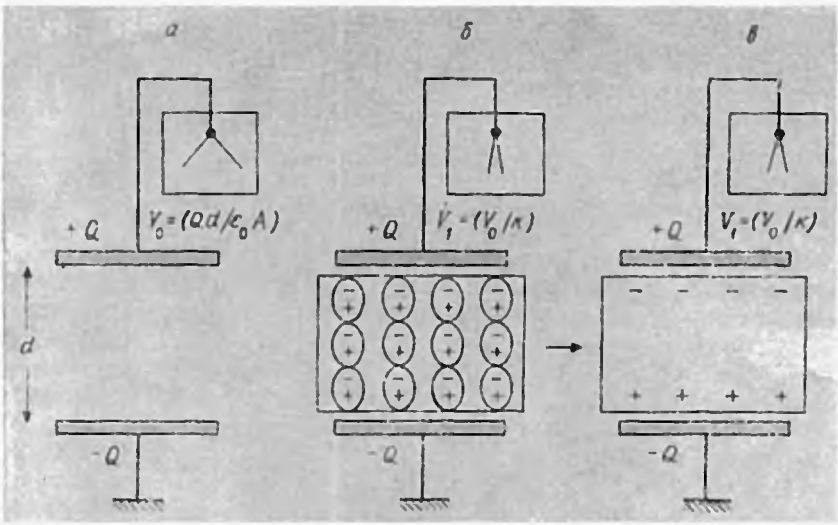


Рис. 5.3. Простая иллюстрация влияния электрических диполей атомных размеров на макроскопическое распределение электрического поля. *а* — плоский конденсатор (одна из его пластин заземлена, потенциал другой пластины измеряется); площадь пластин A, расстояние между ними d, полный заряд пластины Q; *б* — расположение диполей в объеме диэлектрика, помещенного между пластинами; *в* — эквивалент рисунка *б*, на котором изображены полные заряды на поверхности диэлектрика, частично компенсирующие заряды на пластинах, что приводит к ослаблению электрического поля.

Как в теории диэлектрических, так и в теории магнитных явлений, часто оказывается необходимым рассчитать *локальное эффективное поле* на атоме или ионе, определяемое поляризацией. Обычно эта задача решается с помощью метода, предложенного Лоренцем. При этом считается, что рассматриваемый атом окружен воображаемой замкнутой поверхностью. Диполи, расположенные внутри этой поверхности, учитываются каждый в отдельности, а вклад всего остального объема вне поверхности дается в приближении континуума. Таким образом, локальное эффективное поле на атоме в диэлектрике, возникающее под влиянием внешних зарядов, можно записать как сумму четырех слагаемых

$$\mathbf{E}_{\text{лок}} = \mathbf{E}_0 + \mathbf{E}_{\text{деп}} + \mathbf{E}_{\text{пов}} + \mathbf{E}_{\text{дип}}. \qquad (5.9)$$

Здесь $\mathbf{E}_0$ — поле, обусловленное внешними зарядами; $\mathbf{E}_{\text{деп}}$ учитывает деполяризующие эффекты (или любые другие), возникающие за счет индуцированных поверхностных зарядов на внешней поверхности пластины диэлектрика: $\mathbf{E}_{\text{пов}}$ — поле,

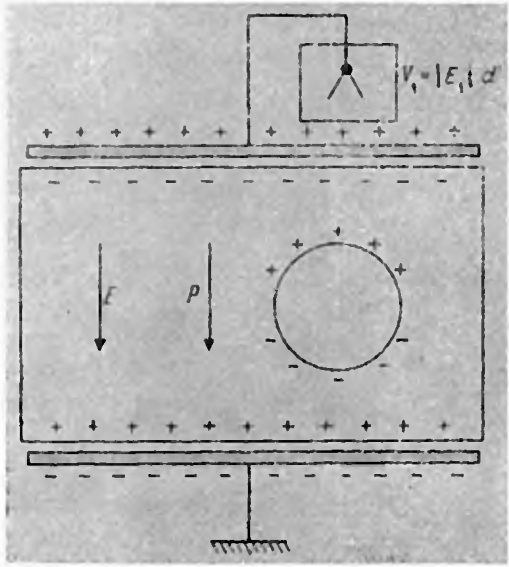


Рис. 5.4. Схема определения величины локального поля в месте расположения атома в поляризованном диэлектрике. В случае кубической решетки поле $\mathbf{E}_{\text{дип}}$ в центре сферической полости равно нулю, а $\mathbf{E}_{\text{пов}} = P/3\varepsilon_0$ для границы сферической полости.

созданное в центре воображаемой полости поверхностными зарядами, наведенными поляризацией на ограничивающую ее поверхность; $\mathbf{E}_{\text{дип}}$ — поле в центре этой полости, возникающее под влиянием отдельных диполей, которые расположены по всем атомным узлам полости, исключая один центральный.

Согласно рис. 5.3, сумма $\mathbf{E}_0 + \mathbf{E}_{\text{деп}}$ равна полю $\mathbf{E}_1$, величина которого точно соответствует (V_1/d). При вычислении локального поля будем считать, что деполяризующие эффекты, обусловленные внешней поверхностью, могут быть учтены с помощью такой же простой геометрии. На рис. 5.4 изображена диэлектрическая пластина с внутренней полостью. Таким образом,

$$\mathbf{E}_{\text{лок}} = \mathbf{E}_1 + \mathbf{E}_{\text{пов}} + \mathbf{E}_{\text{дип}}. \tag{5.10}$$

Влияние поверхностных зарядов, согласно уравнению (5.8), можно учесть путем суммирования по поверхности полости. Для простоты мы выбираем полость сферической формы, и тогда интегрирование (задача 5.2) дает нам

$$\mathbf{E}_{\text{пов}} = (\mathbf{P}/3\varepsilon_0). \tag{5.11}$$

Вторая причина для выбора именно сферической полости связана с тем, что член $\mathbf{E}_{\text{дип}}$ исчезает в случае кубической решетки[2]. В этом случае, согласно методу Лоренца, локальное поле представляется выражением

$$\mathbf{E}_{\text{лок}} = \mathbf{E}_1 + (\mathbf{P}/3\varepsilon_0). \tag{5.12}$$

Вероятно, это поле будет значительно превышать приложенное поле. Используя уравнение (5.1), можем записать

$$\mathbf{P} = \varepsilon_0(\varkappa - 1)\mathbf{E}. \tag{5.13}$$

Теперь, подставляя это соотношение в (5.12), получаем, что локальное поле в атомном узле кубической (изотропной) решетки будет

$$\mathbf{E}_{\text{лок}} = \left(\frac{\varkappa + 2}{3}\right)\mathbf{E}. \tag{5.14}$$

В случае анизотропных твердых тел соответствующие соотношения приобретают более сложный вид, однако, рассматривая изотропные материалы, мы можем получить много полезной информации. В следующем разделе мы обратимся к выражениям (5.12)—(5.14), чтобы получить соотношение *Клаузиуса — Моссотти* между поляризуемостью атомов и ионов и макроскопической диэлектрической постоянной изотропной (т. е. кубической) среды.

Механизмы поляризации в диэлектриках

При обсуждении возможных различий в диэлектрических свойствах твердых тел и вакуума мы рассмотрим три механизма поляризации на атомном уровне.

1. Частичное или полное выстраивание дипольных моментов полярных молекул вдоль локального электрического поля. Как мы уже видели из примеров (см. рис. 5.2), любая асимметричная молекула, состоящая из атомов с различными электроотрицательностями, имеет постоянный дипольный момент. Электрическое поле вызывает повороты таких молекул. Этот процесс называют *дипольной ориентацией* или *параэлектрической восприимчивостью*. Выстраиванию всех постоянных диполей вдоль приложенного электрического поля препятствуют тепловые

[2] Для некубических решеток оценка локального поля оказывается более сложной. Величины вклада $\mathbf{E}_{\text{дип}}$, обусловленного диполями в полости простой формы для тетрагональной и гексагональной решеток выполнены в работах Мюллера: *Mueller H.*— Phys. Rev., **47**, 947 (1935) и Phys. Rev. **50**, 547 (1936). Тот факт, что диполи в сферической полости в кубической решетке вносят нулевой вклад в локальное поле в центре этой полости, обсуждается Джексоном в монографии *Jackson J. D.* Classical Electrodynamics, Wiley, 1962, p. 116. [Имеется перевод: Джексон Дж. Классическая электродинамика.— М.: Мир, 1965.]

колебания. Конкуренцию этих двух процессов мы опишем с помощью функции Ланжевена.

2. Возникновение диполей при относительном смещении положительных и отрицательных ионов под влиянием электрического поля в твердых телах с некоторой долей ионности связи. Этот механизм обычно называется *ионной поляризацией.*

3. Третьим и единственным из механизмов поляризации, который реализуется во *всех* диэлектриках, является процесс *электронной поляризации.* Он состоит в смещении электронов атома относительно его ядра под влиянием электрического поля. Можно сказать, что электрическое поле деформирует электронные оболочки каждого атома; при этом могут произойти некоторые изменения межъядерных расстояний.

В твердом теле, которое содержит постоянные диполи, все три механизма вносят вклад в поляризуемость (и следовательно, в макроскопическую диэлектрическую постоянную) при низких частотах, однако их вклады различны. Если рассматривать диэлектрические свойства при высоких частотах, восприимчивость будет определяться только теми механизмами, которые могут находиться в фазе с быстро изменяющимся электрическим полем. Таким образом, на высоких частотах диэлектрическая постоянная является комплексной величиной, причем ее действительная часть соответствует поляризации диэлектрика, изменяющейся в фазе с приложенным полем. Мнимая часть представляет диэлектрические потери, создаваемые механизмом, который отстает по фазе от приложенного поля при тех частотах, на которых наблюдается дисперсия. Действительная и мнимая части комплексной диэлектрической постоянной для любой среды (изолирующей или какой-либо иной) связаны крайне полезными интегральными зависимостями, известными как *дисперсионные соотношения Крамерса — Кронига*[3].

На рис. 5.5 показано поведение действительной и мнимой частей диэлектрической проницаемости твердого тела, в котором имеются постоянные диполи, способные поворачиваться, и в котором могут быть индуцированы как ионная, так и электронная поляризации. Компонента поляризации, связанная с дипольной ориентацией, не успевает отслеживать быстрые изменения знака электрического поля при частотах более 10^{10} Гц. На рис. 5.5 показано соответствующее уменьшение действительной части

[3] Первое упоминание о соотношениях Крамерса — Кронига можно найти в книге *Kramers H. A.* Collected Scientific, Papers North-Holland, 1956. Соотношение между диэлектрическими и оптическими свойствами твердых тел обсуждаются в гл. 7 и 8 в книге: *Brown F. C.*— Physics of Solids, Benjamin, 1967. Помимо изолирующих твердых тел, дисперсионные соотношения можно применить и к спектрам отражения металлов и полупроводников для определения критических энергий сингулярностей Ван Хова в плотности состояний.

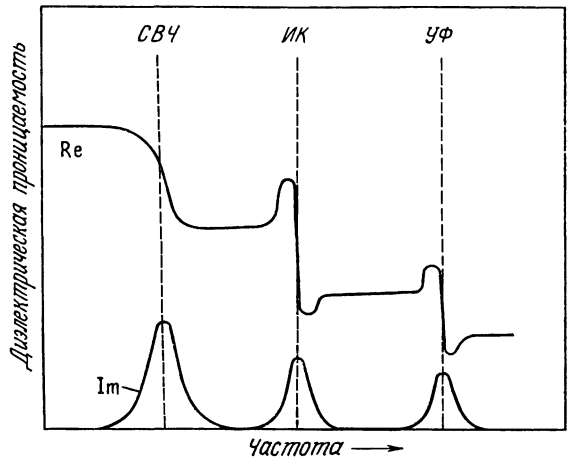


Рис. 5.5. Зависимость действительной (Re) и мнимой (Im) частей комплексной диэлектрической проницаемости от частоты приложенного электрического поля для ионного твердого тела, содержащего постоянные диполи. Мнимая часть диэлектрической проницаемости (коэффициент диэлектрических потерь) возрастает всякий раз с приближением к верхнему частотному пределу для каждого конкретного механизма диэлектрической поляризуемости. В случае дипольной ориентации это возрастание имеет место в микроволновой области (СВЧ), для ионной поляризуемости — в инфракрасной области (ИК), а для смещения связанных электронов — в ультрафиолетовой области (УФ).

диэлектрической проницаемости в микроволновой области спектра. При этом в той же области спектра, мнимая часть диэлектрической проницаемости оказывается отличной от нуля (т. е. появляются значительные диэлектрические потери), поскольку диполи пытаются реагировать на изменения полярности поля, но отстают от них по фазе. При еще больших частотах амплитуды поворотов диполей становятся вообще пренебрежимо малыми.

Для твердых тел с ионной или частично ионной химической связью дальнейшее уменьшение действительной части диэлектрической проницаемости наблюдается во второй спектральной области диэлектрических потерь, которая имеет место при частотах порядка 10^{13} Гц. Это область частот остаточных лучей, которая обсуждалась ранее в разд. 2.3 и иллюстрируется на рисунках 2.14—2.16. Выше этой частоты смещения ионов не могут следовать за изменениями приложенного поля. При более высоких частотах лишь за счет индуцированной электронной поляризации действительная часть диэлектрической проницаемости оказывается больше единицы. Однако при частотах много больших 10^{15} Гц и этот механизм также не может не

отставать от поля. Выше этой частоты твердое тело имеет диэлектрическую проницаемость, очень близкую к единице.

Удобно выразить действительную часть диэлектрической восприимчивости через атомные поляризуемости. Эта задача оказывается достаточно простой в случае кубической (и следовательно, изотропной) среды. При этом мы используем несколько сомнительное приближение, которое заключается в том, что существует локальное поле $\mathbf{E}_{\text{лок}}$ [см. уравнение (5.12)], поляризующее атомы или ионы диэлектрика. Очевидно, такое представление не может считаться вполне корректным, так как понятие $\mathbf{E}_{\text{лок}}$ вводится без учета именно того атома, поляризацию которого мы обсуждаем. В действительности же должна существовать некоторая обратная связь между индуцированным дипольным моментом и локальным полем. Тем не менее определим поляризуемость α_j атомов j-го типа твердого тела через его дипольный момент в виде

$$\mathbf{p} = \alpha_j \mathbf{E}_{\text{лок}}. \tag{5.15}$$

Тогда полная поляризуемость единицы объема

$$\mathbf{P} = \mathbf{E}_{\text{лок}} \sum_r N_j \alpha_j, \tag{5.16}$$

где суммирование проводится по различным типам поляризуемых атомов, а N_j — объемная плотность атомов j-го типа.

Теперь, используя соотношение (5.14), получим связь между полной поляризуемостью и приложенным полем

$$\mathbf{P} = \left(\frac{\varkappa + 2}{3}\right) \mathbf{E} \sum_r N_j \alpha_j. \tag{5.17}$$

Если мы перепишем это уравнение в виде

$$\sum_r N_j \alpha_j = 3\,\mathbf{P}/(\varkappa + 2)\,\mathbf{E} \tag{5.18}$$

и используем уравнение (5.13), то получим, что

$$\sum_r N_j \alpha_j = 3\varepsilon_0 \left(\frac{\varkappa - 1}{\varkappa + 2}\right). \tag{5.19}$$

Это выражение известно как *соотношение Клаузиуса — Мосотти* между атомной поляризацией и макроскопической диэлектрической проницаемостью. Необходимо помнить, что это соотношение получено для твердого тела, которое кристаллизуется в кубической решетке. Для менее изотропных твердых тел связь между атомными и макроскопическими параметрами оказывается более сложной.

Дипольная ориентация

Приведенное здесь рассмотрение поляризуемости, возникающей в результате поворотов постоянных диполей скорее соответствует ситуации, которая характерна для газов и молекул. Полярные молекулы в твердых телах не могут свободно поворачиваться в ответ на воздействие электрического поля. Более корректно такой процесс может быть описан как отдельные резкие скачки в угловой ориентации молекул из одного стабильного положения в другое, при которых угол между дипольным моментом и электрическим полем уменьшается. Кратко рассмотрим изменения в угловой ориентации, допускаемые симметрией кристалла, и время релаксации, которое определяет частоту скачков. Однако сначала определим диэлектрическую восприимчивость, предполагая, что молекулы *могут* поворачиваться свободно. Приближение, которое мы будем использовать, развито и используется до сих пор в теории парамагнетизма при описании ориентации магнитных диполей.

Рассмотрим некоторую среду с N постоянными электрическими диполями в единице объема, каждый из которых имеет момент p. В отсутствие внешнего поля диполи имеют случайную ориентацию. Теперь предположим, что статическое поле $\mathbf{E}$ стремится упорядочить диполи. Тогда поляризация единичного объема может быть записана как

$$P_{dp} = \sum_{N} p \cos\theta = Np\langle\cos\theta\rangle. \tag{5.20}$$

Здесь θ — угол между направлением каждого диполя и приложенным полем.

Процессу упорядочения диполей препятствует тепловое движение частиц среды. Допустим, что тепловое движение можно описать больцмановским распределением по энергиям. Тогда ожидаемое значение $\langle\cos\theta\rangle$ определяется отношением двух больцмановских интегралов, т. е.

$$\langle\cos\theta\rangle = \frac{\int_0^{\pi} 2\pi \sin\theta\cos\theta \exp(-U/k_0T)\, d\theta}{\int_0^{\pi} 2\pi \sin\theta \exp(-U/k_0T)\, d\theta}, \tag{5.21}$$

где U — потенциальная энергия диполя в поле $\mathbf{E}$ для угла θ;

$$U = -\mathbf{p}\cdot\mathbf{E} = -pE\cos\theta. \tag{5.22}$$

Интегралы уравнения (5.21) вычисляются проще, если произвести подстановки

$$y = (-U/k_0T) = +(pE/k_0T)\cos\theta,$$
$$w = (pE/k_0T). \tag{5.23}$$

Тогда уравнение (5.21) принимает вид

$$\langle \cos\theta \rangle = \frac{\int\limits_{-w}^{w} y/w) \exp(y)\, dy}{\int\limits_{-w}^{w} \exp(y)\, dy}. \qquad (5.24)$$

Числитель легко вычисляется путем интегрирования по частям и

$$\langle \cos\theta \rangle = \left[\frac{1+\exp(-2w)}{1-\exp(-2w)} - \frac{1}{w}\right] = \left[\operatorname{cth} w - \frac{1}{w}\right] \equiv L(w). \qquad (5.25)$$

Функция $L(w)$, определяемая выражением (5.25), введена Ланжевеном в 1905 г. при создании теории парамагнитной восприимчивости и мы опять встретимся с ней в разд. 5.2. Заметим, что при описании параэлектрической восприимчивости для больших значений w, т. е. когда приложенное поле E велико по сравнению с (k_0T/p), эта функция стремится к единице. Как показано на рис. 5.6, в пределе слабых полей функция Ланжевена принимает вид

$$\langle \cos\theta \rangle \equiv L(pE/k_0T) \approx (pE/3k_0T), \qquad E \ll (k_0T/p). \qquad (5.26)$$

В задаче 5.3 доказывается, что эта аппроксимация функции Ланжевена справедлива при обычных температурах и для случая достаточно больших электрических полей. При обычных температурах электрическое поле, для которого соотношение (5.26) перестает быть справедливым, оказывается настолько большим, что у значительной части твердых тел наступает электрический пробой [4].

Таким образом, полная поляризация единицы объема, возникающая вследствие переориентации постоянных диполей, есть

$$P_{dp} = Np^2E/3k_0T, \qquad (5.27)$$

[4] В идеальных условиях характерная для твердого диэлектрика прочность на пробой составляет $\sim 10^9$ В/м. Превышающее это значение электрическое поле неизбежно приводит к явлению умножения числа свободных электронов (лавинному пробою) или к тепловому пробою, так как в любом твердом диэлектрике всегда имеется *некоторое* количество свободных электронов и *несколько* источников джоулева тепла. Даже при полном отсутствии свободных электронов в равновесном состоянии в полях, превышающих некоторое критическое значение, благодаря процессу туннелирования Зинера через широкую запрещенную зону диэлектрика образуются электронно-дырочные пары. Для любого реального диэлектрика поле пробоя всегда меньше, чем критическое значение, поскольку существование любых незначительных неоднородностей приводит к концентрации огромных напряжений в малых областях твердого тела. Различные явления пробоя обсуждаются в книге: *O'Dwyer J.* The Theory of Electrical Conduction and Breakdown in Solid Dielectrics, Oxford University Press, 1973.

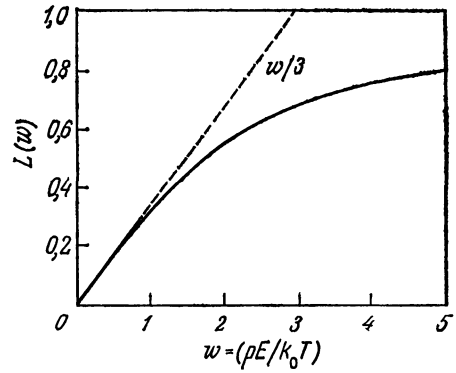


Рис. 5.6. Функция Ланжевена $L(w)$, используемая для описания ориентации диполей в параэлектрических и парамагнитных явлениях.

а соответствующая восприимчивость

$$\chi_{dp} = (P_{dp}/\varepsilon_0 E) = Np^2/3\varepsilon_0 k_0 T. \qquad (5.28)$$

Этот вклад в полную поляризацию мал по сравнению с единицей для газов, но для полярных жидкостей или твердых тел, в которых возможна переориентация диполей, он может быть сравним с единицей. Таким образом в материале с $N = 10^{28}$ полярных молекул в кубическом метре, каждая из которых обладает постоянным моментом в 1 дебай, восприимчивость $\chi_{dp} \approx \approx (300/T)$.

Безусловно, *индуцированные* ионные диполи и неизбежно возникающая электронная поляризация также вносят вклад в статическую диэлектрическую проницаемость, которую нужно записать в виде

$$\varkappa = 1 + \chi_{dp} + \chi_i + \chi_e. \qquad (5.29)$$

Здесь χ_{dp} определяется соотношением (5.28). Так как χ_i и χ_e не зависят от температуры, величину момента постоянных диполей можно определить из измерений температурной зависимости статической (или низкочастотной) диэлектрической проницаемости.

Рис. 5.7 на примере сероводорода иллюстрирует поведение низкочастотной диэлектрической проницаемости в реальных полярных твердых телах. В интервал температур, представленный на этом рисунке, попадает и точка плавления $T_{пл}$ сероводорода. Вклад в полную диэлектрическую проницаемость, связанный с ориентацией диполей, *возрастает* при охлаждении. Но это возрастание не является монотонным, поскольку диполи не имеют возможности вращаться совершенно свободно; скорее кристаллическая решетка допускает перескоки из одного ограниченного набора возможных ориентаций в другой. Вероятно, при

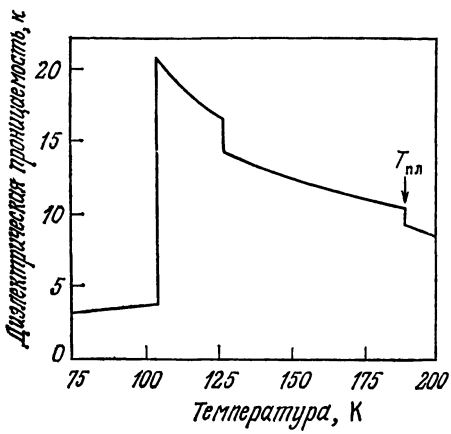


Рис. 5.7. Температурная зависимость диэлектрической проницаемости H_2S, измеренная при низких частотах (5 кГц) [*Smyth C. P., Hitchcock C. S.*— J. Am. Chem. Soc., **56**, 1084 (1934)].

температурах ниже 103 К возможность для совершения таких дискретных ориентационных перескоков исчезает вообще.

Предположим, что энергия полярной молекулы в кристалле (в отсутствие внешнего поля) может быть представлена верхней кривой рис. 5.8. По отношению к определенному направлению, вдоль которого позднее будет приложено электрическое поле, имеются два положения равновесия с углами θ_1 и θ_2. Для любой другой ориентации молекулы необходима слишком большая энергия. Перескоки между двумя возможными ориентациями требуют преодоления барьера высотой Y. Скорость, с которой происходят такие перескоки, может быть охарактеризована *постоянной времени релаксации* τ, которая выражается как

$$\tau = (2\pi\nu_D)^{-1} \exp(Y/k_0T), \tag{5.30}$$

где ν_D — собственная частота перескоков, выражаемая через максимальную частоту в спектре колебаний кристаллической решетки, т. е.

$$\nu_D \approx k_0\theta_D/2\pi\hbar. \tag{5.31}$$

Обычно ν_D лежит в интервале от 10^{11} до 10^{13} Гц. Однако постоянная времени τ может быть существенно больше, чем 10^{-11} с, если высота барьера Y значительно превышает k_0T.

При приложении внешнего электрического поля энергетические минимумы перестают быть эквивалентными и переходы способствуют повышению заселенности одного из них за счет уменьшения заселенности другого. Если поле приложено в течение времени, достаточного для установления стационарного состояния, отношение заселенностей для ориентаций 1 и 2, со-

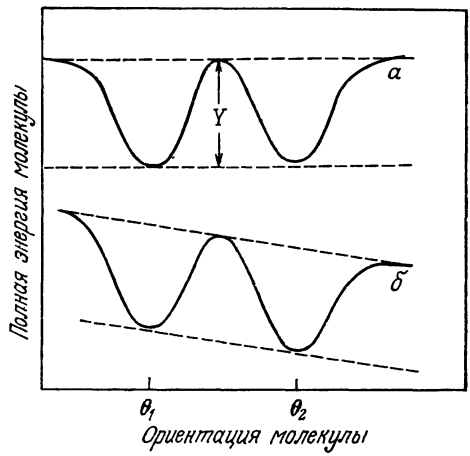


Рис. 5.8. *а* — возможная зависимость энергии полярной молекулы в твердом теле от ее ориентации в отсутствие внешнего поля; *б* — изменение относительного положения минимумов энергии внешним электрическим полем.

гласно соотношению (5.22) и закону распределения Больцмана, будет

$$\frac{N_1}{N_2} = \exp\left[\frac{\cos\theta_2 - \cos\theta_1}{k_0 T/pE}\right]. \qquad (5.32)$$

Разность заселенностей (N_2—N_1) приводит к появлению объемной поляризации, которую мы ранее описывали уравнением (5.27), допуская возможность свободных поворотов.

Теперь допустим, что мы прикладываем переменное электрическое поле, частота которого настолько велика, что перескоки постоянных диполей не вполне успевают его отслеживать [т. е. частота поля сравнима с величиной τ^{-1}, которая определяется выражением (5.30)]. Однако частота приложенного поля оказывается не настолько большой, чтобы наведенные ионная и электронная поляризации не смогли бы следовать за ним. Как впервые было показано в 1929 г. Дебаем, в этом случае диэлектрическую проницаемость следует записать как комплексную величину

$$\varkappa = (\varkappa_1 + i\varkappa_2) = A + \frac{B}{1 - i\omega\tau}. \qquad (5.33)$$

В рассматриваемом интервале частот *индуцированные* ионная и электронная поляризации не зависят от частоты и не вносят вклад в диэлектрические потери; с ними связано первое слагаемое в правой части уравнения (5.33). Запаздывание при повороте постоянных диполей приводит к появлению второго слагаемого в соотношении (5.33); при этом величина B зависит от

плотности постоянных диполей и от того, насколько возможно для них вращательное движение. Отметим, что соотношение (5.28) описывает χ_{dp} для случая свободного вращения диполей в жидкости.

Действительная и мнимая части диэлектрической проницаемости могут быть представлены соответственно в виде

$$\varkappa_1 = \left[A + \frac{B}{1+\omega^2\tau^2}\right], \tag{5.34}$$

$$i\varkappa_2 = \frac{i\omega\tau B}{1+\omega^2\tau^2}. \tag{5.35}$$

Выражения (5.34) и (5.35) обычно называют *уравнениями Дебая*. Уменьшение действительной части диэлектрической проницаемости $\varkappa_1$ с ростом частоты, которое имеет место при частотах порядка обратного времени релаксации, определенного равенством (5.30), на верхней кривой рис. 5.5 представлено первой ступенью. Расположенный под ней первый пик на частотной зависимости мнимой части соответствует выражению (5.35).

Эффекты, связанные с тем, что частота ориентационных скачков постоянных диполей ограничена, проявляются при измерениях на фиксированной частоте как уменьшение $\varkappa_1$ при охлаждении. При этом становятся ощутимыми диэлектрические потери. Связь между частотной зависимостью при данной температуре и температурной зависимостью при данной частоте для диэлектрической проницаемости исследуется ниже в задаче 5.4.

В качестве заключительного замечания по ориентации постоянных диполей добавим, что ее рассмотрение с точки зрения параэлектрической восприимчивости намеренно проводилось без использования соотношения Клаузиуса — Мосотти для поляризуемости. Причина этого кроется в том, что метод Лоренца для вычисления «эффективного локального поля» пригоден для индуцированных диполей и не пригоден для постоянных диполей. Соотношение (5.14) получено в предположении, что каждый индуцированный диполь направлен вдоль внешнего поля, в то время как большинство постоянных диполей в полярных твердых телах имеет целый набор направлений. Вероятно, *локальное поле*, которое оказывается *эффективным* для ориентации диполей, будет мало отличаться от внешнего приложенного поля.

Теперь нам необходимо вернуться к ситуации с индуцированными электронными и ионными диполями, при рассмотрении которой мы вполне можем использовать представление о локальном поле Лоренца.

Электронная поляризуемость

В отличие от поведения постоянных диполей, которое имеет релаксационный характер, диэлектрические свойства индуцированных диполей относятся к числу резонансных явлений. Рассмотрим электрон, смещающийся под действием синусоидального внешнего поля, которое представим через локальное поле Лоренца (соотношение (5.14)). Если постоянную возвращающей силы обозначить через β, то уравнение можно записать в следующем виде:

$$m\frac{d^2x}{dt^2}+\beta x = eE_{\text{лок}} = eE_{\text{амп}}\exp(i\omega t). \tag{5.36}$$

Это уравнение движения для гармонического осциллятора с собственной угловой частотой $\omega_0=(\beta/m)^{1/2}$. Поэтому его можно переписать в виде

$$m\frac{d^2x}{dt^2}+m\omega_0^2 x = eE_{\text{лок}}. \tag{5.37}$$

Обычное решение для амплитуды вынужденного колебательного движения с угловой частотой ω имеет вид

$$x_{\text{амп}} = \frac{eE_{\text{амп}}}{m(\omega_0^2-\omega^2)}, \tag{5.38}$$

что соответствует дипольному моменту с амплитудой $p=ex_{\text{амп}}$ или электронной поляризуемости [согласно равенству (5.15)]

$$\alpha_e = ex/E_{\text{лок}} = \frac{e^2}{m(\omega_0^2-\omega^2)}. \tag{5.39}$$

Вклад этого процесса в диэлектрическую проницаемость одинаков для всех частот, много меньших, чем $(\omega_0/2\pi)$. В видимой (оптической) области индуцированная электронная поляризация является единственным механизмом, благодаря которому диэлектрическая постоянная $\varkappa$ и показатель преломления $n=\sqrt{\varkappa}$ отличаются от единицы.

Величина показателя преломления может быть использована для определения электронной поляризуемости и, следовательно, резонансной частоты ω_0. В видимой спектральной области соот-

ношение Клаузиуса — Мосотти (5.19) можно записать в следующей форме:

$$\alpha_e = \frac{3\varepsilon_0}{N_e}\left(\frac{\varkappa - 1}{\varkappa + 2}\right) = \frac{3\varepsilon_0}{N_e}\left(\frac{n^2 - 1}{n^2 + 2}\right). \tag{5.40}$$

В реальных твердых телах типичное значение показателя преломления $n = 1{,}7$, а плотность электронов на внешних оболочках атомов $N_e = 10^{29}$ м$^{-3}$, что приводит к значению $\alpha_e \simeq$ $\simeq 10^{-40}$ Ф·м^2. Подставив это значение в соотношение (5.39), можем получить, что $\omega_0 \sim 1{,}7 \cdot 10^{16}$ рад/с. Этой частоте соответствует длина волны 110 нм, которая попадает в вакуумную ультрафиолетовую область электромагнитного спектра.

Согласно уравнению (5.39), поляризуемость должна иметь некоторую особенность при собственной резонансной частоте ω_0, но реальная диэлектрическая проницаемость не достигает ни $+\infty$ ни $-\infty$. Отсутствие этой особенности связано с диэлектрическими потерями, которые мы не учли в уравнении (5.36). Так как колеблющийся электрон периодически ускоряется или замедляется, он должен излучать энергию с интенсивностью, пропорциональной (dx/dt). Потери возникают и вследствие неупругих столкновений электронов; вероятность этого процесса также пропорциональна (dx/dt). Таким образом, уравнение движения электрона должно включать в себя некоторый диссипативный *вязкостный член*, который содержит *постоянную диссипации* γ. Можно сделать так, что величина γ будет иметь размерность частоты. Тогда полное уравнение движения электрона запишется в виде

$$m\frac{d^2x}{dt^2} + m\gamma\frac{dx}{dt} + \beta x = eE_{\text{лок}},$$

или

$$m\frac{d^2x}{dt^2} + m\gamma\frac{dx}{dt} + m\omega_0^2 x = eE_{\text{лок}}. \tag{5.41}$$

Это уравнение также имеет решение для x, изменяющегося синусоидально в соответствии с изменением локального поля, но теперь решение имеет вид

$$x = \frac{eE_{\text{лок}}}{m(\omega_0^2 - \omega^2 - i\gamma\omega)}. \tag{5.42}$$

Таким образом, электронная поляризуемость является комплексной величиной:

$$\alpha_e = \frac{ex}{E_{\text{лок}}} = \frac{e^2}{m(\omega_0^2 - \omega^2 + i\gamma\omega)}. \tag{5.43}$$

Теперь для получения комплексной диэлектрической проницаемости используем соотношение Клаузиуса — Мосотти (5.19). Это приводит к следующему выражению:

$$\varkappa = \varkappa_1 + i\varkappa_2 = \left[1 + \frac{e^2 N_e}{\varepsilon_0 m\left(\omega_0^2 - \omega^2 + i\gamma\omega\right) - (e^2 N_e/3)}\right], \quad (5.44)$$

получить которое мы предлагаем читателю в задаче 5.5. Очевидно, большая плотность поляризуемых электронов вызывает сдвиг резонансной частоты, что отражает следующее соотношение:

$$\omega_1 = \left[\omega_0^2 - (e^2 N_e/3m\varepsilon_0)\right]^{1/2}. \quad (5.45)$$

Используя величину ω_1, выразим действительную (совпадающую по фазе) и мнимую (не совпадающую по фазе) составляющие диэлектрической проницаемости:

$$\varkappa_1 = \left[1 + \frac{(e^2 N_e/m\varepsilon_0)\left(\omega_1^2 - \omega^2\right)}{\left(\omega_1^2 - \omega^2\right)^2 + \gamma^2\omega^2}\right], \quad (5.46)$$

$$i\varkappa_2 = i\left[\frac{(e^2 N_e/m\varepsilon_0)\,\gamma\omega}{\left(\omega_1^2 - \omega^2\right)^2 + \gamma^2\omega^2}\right]. \quad (5.47)$$

Эти уравнения описывают зависимости, иллюстрируемые рис. 5.5 для ультрафиолетовой области. Для действительной части диэлектрической проницаемости здесь представлена последняя осцилляция, после которой для частот, превышающих ω_1, величина $\varkappa_1$ оказывается очень близкой к единице. В задаче 5.5 читателю предлагается исследовать поведение диэлектрической проницаемости в этой области частот.

Ионная поляризуемость

Вклад индуцированных ионных диполей в диэлектрическую проницаемость твердых тел с ионной связью можно рассмотреть по аналогии с вкладом индуцированной электронной поляризации. Некоторые аспекты этой проблемы мы уже обсуждали в разд. 2.3 в связи с взаимодействием электромагнитных волн с колебаниями решетки при частоте остаточных лучей ω_t.

Рассмотрим твердое тело с ионной связью, в котором имеется N_e поляризующихся электронов и N_i поляризующихся ионных пар в кубическом метре. Уравнение Клаузиуса — Мосотти позволяет записать связь между статической величиной диэлектрической проницаемости $\varkappa_0$ и поляризуемостями в виде

$$3\varepsilon_0\left(\frac{\varkappa_0 - 1}{\varkappa_0 + 2}\right) = N_i\alpha_i + N_e\alpha_e. \quad (5.48$$

Теперь, как и в соотношении (2.43) разд. 2.3, используем символ $\varkappa_\infty$ для обозначения диэлектрической проницаемости при частотах, слишком высоких для какого-либо проявления вклада индуцированных ионных диполей, но недостаточно высоких для какого-либо уменьшения электронной компоненты. Тогда

$$3\varepsilon_0\left(\frac{\varkappa_\infty - 1}{\varkappa_\infty + 2}\right) = N_e\alpha_e. \tag{5.49}$$

Таким образом, если известны $\varkappa_0$ и $\varkappa_\infty$, ионная поляризуемость может быть определена из разности соотношений (5.48) и (5.49):

$$\alpha_i = (3\varepsilon_0/N_i)\left[\frac{(\varkappa_0 - 1)}{(\varkappa_0 + 2)} - \frac{(\varkappa_\infty - 1)}{(\varkappa_\infty + 2)}\right]. \tag{5.50}$$

В твердых телах с ионной связью эта величина порядка 10^{-40} Ф·м^2. Так, для кристалла NaCl число ионных пар равно $2{,}2\cdot 10^{28}$ м$^{-3}$, а диэлектрическая проницаемость изменяется от 5,6 для области ниже частоты остаточных лучей до 2,25 для области выше этой частоты. Подставляя эти значения в выражение (5.50), получим, что для NaCl ионная поляризуемость $\alpha_i = 3{,}8\cdot 10^{-40}$ Ф·м^2.

Для пар ионов с массами M_+ и M_- уравнение движения вынужденных колебаний, индуцированных электрическим полем, записывается в виде

$$\left(\frac{M_+M_-}{M_+ + M_-}\right)\left[\frac{d^2x}{dt^2} + \gamma_i\frac{dx}{dt} + \omega_t^2 x\right] = eE_{\text{лок}}, \tag{5.51}$$

где γ_i, как и в уравнении (5.41), отражает диссипативный процесс, а ω_t — собственная резонансная частота (длинноволновая граничная частота поперечных оптических колебаний), определяемая силовыми постоянными твердого тела. Эффективное поле $E_{\text{лок}}$ представляет приложенное извне поле, а также отражает влияние электронной и ионной поляризации. Решение этого уравнения для x, как и в случае чисто электронной поляризации [уравнение (5.41)], для частот, сравнимых с ω_t, оказывается комплексным. Соотношение Клаузиуса — Мосотти при учете обоих типов поляризации дает следующее выражение для диэлектрической проницаемости:

$$\varkappa(\omega) = \varkappa_\infty + \frac{(\varkappa_0 - \varkappa_\infty)\,\omega_t^2}{(\omega_t^2 - \omega^2 + i\gamma_i\omega)}, \tag{5.52}$$

которое эквивалентно приведенному ранее выражению (2.43). Теперь можно выделить действительную и мнимую части точно так же, как это было сделано для случая электронной поляри-

зуемости [выражение (5.44)]. Таким образом, изменение действительной части диэлектрической проницаемости с увеличением частоты аналогично ее поведению в случае электронной поляризуемости, которое имеет место при более высоких частотах, и отражает наличие затухания. Отметим также, что для твердых тел, в которых возможно возникновение индуцированных ионных диполей, у мнимой компоненты $i\varkappa_2$ наблюдается довольно высокий максимум, расположенный при частоте ω_t. Именно он определяет хорошо известные оптические свойства таких твердых тел в данной спектральной области.

Пьезоэлектричество и электрострикция

В некоторых кристаллических твердых телах под действием упругой деформации возникает электрическая поляризация. Существует и обратное явление — в результате приложения внешнего электрического поля в определенных направлениях эти же твердые тела деформируются. Такой *пьезоэлектрический эффект* возможен только для ионных кристаллов, которые кристаллизуются в структурах, не обладающих центром инверсии.

Последнее утверждение можно проиллюстрировать на примере молекулы некоторого гипотетического твердого тела с ионной связью, в которой три электрических диполя равной величины в равновесном состоянии повернуты относительно друг друга на 120°. Это состояние молекулы представлено на рис. 5.9, *а*. Такая молекула имеет группу точечной симметрии $3m$ (см. рис. 1.19). Полный дипольный момент этой молекулы в силу симметрии равен нулю. Однако дипольный момент должен появиться, если эту молекулу (вместе со всем твердым телом) растянуть или сжать вдоль или поперек одного из трех направлений от центра к вершинам. Аналогично электрическое поле, приложенное вдоль одного из этих трех направлений, должно привести к деформации каждой молекулы. В свою очередь это вызовет сокращение или удлинение кристалла вдоль направления приложенного поля, сопровождаемое изменением противоположного знака его линейных размеров в поперечном направлении. Однако если электрическое поле приложить перпендикулярно к одному из этих трех направлений (например, в горизонтальном направлении на рис. 5.9), то изменений размера кристалла не должно произойти, поскольку в данном случае поле перпендикулярно плоскости зеркальной симметрии.

Из 32 точечных групп, которые рассматриваются в трехмерном пространстве, в 21 отсутствует центр инверсии. Одна из них обладает настолько высокой степенью симметрии, что исключает возникновение пьезоэлектрических эффектов. Для остальных 20 точечных групп пьезоэлектрические свойства ока-

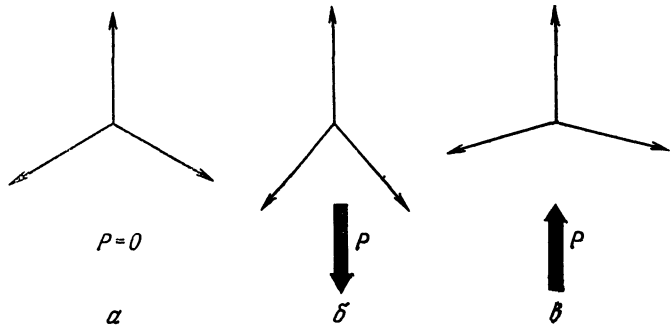


Рис. 5.9. Реакция пьезоэлектрической молекулы на деформацию, *а* — молекула в состоянии равновесия в неискаженном кристалле; *б* — поляризация, возникающая при растяжении в вертикальном или сжатии в горизонтальном направлениях; *в* — поляризация, возникающая при сжатии в вертикальном или растяжении в горизонтальном направлениях.

зываются возможными. Однако лишь относительно малая часть твердых тел, отвечающих этим 20 точечным группам, обладает пьезоэлектрическим эффектом, достаточно большим для регистрации. В табл. 5.1 приводятся наиболее известные пьезоэлектрические материалы в порядке уменьшения их пьезоэлектрических коэффициентов. Как видно из сопоставления с табл. 5.2, некоторые из этих твердых тел при температурах ниже температуры перехода T_c становятся сегнетоэлектриками.

Пьезоэлектрические коэффициенты кварца весьма малы, однако высокая механическая и термическая прочность этого материала делают его незаменимым для высокостабильных пьезоэлектрических генераторов. Эти приборы стабилизируют частоту радиопередатчиков, кварцевых часов и в последнее время наручных кварцевых часов.

Для других практических целей оказывается необходимой высокая пьезоэлектрическая эффективность. По этой причине сегнетова соль на протяжении многих лет использовалась в качестве материала для чувствительных преобразователей в недорогих проигрывателях. В более новых моделях применяются керамические пластины специальной формы из титаната бария — стронция или керамики из PZT*, так как эти материалы обладают большей пьезоэлектрической эффективностью и, кроме того, являются стойкими к воздействию нагрева и влаги. Блоки много больших размеров из титаната бария — стронция или PZT используются как источники ультразвука

* PZT означает соединение титаната циркония с добавками свинца (см. табл. 5.1). — *Прим. ред.*

Таблица 5.1. **Некоторые пьезоэлектрические материалы**

Материал	Химическая формула	Относительная величина пьезоэлектрических коэффициентов
Сегнетова соль	$NaKC_4H_4O_6 \cdot 4H_2O$	Очень большой
Титанат бария — стронция	$Ba_xSr_{1-x}TiO_3$	Большой
Титанат свинца — циркония (PZT)	$Pb_xZr_{1-x}TiO_3$	Большой
Дигидрофосфат аммония (ADP)	$NH_4H_2PO_4$	Большой
Дигидрофосфат калия (KDP)	KH_2PO_4	Средний
Диаминэтилентартрат*	$C_6H_{14}N_2O_6$	Средний
Турмалин	$(FeCrLiNaK)_4Mg_{12}B_6Al_{16}H_8Si_{12}O_{63}$	Маленький
α-кварц	SiO_2	Маленький

* Имеется в виду *D*-винная тартроновая кислота с двумя аминными группами и этиленом $C_4H_6O_6 \cdot 2(NH_2) \cdot C_2H_4$. — *Прим. ред.*

в ваннах для очистки и как передатчики и приемники в подводных сонарах.

В первом приближении деформация пьезоэлектрика в электрическом поле имеет линейный характер, а поляризация, индуцированная механическими напряжениями, пропорциональна возникающей деформации. Для *любого* ионного твердого тела вне зависимости от того, является ли оно пьезоэлектриком или нет, наблюдается *много меньшее* по величине сжатие в электрическом поле, которое пропорционально *квадрату* напряженности поля. Это более общее явление *электрострикции* связано с нарушением закона Гука при описании изменения межионных расстояний с приложением внешнего поля. Таким образом, можно полагать, что в случае твердого тела, у которого наблюдается значительная электрострикция, ангармонические эффекты оказывают заметное влияние на колебательные свойства кристаллической решетки.

Сегнетоэлектрические свойства

У твердого тела в сегнетоэлектрическом состоянии даже в отсутствие какого-либо приложенного поля центры распределения положительных и отрицательных зарядов не совпадают. Таким образом, наблюдается объемная электрическая поляризация вне зависимости от того, существует внешнее поле или нет.

Таблица 5.2. **Некоторые сегнетоэлектрические соединения***

Вещество	Химическая формула	T_c, К	P_s, Кл/м²
Триглицинсульфат	$(NH_2CH_2COOH)_3 \cdot H_2SO_4$	322	$2{,}8 \cdot 10^{-2}$ при 275 К
KDP	KH_2PO_4	123	$4{,}7 \cdot 10^{-2}$ при 100 К
Дейтерированный KDP	KD_2PO_4	213	$5{,}5 \cdot 10^{-2}$ при 100 К
Титанат бария	$BaTiO_3$	393	$2{,}6 \cdot 10^{-1}$ при 300 К
Титанат стронция	$SrTiO_3$	32	$3{,}0 \cdot 10^{-2}$ при 4,2 К
Ниобат калия	$KNbO_3$	710	$3{,}0 \cdot 10^{-1}$ при 600 К
Сульфат аммония	$(NH_4)_2 \cdot SO_4$	223	$4{,}5 \cdot 10^{-3}$ при 220 К
Сегнетова соль	$NaKC_4H_4O_6 \cdot 4H_2O$	296 (верхняя) 255 (нижняя)	$2{,}5 \cdot 10^{-3}$ при 275 К

* Данные взяты из *Jona F., Shirane. G.* Ferroelectric Crystals, Pergamon, 1962.

Одним из необходимых условий возникновения сегнетоэлектрических свойств является отсутствие центра инверсии в кристаллической структуре. Кроме того, необходимо, чтобы кристаллическая структура имела хотя бы одно неэквивалентное направление. Из 20 точечных групп, которые удовлетворяют условиям существования пьезоэлектрических свойств, 10 могут удовлетворить и второму условию. Таким образом сегнетоэлектрический материал должен быть пьезоэлектрическим, однако обратное необязательно.

В качестве общего требования для установления сегнетоэлектрического состояния необходимо отметить, что *упорядоченное* расположение *индуцированных* ионных диполей должно обладать меньшей полной энергией (даже в отсутствие электрического поля), чем для случая нулевого дипольного момента. В свою очередь это приводит к требованию большой плотности сильно поляризуемых ионов и не слишком высокой температуры. Сегнетоэлектрическое состояние разрушается выше некоторой характеристической температуры, *точки Кюри* T_c, поскольку выше этой температуры амплитуда тепловых колебаний становится достаточно большой для того, чтобы предотвратить образование упорядоченного расположения диполей. Твердое тело, которое при низких температурах является сегнетоэлектриком, при температурах выше точки Кюри становится параэлектриком с восприимчивостью, которая изменяется пропорционально $(T - T_c)^{-1}$.

В табл. 5.2 приведено несколько примеров хорошо известных сегнетоэлектрических кристаллов, их характеристические температуры и величины плотности спонтанной поляризации в отсут-

ствие поля. Как отмечается в таблице, спонтанная поляризация сегнетовой соли снова исчезает при температурах ниже второй характеристической температуры, но эта особенность не встречается у других сегнетоэлектриков. Последняя колонка табл. 5.2 содержит значения спонтанной объемной поляризации P_s в Кл/м^2 — единицах системы СИ. Можно оценить, каковы должны быть расстояния между зарядами, чтобы обеспечить значения поляризации, указанные в таблице, если учесть, что 10^{29} пар ионов на кубический метр с расстоянием в 0,25 Å между положительными и отрицательными зарядами обеспечивают поляризацию $P_s = 0{,}4$ Кл/м^2. Все значения поляризации, приведенные в табл. 5.2, меньше этой величины, некоторые даже в 100 раз.

Рассмотрим источник спонтанной поляризации $BaTiO_3$. Это соединение кристаллизуется в структуре перовскита (см. рис. 1.36). Решетка $BaTiO_3$ при температурах выше точки Кюри 393 К остается кубической, а переход в сегнетоэлектрическое состояние сопровождается появлением тетрагональных искажений куба. При температурах ниже T_c в элементарной ячейке происходят следующие изменения — она растягивается вдоль одного из псевдокубических направлений (ось c) примерно на 1 % и сжимается приблизительно на 0,5 % вдоль каждой из осей, перпендикулярных этому направлению. Само по себе тетрагональное искажение еще не приводит к поляризации; в действительности (при низких температурах во всяком случае) происходит следующее: подрешетка всех катионов бария и титана смещается вдоль оси c вверх или вниз по отношению к подрешетке анионов кислорода, что приводит к понижению энергии кристалла. Как установлено в результате структурных исследований[5], относительное смещение двух подрешеток приблизительно равно 0,1 Å, что оказывается вполне достаточным для возникновения больших значений объемной поляризации.

Для данного направления вдоль оси c в тетрагональном кристалле $BaTiO_3$ поляризация P_s может быть направлена либо «вверх», либо «вниз» в зависимости от направления, в котором одна подрешетка смещается относительно другой. Подчеркнем, что каждый ион титана (или бария) имеет два положения в кристаллической решетке, в которых его энергия минимальна, причем они разделены энергетическим барьером. Центр этого барь-

[5] Различные эксперименты в области структурной кристаллографии, на основании которых была установлена величина этого смещения, описаны в гл. 4 монографии: *Jona F., Shirane G.* Ferroelectric Crystals, Pergamon, 1962. [Имеется перевод: *Иона Ф., Ширане Д.* Сегнетоэлектрические кристаллы.— М.: Мир, 1965.] В действительности смещения подрешеток являются гораздо более сложными, так как ионы титана, бария и кислорода сдвигаются на различные расстояния при различных направлениях поляризации

ера соответствует тому положению иона титана (или бария), в котором он оказывается в результате тетрагональных искажений кристаллической решетки (рис. 1.36). Только при температурах выше точки Кюри этот барьер исчезает и наиболее низкое значение энергии решетки соответствует одному положению для каждого иона. Зависимость энергии иона от его положения в решетке в виде кривой с двумя минимумами является следствием существования локального поля $E_{\text{лок}}$, действие которого испытывает на себе каждый ион.

Объяснения природы сегнетоэлектричества в значительной степени неодинаковы для различных групп сегнетоэлектрических веществ, однако все они используют представление о том, что зависимость энергии кристалла от положения ионов имеет вид кривой с двумя минимумами. В KDP и ряде других подобных ему соединений определенные атомы водорода или дейтерия участвуют в образовании как водородных, так и обычных связей. Сегнетоэлектричество в этих кристаллах обусловлено смещением связанных водородными связями протонов (или дейтронов) из положений, при которых поляризация отсутствует, в тот или иной набор положений, для которых характерна более низкая энергия.

Чтобы показать, каким образом энергия может стать меньше при наличии индуцированных диполей, чем в их отсутствие, преобразуем соотношение Клаузиуса — Мосотти

$$[N_i\alpha_i + N_e\alpha_e] = 3\varepsilon_0\left(\frac{\varkappa - 1}{\varkappa + 2}\right) \tag{5.53}$$

к виду

$$(\varkappa - 1) = \frac{3[N_i\alpha_i + N_e\alpha_e]}{3\varepsilon_0 - [N_i\alpha_i + N_e\alpha_e]}. \tag{5.54}$$

Тогда диэлектрическая проницаемость будет стремиться неограниченно возрастать (приводя к возможности возникновения спонтанной поляризации) при условии, что

$$[N_i\alpha_i + N_e\alpha_e] \geqslant 3\varepsilon_0. \tag{5.55}$$

Соотношение (5.55) содержит как знак «больше», так и знак равенства, поскольку избыточная поляризация не исключает возникновения сегнетоэлектричества. Если поляризуемость такова, что знаменатель выражения в правой части (5.54) стремится к нулю, ангармонические эффекты будут устанавливать верхний предел для смещений ионов.

Когда реализуются условия для возникновения спонтанной поляризации, плотность энергии поляризующего поля превышает энергию, необходимую для образований диполей вне за-

висимости от того, приложено внешнее поле или нет. Допустим, что твердое тело содержит N_i поляризуемых пар ионов, причем каждая из них реагирует на локальное поле, которое, согласно (5.12), определяется соотношением

$$E_{\text{лок}} = (E_1 + P/3\varepsilon_0). \tag{5.56}$$

Тогда для зарядов $+q$ и $-q$, находящихся на расстоянии x друг от друга, дипольный момент равен

$$p = qx = \alpha_i E_{\text{лок}} = \frac{\alpha_i F}{q}, \tag{5.57}$$

где α_i — поляризуемость, а F — возвращающая сила, которая стремится свести вместе два заряда диполя. Работа, необходимая для *создания* N_i таких диполей на единицу объема, составляет

$$W_1 = N_i \int F dx = \frac{N_i q^2}{\alpha_i} \int x dx = \frac{N_i q^2 x^2}{2\alpha_i} = \frac{N_i p^2}{2\alpha_i} = \frac{P^2}{2N_i \alpha_i}. \tag{5.58}$$

Эту величину необходимо сравнить с энергией электрического смещения

$$W_2 = \int E_{\text{лок}} dP = \int [E_1 + P/3\varepsilon_0]\, dP = (P^2/6\varepsilon_0) + \int E_1 dP. \tag{5.59}$$

Таким образом, полная энергия поляризованного диэлектрика равна

$$(W_2 - W_1) = (P^2/2N_i\alpha_i)\,[(N_i\alpha_i/3\varepsilon_0) - 1] + \int E_1 dP. \tag{5.60}$$

Члены в правой части уравнения (5.60) записаны таким образом, чтобы подчеркнуть, что полная энергия сегнетоэлектрика, для которого выполняется неравенство (5.55), положительна даже при отсутствии внешнего поля. Таким образом, спонтанное упорядоченное расположение сильно поляризуемых диполей оказывается энергетически выгодным. При достаточно высоких температурах тепловые колебания решетки препятствуют этому упорядочению, и величина спонтанной поляризации стремится к нулю при приближении температуры к T_c со стороны низких температур.

При температурах выше T_c спонтанная поляризация сегнетоэлектрика исчезает, однако такое твердое тело обладает очень большой диэлектрической постоянной. Так, для керамики из $(Ba, Sr)TiO_3$, которая используется в малогабаритных конденсаторах большой емкости, величина $\varkappa$ может достигать значений вплоть до 6000. В других практических целях используется статическая поляризация сегнетоэлектриков при температурах ниже T_c. Например, конденсаторный микрофон с поляризованным сегнетоэлектриком, который заполняет бóльшую часть пространства между пластинами, снова пользуется боль-

шой популярностью. Электрооптические характеристики таких сегнетоэлектриков, как $BaTiO_3$ и KDP, используются для модуляции и отклонения лазерного луча как внутри, так и вне оптической полости. Одно время выражалась надежда, что монокристаллы $BaTiO_3$ могут быть использованы для создания запоминающих устройств с произвольной выборкой для компьютеров. «Биты» в тонких монокристаллических пластинках $BaTiO_3$ должны быть представлены поляризацией P, направленной вверх или вниз в зависимости от конструкции ячейки памяти. Знак этой поляризации должен считываться без ее разрушения по мере необходимости и изменение направления поляризации должно осуществляться приложением достаточно сильного электрического поля в противоположном направлении. Реализация этой идеи оказалась непростой задачей из-за наличия доменной структуры в $BaTiO_3$, которая будет обсуждаться ниже. Тем временем были разработаны альтернативные системы памяти для компьютеров, основанные на постоянном ферримагнетизме ферритов и на полупроводниковых триггерных ячейках.

Сегнетоэлектрические домены

Большой монокристалл сегнетоэлектрика не обязательно должен проявлять макроскопические признаки спонтанной поляризации, поскольку он, видимо, представляет собой совокупность доменов с различными направлениями поляризации. Наиболее типичным оказывается такое расположение, когда домен окружен соседними доменами с противоположным направлением поляризации P_s. В этом случае говорят о наличии 180-градусных *доменных стенок*. При приложении внешнего электрического поля доменные стенки получают возможность перемещаться. Это приводит к росту доменов, в которых направление P_s совпадает или почти совпадает с направлением E_1, в то время как домены с поляризацией P_s, направленной в противоположную сторону, имеют тенденцию сокращаться. Доменная структура сегнетоэлектриков до некоторой степени подобна доменной структуре ферромагнитных материалов, однако между ними имеется также и существенное различие. Блоховские стенки между магнитными доменами имеют толщину 750 Å (и обладают относительно низкой энергией), в то время как толщина стенок между сегнетоэлектрическими доменами равна одному или двум межатомным расстояниям и для них характерна высокая плотность энергии.

Рис. 5.10 иллюстрирует более сложный вариант доменной структуры в сегнетоэлектрическом кристалле опять же на примере тетрагонального $BaTiO_3$. В изображенном здесь кристалле имеются 180-градусные доменные стенки между областями

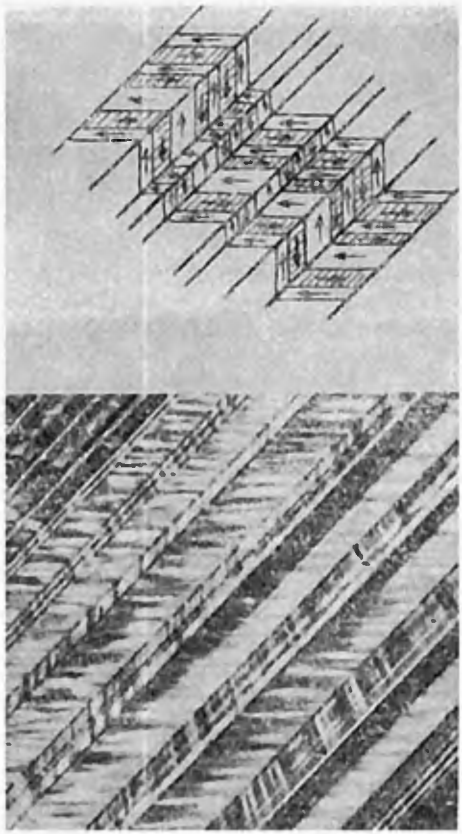


Рис. 5.10. Сегнетоэлектрические домены со 180- и 90-градусными доменными стенками в кристалле тетрагонального титаната бария [*Hooton J. A., Merz W. J.*— Phys. Rev., 98, 409, (1955)].

с антипараллельными поляризациями, однако наблюдаются и 90-градусные доменные стенки между участками кристалла с взаимно перпендикулярными или почти взаимно перпендикулярными поляризациями. Происхождение 90-градусных стенок в этом материале можно объяснить, если вспомнить, что при охлаждении ниже точки Кюри происходит небольшое тетрагональное растяжение одной из осей кубической решетки. Весьма вероятно, что в разных частях большого кристалла такое растяжение реализуется вдоль разных осей. Это должно привести к появлению сильно напряженных границ, вытянутых вдоль псевдоплоскостей (110), и возникновению доменов с поляриза-

циями, направленными под углом 45° к этим границам. Как установили Иона и Ширанэ (см. примечание 5), при тетрагональных искажениях решетки угол между направлениями поляризации доменов, разделенных 90-градусными стенками, в действительности равен 89°24′.

Для многих других сегнетоэлектрических материалов микроскопическая доменная структура оказывается не менее сложной.

Явления антисегнетоэлектричества

Таким образом, в сегнетоэлектрическом материале при температурах ниже температуры перехода возникает упорядоченное расположение индуцированных диполей, что приводит к понижению энергии кристалла. Для *антисегнетоэлектрических* твердых тел также характерно упорядочение индуцированных диполей ниже температуры перехода, но этот класс веществ не обладает объемной спонтанной поляризацией, поскольку каждый диполь *антипараллелен* соседним диполям. Требования к структуре для этого типа веществ те же, что и для истинных сегнетоэлектриков, и хорошо известные антисегнетоэлектрические материалы изоморфны некоторым сегнетоэлектрикам, упоминавшимся выше. Так, ниобат натрия $NaNbO_3$ и цирконат свинца $PbZrO_3$ кристаллизуются в структуре перовскита, однако антипараллельное выстраивание диполей соседних цепочек обеспечивает ниже определенной температуры более низкую полную энергию, чем для параллельной ориентации диполей в $BaTiO_3$ и $KNbO_3$ или для состояния с полным *отсутствием* индуцированных диполей. Точно так же дигидрофосфат калия (KDP) является сегнетоэлектриком, а изоморфный ему дигидрофосфат аммония (ADP) — антисегнетоэлектриком.

5.2. Магнитные свойства твердых тел

Некоторые из терминов, необходимых для обсуждения магнитных свойств, уже были определены при рассмотрении уравнений (5.2) и (5.3). Магнитные свойства большинства твердых тел в отличие от диэлектрических свойств определяются в основном ориентацией постоянных диполей, и мы будем иметь дело с *магнитной восприимчивостью* χ_m и *намагниченностью* $M=\chi_m H$ ансамбля постоянных диполей в приложенном поле. «Магнитная поляризуемость» обычно не рассматривается, поскольку для большинства твердых тел вклад индуцированных диполей совершенно незначителен.

Моменты магнитных диполей

Согласно закону Ампера, магнитный момент витка с током равен произведению площади витка на ток. Поэтому магнитный дипольный момент электрона, движущегося по круговой орбите радиуса r с угловой частотой ω равен

$$\mu_{\text{орб}} = -\frac{1}{2} er^2\omega. \tag{5.61}$$

При движении электрона в атоме соотношение между ω и r ограничивается условием квантования: орбитальный угловой момент должен быть кратным $\hbar$. Следовательно, необходимо, чтобы магнитный момент, связанный с орбитальным движением, был кратным *магнетону Бора*:

$$\mu_B = \frac{\hbar e}{2m} = 9{,}274 \cdot 10^{-24} \text{ Дж/Тл}. \tag{5.62}$$

Угловой момент, связанный со спином электрона, может быть охарактеризован спиновым квантовым числом $s = \pm 1/2$. Магнитный момент, связанный со спином, обычно записывается в виде

$$\mu_{\text{сп}} = gs\mu_B, \tag{5.63}$$

где g — величина, называемая *фактором спектроскопического расщепления* или просто *g-фактором*. Поскольку для свободного электрона $g = 2{,}0023$, можно считать, что магнитный момент, связанный со спином электрона, почти точно равен одному магнетону Бора.

Полный угловой момент многоэлектронного атома может быть получен векторным сложением орбитальных и спиновых угловых моментов. Эти вопросы подробно обсуждаются в учебниках по атомной физике. Допустим, что сложение векторов происходит согласно схеме, предложенной в 1925 г. Расселом и Саундерсом[6]*. В их схеме орбитальные квантовые моменты электронов векторно складываются, что дает полный орбитальный угловой момент атома $\hbar\mathbf{L}$; аналогично складываются и их спины, образуя полный спиновый угловой момент $\hbar\mathbf{S}$. В свою

[6] Рассел-саундерсовская связь спинового и орбитального угловых моментов и фактор расщепления Ланде обсуждаются во всех книгах по атомной физике, включая книги Бокхоффа, Френча и Тейлора, указанные в основном списке литературы, приведенном в начале этой книги.

* Из литературы на русском языке, в которой читатель может ознакомиться с этим вопросом, отметим книгу: *Ландау Л. Д., Лифшиц Е. М.* Квантовая механика. — М.: Физматгиз, 1963. — *Прим. перев.*

очередь $\hbar\mathbf{L}$ и $\hbar\mathbf{S}$ также складываются, образуя полный угловой момент системы электронов атома, который равен

$$\hbar\mathbf{J} = \hbar\mathbf{L} + \hbar\mathbf{S}. \tag{5.64}$$

Распределение электронов по различным квантовым состояниям в неполностью заполненной оболочке определяется принципом Паули и *правилами Хунда*[7]. Магнитный атом имеет частично заполненную оболочку и, следовательно, имеет некоторое число непарных электронов. Такой атом обладает отличным от нуля постоянным магнитным моментом, что является необходимым условием для возникновения парамагнетизма, ферромагнетизма, антиферромагнетизма и ферримагнетизма.

Если мы будем считать, что g-фактор для электрона равен 2,0000, а не 2,0023, то полный *магнитный момент* атома или иона есть

$$\boldsymbol{\mu} = \mu_B'(\mathbf{L} + 2\mathbf{S}). \tag{5.65}$$

Этот момент прецессирует вокруг направления $\mathbf{J}$, и, согласно предложенной в 1923 г. Ланде форме записи, этот момент можно представить в виде

$$\mu = g\mu_B J. \tag{5.66}$$

Здесь

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \tag{5.67}$$

— *фактор расщепления Ланде* или *g-фактор* для системы электронов, в которой орбитальная и спиновая составляющая связаны по схеме Рассела — Саундерса.

Очень малый магнитный момент связан со спиновым угловым моментом атомного ядра. Он измеряется в единицах *ядерного магнетона*

$$\mu_n = \frac{\hbar e}{2M_p} = 5{,}051 \cdot 10^{-27}\ \text{Дж/Тл}, \tag{5.68}$$

который меньше магнетона Бора во столько раз, во сколько масса протона больше массы электрона. Магнитный момент протона равен 2,793 μ_n, и он меньше, чем магнитные моменты большинства более тяжелых ядер. При рассмотрении статисти-

[7] Правила Хунда устанавливают, что электроны в частично заполненной оболочке располагаются (согласно принципу Паули) так, чтобы оболочка обладала максимальным спиновым моментом S и максимальным орбитальным моментом L (при данном S). Тогда $J = (L—S)$ для оболочки, которая заполнена менее чем наполовину (все спины параллельны) и $J = (L+S)$ для оболочки, которая заполнена более чем наполовину. Изложение правил Хунда можно найти в книгах, указанных в примечании 6.

ческой намагниченности ядерными моментами почти всегда можно пренебречь[8] по сравнению с вкладом электронной компоненты. Однако в разд. 5.3 мы увидим, что ядерные магнитные моменты обеспечивают возможность наблюдения спектров ЯМР и с их существованием связана сверхтонкая структура спектров ЭПР.

Диамагнетизм

Наиболее интересными типами магнетизма являются его основные типы, возникающие в веществах, у которых по меньшей мере часть атомов является магнитными. Приступая к рассмотрению этих вопросов, сначала необходимо хотя бы кратко остановиться на диамагнетизме, который присущ каждому электрону, движущемуся по орбите. У *каждого* электрона в *каждом* атоме при приложении внешнего магнитного поля проявляется индуцированная магнитная восприимчивость, однако эти индуцированные моменты оказываются очень слабыми. Диамагнетизм всегда присутствует, но его можно выявить лишь в тех случаях, когда все электроны вещества спарены, что исключает существование постоянных диполей.

Для изучения диамагнетизма можно воспользоваться моделью атома (или ионного остова атома), в которой все электроны оказываются спаренными (имея в виду, что аналогичная слабая реакция на магнитное поле должна также иметь место и для магнитных атомов). В отсутствие магнитного поля движение электронов сферически симметрично; угловой момент равен нулю, как и циркулирующий вокруг ядра электрический ток и магнитный момент. Если теперь приложить магнитное поле $\mathbf{H}$, то соответствующая ему величина $\mathbf{B}$ будет определять действующую на каждый электрон силу Лоренца $e\mathbf{v}\times\mathbf{B}$. *Правило Ленца* описывает это взаимодействие: при изменении магнитного потока в контуре индуцируется электрический ток такого направления, что он препятствует изменению потока. Магнитная сила нарушает равновесие между центробежной и центростремительной силами таким образом, что частота орбитального движения электрона, обладающего орбитальным магнитным моментом, параллельным полю, уменьшается, а частота орбитального движения электрона с орбитальным магнитным моментом, антипараллельным полю, возрастает.

---

[8] Исключением является ядерный парамагнетизм твердого водорода, который был измерен В. Лазаревым и Л. Шубниковым [*Lazarew W., Schubnikow L.*— Phys. Z. Sowjetunion, **11**, 445 (1937)]. Очень слабый парамагнетизм протонов этого вещества удалось измерить благодаря диамагнетизму системы электронов.

Изменение частоты орбитального движения можно определить из простого классического анализа (задача 5.7); это изменение есть *ларморова частота прецессии*:

$$\Delta\omega = \frac{eB}{2m}. \tag{5.69}$$

Все Z электронов атома прецессируют в одном и том же направлении, что эквивалентно появлению циркулирующего тока

$$I = -\left(\frac{Ze\Delta\omega}{2\pi}\right) = -\left(\frac{Ze^2B}{4\pi m}\right). \tag{5.70}$$

Мы уже приводили формулировку закона Ампера, согласно которому магнитный момент витка с током равен произведению величины циркулирующего тока на площадь витка. Таким образом, индуцированный ток, определенный соотношением (5.70), приводит к появлению магнитного диполя.

Если средний квадрат расстояния электронов от ядра атома равен $\langle r^2\rangle$, то средний квадрат смещения в плоскости, перпендикулярной **B**, есть $(2\langle r^2\rangle/3)$, а эффективная площадь витка с током равна $(2\pi\langle r^2\rangle/3)$. Следовательно, атомный дипольный момент равен

$$\mu_{\text{диа}} = (2\pi I\langle r^2\rangle/3) = -\left(\frac{Ze^2\langle r^2\rangle B}{6m}\right). \tag{5.71}$$

Из выражения (5.71) видно, что индуцированная намагниченность отрицательна и, следовательно, объемная диамагнитная восприимчивость $\chi_{\text{диа}} = (+M/H) \simeq (+\mu_0 M/B)$ также отрицательна. Величина χ_m безразмерна и для диамагнитных веществ много меньше единицы. Чтобы получить выражение для $\chi_{\text{диа}}$, заменим $(\mu_0 M/B)$ на (M/H). Из формулы (5.71) можно получить, что для вещества, в каждой единице объема которого содержится N атомов одного сорта,

$$\chi_{\text{диа}} = -(\mu_0 NZe^2\langle r^2\rangle/6m). \tag{5.72}$$

Тот же результат получается и при решении квантовомеханической задачи, как показано, например, в книге Ван-Флека, которая приведена в библиографии к этой главе.

Используя типичные для твердого тела величины N и $\langle r^2\rangle$, с помощью выражения (5.72), получаем, что диамагнитная восприимчивость $\chi_{\text{диа}} \approx -10^{-5}$. Эта величина сравнима с экспериментально определенными значениями диамагнитной восприимчивости, как показывают последние пять строк табл. 5.3. Диа-

Таблица 5.3. **Некоторые парамагнетики и диамагнетики**

Материал	Парамагнитная восприимчивость χ_m (при $T=300$ К)	Диамагнитная восприимчивость χ_m
Платина	$+2{,}9\cdot10^{-4}$	
Хром	$+2{,}7\cdot10^{-4}$	
Ниобий	$+2{,}6\cdot10^{-4}$	
Алюминий	$+2{,}2\cdot10^{-5}$	
Литий	$+2{,}1\cdot10^{-5}$	
Кальций	$+1{,}9\cdot10^{-5}$	
Натрий	$+9{,}1\cdot10^{-6}$	
Кислород (при нормальных условиях)	$+1{,}9\cdot10^{-6}$	
Азот (при нормальных условиях)		$-5{,}0\cdot10^{-9}$
Кремний		$-4{,}2\cdot10^{-6}$
Медь		$-9{,}5\cdot10^{-6}$
Галлий		$-2{,}3\cdot10^{-5}$
Ртуть		$-2{,}9\cdot10^{-5}$
Висмут		$-1{,}6\cdot10^{-4}$

магнетизм присущ инертным газам (как в твердом, так и в газообразном состоянии), а также положительным и отрицательным ионам твердых тел с заполненными электронными оболочками.

Более сложные модели диамагнетизма, включая диамагнетизм молекул, в которых атомы связаны ковалентными связями, описаны в монографии Ван-Флека.

Различные формы проявления постоянных дипольных моментов

Рис. 5.11 иллюстрирует четыре различных вида магнетизма к которым приводит существование постоянных диполей в твердом теле. В дальнейшем мы рассмотрим их более подробно, однако вначале полезно вспомнить качественную картину возможных типов упорядочения.

Парамагнетизм имеет место, когда магнитные моменты различных атомов в отсутствие магнитного поля не согласованы между собой. Конечно, в магнитном поле диполи стремятся к упорядочению и намагниченность определяется функцией, до некоторой степени подобной функции Ланжевена. В действительности ситуация оказывается несколько более сложной из-за

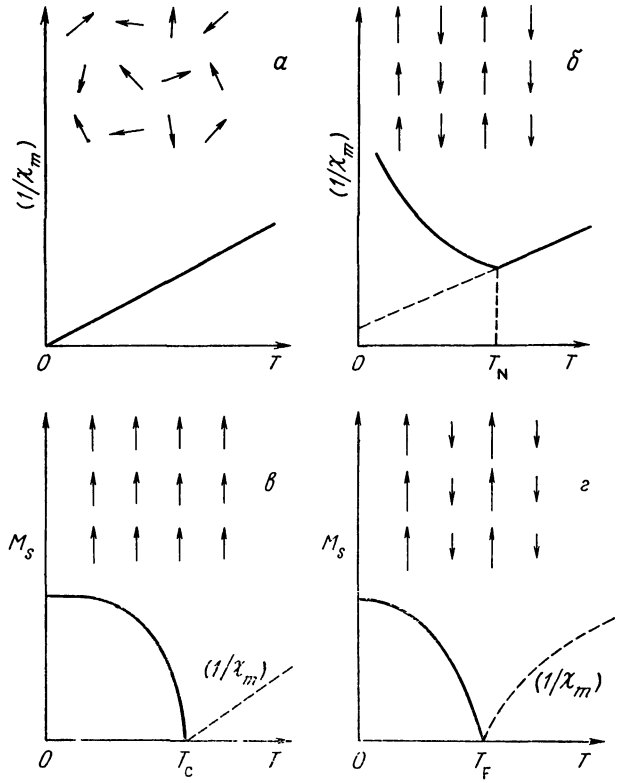


Рис. 5.11. Возможное низкотемпературное упорядочение соседних диполей и соответствующее поведение спонтанной намагниченности и/или восприимчивости для парамагнетика (*а*), антиферромагнетика (*б*), ферромагнетика (*в*), ферримагнетика (*г*) [*Chikazumi S.*— Physics of Magnetism, Wiley, 1964].

квантовых ограничений на величину проекции углового момента на направление внешнего поля[9], однако в слабом поле магнитная восприимчивость описывается *законом Кюри:*

$$\chi_m = \frac{C}{T}. \tag{5.73}$$

[9] Описание парамагнетизма с помощью функции Ланжевена в общем аналогично использованию этой функции для описания простой картины, отражающей реакцию постоянных электрических диполей в полярном твердом теле на внешнее электрическое поле. В случае *параэлектрического* твердого тела сложности были связаны с ограничениями на ориентации диполей, налагаемыми кристаллической решеткой. В *парамагнетиках* ситуация осложняется совсем по другой причине (квантование углового момента) и функция Ланжевена должна быть заменена на *функцию Бриллюэна.*

Для *антиферромагнетиков* полная энергия кристалла при низких температурах в отсутствие внешнего магнитного поля минимальна, когда имеет место чередование противоположно направленных магнитных моментов. Такое расположение магнитных моментов очень устойчиво при низких температурах и магнитная восприимчивость в приложенном поле мала (значение $(1/\chi_m)$, как показано на рис. 5.11, *б*, велико). При повышении температуры эффективность взаимодействий между диполями уменьшается и восприимчивость возрастает до тех пор, пока при температуре Нееля T_N спины не станут «свободными», чтобы реагировать на магнитное поле. При еще более высоких температурах твердое тело ведет себя как парамагнетик и его восприимчивость описывается модифицированным законом Кюри:

$$\chi_m = \frac{C}{(T + \theta)}. \tag{5.74}$$

Ферромагнетики при температурах ниже температуры Кюри T_c обладают упорядоченной структурой с параллельными спинами, что приводит к возникновению спонтанной намагниченности M_s. При температуре Кюри эта объемная намагниченность падает до нуля, и при более высоких температурах магнитная восприимчивость парамагнетиков для системы неупорядоченных спинов подчиняется *закону Кюри — Вейсса*:

$$\chi_m = \frac{C}{(T - T_c)}. \tag{5.75}$$

Упорядочение *ферримагнитных* материалов при низких температурах аналогично упорядочению антиферромагнетиков. Однако направленные навстречу друг другу магнитные моменты двух спиновых систем не равны, и состояние с наименьшей энергией обладает спонтанной намагниченностью. При нагревании такого твердого тела до точки Кюри T_f его намагниченность уменьшается до нуля, и при более высоких температурах оно ведет себя как парамагнетик. Температурная зависимость парамагнитной восприимчивости ферримагнитных твердых тел оказывается более сложной, чем в трех других случаях, приведенных на рис. 5.11.

Парамагнетизм

Рассмотрим твердое тело, содержащее N магнитных атомов в единице объема, каждый из которых обладает магнитным моментом μ [выражение (5.66)]. В отсутствие магнитного поля состоянием с наименьшей энергией для этого твердого тела

будет состояние с магнитными моментами, ориентированными случайным образом. Если теперь твердое тело поместить в магнитное поле **H**, энергия диполя оказывается равной $-\boldsymbol{\mu}\cdot\mathbf{B}$ и увеличение **H** и **B** должно приводить к тому, что ориентация дипольных моментов вдоль поля оказывается предпочтительной. По аналогии с ориентацией электрических диполей [см. выражения (5.20)—(5.28)] можно ожидать, что намагниченность записывается в виде

$$M = N\mu L(\mu B/k_0 T). \tag{5.76}$$

Здесь $L(w)$ — функция Ланжевена, приведенная ранее выражением (5.25). Известно, что $L(w)\simeq(w/3)$, когда $w\ll 1$. Таким образом, для некоторой комбинации приложенного поля и температуры, которая удовлетворяет приближению *слабого поля*, из равенства (5.76) можно получить выражение для парамагнитной восприимчивости

$$\chi_m = (M/H) \simeq (\mu_0 M/B) = C/T \text{ при } \mu B \ll k_0 T, \tag{5.77}$$

где

$$C = (\mu_0 \mu^2 N/3k_0). \tag{5.78}$$

Выражение (5.77) известно как *закон Кюри*, а величина C — *постоянная Кюри* для данного твердого тела. Прямая линия на рис. 5.11, *а*, описывает соответствующую температурную зависимость. Согласно выражению (5.77), при концентрации магнитных атомов $N\simeq 5\cdot 10^{28}$ м$^{-3}$, каждый из которых обладает магнитным моментом в один магнетон Бора, восприимчивость $\chi_m\simeq(0{,}13/T)$, что вполне типично для целого ряда практических случаев. В верхней части табл. 5.3 представлена парамагнитная восприимчивость для некоторых химических элементов.

Закону Кюри очень хорошо следуют многие парамагнетики, однако для других твердых тел при низких температурах наблюдаются отклонения от него. Эти отклонения следует считать скорее правилом, чем исключением, так как парамагнитная ориентация усложняется правилами квантовой механики, определяющими возможные значения проекции углового момента на выбранное направление.

Согласно квантовой механике, вектор полного углового момента атома равен $\hbar\sqrt{J(J+1)}$, а его проекция на направление магнитного поля может иметь одно из значений $\hbar m_J$, где азимутальное квантовое число m_J принимает одно из значений J, $(J-1)$, $(J-2)$, ..., $(1-J)$, $-J$. Тогда максимальное значение $\hbar m_J$ оказывается меньше величины вектора полного углового момента; при этом имеется некоторый поперечный компонент.

Для каждого значения m_J полному магнитному моменту μ [выражение (5.66)] соответствует проекция на направление поля, которая равна $g\mu_B m_J$ [здесь g — фактор спектроскопического расщепления Ланде, представленный соотношением (5.67)]. Такой проекции магнитного момента и магнитной индукции B соответствует потенциальная энергия

$$U_{J,m} = -g\mu_B m_J B. \tag{5.79}$$

Теперь для того, чтобы определить намагниченность для данной температуры и данного магнитного поля, надо набор энергий, представленный выражением (5.79), сопоставить с больцмановской энергией k_0T. Задача 5.8 как раз основана на очень простом примере такого определения намагниченности. В общем случае магнитный момент на единицу объема следует записать в виде

$$M = N\left[\frac{\sum_{-J}^{J}(g\mu_B m_J)\exp(g\mu_B m_J B/k_0T)}{\sum_{-J}^{J}\exp(g\mu_B m_J B/k_0T)}\right]. \tag{5.80}$$

Перепишем это выражение таким образом:

$$M = Ng\mu_B J B_J(y), \tag{5.81}$$

где безразмерная переменная $y \equiv (g\mu_B JB/k_0T)$, а $B_J(y)$ — *функция Бриллюэна*. В результате суммирования в выражении (5.80) функцию Бриллюэна можно записать в виде

$$B_J(y) = \left\{\left(\frac{2J+1}{2J}\right)\mathrm{cth}\left[\frac{(2J+1)}{2J}y - \left(\frac{1}{2J}\right)\mathrm{cth}\left(\frac{y}{2J}\right)\right]\right\}, \tag{5.82}$$

и для очень больших значений J она асимптотически переходит в функцию Ланжевена (5.25). Поскольку для реальных атомов величина J конечна, возможно лишь ограниченное число ориентаций по отношению к направлению поля, причем ни одна из них не совпадает с этим направлением.

Функция Бриллюэна изменяется от нуля при отсутствии внешнего поля, до единицы в бесконечно большом поле[10]. Таким образом, намагниченность насыщения парамагнетика $M_{\text{макс}} = Ng\mu_B J$. На рис. 5.12 представлены результаты измерений при очень низких температурах и в различных магнитных полях вплоть до полей большой напряженности. Этот рисунок

[10] Функции Бриллюэна для различных значений J сведены в таблицу в книге: *Smart J. S.* Effective Field Theories of Magnetism, Saunders, 1966. [Имеется перевод: *Смарт Дж.* Эффективное поле в теории магнетизма — М: Мир, 1968.]

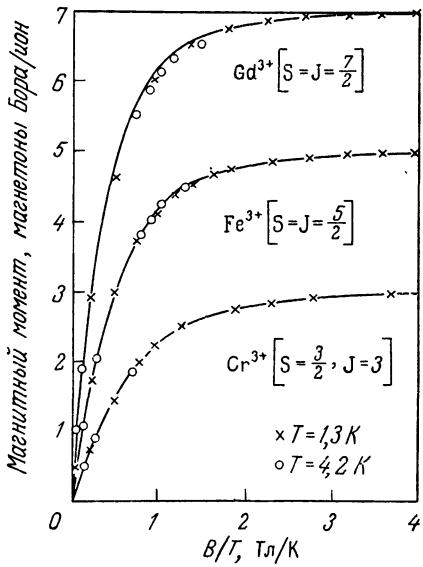


Рис. 5.12. Зависимость магнитных моментов от магнитного поля для парамагнитных ионов в непроводящих солях. По оси абсцисс отложено отношение B/T. Ион Gd^{3+} находится в октагидратсульфате гадолиния, Fe^{3+} — в ферроаммонийном алюминии, а Cr^{3+} — в калийхромовокислом алюминии. Кривые на рисунке представляют функции Бриллюэна для соответствующих J (или соответствующих значений S) [*Warren Henry.*— Phys. Rev., **88**, 559 (1952)].

показывает, что намагниченность парамагнитных солей изменяется в соответствии с функцией Бриллюэна.

При выполнении условия слабого поля функция Бриллюэна асимптотически стремится к пределу

$$B_J(y) \approx [y(J+1)/3J], \qquad y \ll 1, \tag{5.83}$$

а при условии, что $B \simeq \mu_0 H$ мало по сравнению с $(k_0 T/g\mu_B J)$, поведение магнитной восприимчивости описывается законом Кюри:

$$\chi_m = (M/H) \simeq (\mu_0 M/B) = [\mu_0 \mu_B^2 g^2 NJ(J+1)/3k_0 T]. \tag{5.84}$$

Парамагнитная восприимчивость обычно является достаточно малой величиной, поэтому замена $\mu_0 H$ на B оказывается хорошим приближением. Сопоставление выражения (5.84) с законом Кюри [выражениями (5.73) и (5.77)] позволяет установить, что константа Кюри имеет вид

$$C = [\mu_0 \mu_B^2 g^2 NJ(J+1)/3k_0]. \tag{5.85}$$

Сравнив выражения (5.78) и (5.85) для константы Кюри, можно сделать вывод, что каждый магнитный атом ведет себя так, будто он обладает *эффективным дипольным моментом*

$$\mu_{\text{эфф}} = g\mu_B \sqrt{J(J+1)}, \tag{5.86}$$

если рассматривать его поведение в *слабом магнитном поле.*

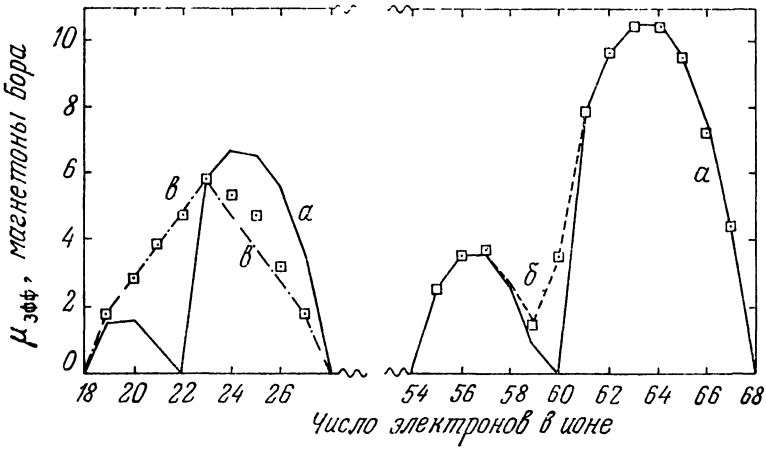


Рис. 5.13. Магнитный момент как функция числа электронов для парамагнитных ионов переходных элементов группы железа и редкоземельных элементов [*Bates L. F.* Modern Magnetism, Cambridge, 4th ed., 1961]. Точками представлены экспериментальные данные. Они сопоставляются со сплошной кривой для случая рассел-саундерсовской связи, когда магнитный момент в магнетонах Бора равен $g\sqrt{J(J+1)}$ (*а*); пунктирной кривой, соответствующей более сложной модели для редкоземельных ионов, которая развита Ван-Флеком (см. библиографию) (*б*); штрихпунктирной кривой для элементов группы железа в случае, если орбитальный момент полностью заморожен и магнитный момент в магнетонах Бора равен $2\sqrt{S(S+1)}$ (*в*).

Исследование парамагнитных свойств удобнее всего проводить на непроводящих соединениях, в состав которых входят переходные элементы (исследования на самих переходных металлах осложняются вкладом парамагнетизма Паули свободных электронов). Такие измерения в широком масштабе были проведены на соединениях переходных элементов группы железа (частично заполненная 3*d*-оболочка) и элементов редкоземельной группы (4*f*-оболочка); в меньшем объеме эти измерения проводились для соединений элементов группы палладия (4*d*-оболочка), группы платины (5*d*-оболочка) и группы урана (5*f*-оболочка).

Магнитные моменты атомов, полученные из этих экспериментов, достаточно хорошо согласуются с величинами моментов, рассчитанными на основе выражений (5.66) и (5.67) для элементов редкоземельной группы, но не согласуются для элементов группы железа, что наглядно продемонстрировано на рис. 5.13, где приведены расчетные кривые и экспериментальные данные. Для элементов с частично заполненной 3*d*-оболочкой *действующие* моменты гораздо ближе к тем величинам, при расчете которых учитывается наличие одного лишь спина,

а орбитальный момент оказывается «замороженным». Кривые для Gd^{3+} и Fe^{3+} (рис. 5.12) в этом отношении неинформативны, поскольку для каждого из этих ионов орбитальный момент L равен нулю, что приводит к $J=S$ и фактору Ланде $g=2$. Однако экспериментальные данные для Cr^{3+} на рис. 5.12 очень хорошо согласуются с расчетной кривой для значений $J\approx 3/2$ и $g\approx 2$ (орбитальный момент «заморожен»), а не с кривой для $J=3$ и $g=0{,}4$, которую можно ожидать в случае взаимодействия Рассела—Саундерса.

Нарушение рассел-саундерсовской связи для элементов группы железа можно объяснить как результат взаимодействия со значительным неоднородным локальным электрическим полем, создаваемым соседними атомами, так как $3d$-электроны этой группы элементов находятся на внешней оболочке и не защищены от внешнего влияния. $4f$-электроны, с которыми связан парамагнетизм редкоземельных элементов, экранированы от соседних атомов $5s$- и $5p$-электронами, что уменьшает *расщепление кристаллическим полем* и практически не нарушает связь Рассела—Саундерса.

Ферромагнетизм

Материалы с постоянной намагниченностью при комнатной температуре в отсутствие внешнего магнитного поля известны уже много лет. Еще древние мореплаватели использовали кусочки спонтанно намагниченного магнетита (Fe_3O_4) в своих примитивных компасах. Большинство ферромагнитных материалов, известных до нашего столетия,— металлы и их сплавы. Значительное число непроводящих ферромагнитных соединений было открыто в течение нескольких последних десятилетий [11].

Из сказанного ясно, что ферромагнетизм основан на коллективном выстраивании постоянных атомных диполей, которые возникают в атомах, имеющих непарные электроны. Момент каждого отдельного диполя μ равен нескольким магнетонам Бора, но полностью упорядоченное расположение таких моментов приводит к большой спонтанной намагниченности M_s, как можно видеть из табл. 5.4. Число магнетонов Бора (μ/μ_B) для большинства ферромагнитных материалов почти целиком определяется спином непарных электронов с малой поправкой, связанной с орбитальным движением.

В ряде случаев число магнетонов можно объяснить, принимая во внимание спектр электронных энергетических зон

[11] Таблица ферромагнитных материалов, содержащая и некоторые непроводящие ферромагнитные соединения, приведена в книге: *Смарт Дж.* Эффективное поле в теории магнетизма. — М.: Мир, 1968. Дополнительную информацию можно найти в книге: American Institute of Physics Handbook, McGraw-Hill, 3rd Ed., 1971.

Таблица 5.4. **Ферромагнитные элементы***

Материал	Кристаллическая структура	Точка Кюри T_c, К	Намагниченность насыщения при $T = 0$ M_s, А/м	$B_r = \mu_0 M_s$, Тл	Число магнетонов Бора на атом (μ/μ_B)
Fe	о. ц. к.	1043	$1{,}74 \cdot 10^6$	2,18	2,22
Co	г. п. у.	1395	$1{,}45 \cdot 10^6$	1,82	1,72
Ni	г. ц. к.	631	$5{,}11 \cdot 10^5$	0,643	0,606
Gd	г. п. у.	289	$2{,}11 \cdot 10^6$	2,65	7,1
Dy	г. п. у	85	$2{,}92 \cdot 10^6$	3,67	10,0

* Данные взяты из книги: *Смарт Дж.* Эффективное поле в теории магнетизма. — М.: Мир, 1968.

в твердом теле [12]. Рассмотрим, например, распределение электронов в $3d$- и $4s$-зонах никеля. Изолированный атом никеля имеет конфигурацию $3d^8 4s^2$, однако в металлическом никеле вплоть до энергии Ферми имеется 9,46 электронов на атом в $3d$-зоне и только 0,54 электронов на атом в перекрывающейся с ней $4s$-зоне. Когда никель находится в парамагнитном состоянии (при температурах выше его точки Кюри T_c), степень заполнения каждой из зон состояниями со *спином вверх* и со *спином вниз* оказывается одинаковой, что и проиллюстрировано на рис. 5.14, *а*. В этом случае произвольное внешнее магнитное поле приводит к очень слабому относительному сдвигу этих двух наборов состояний, что уже отмечалось при рассмотрении парамагнетизма Паули в разд. 3.3. При температуре, достаточно низкой для возникновения ферромагнитного упорядочения в никеле, очень сильное эффективное молекулярное поле вызывает значительный энергетический сдвиг состояний со спином верх по отношению к состояниям со спином вниз. На рис. 5.14, б показано, что $3d$-зона для одного направления спина оказывается полностью заполненной (5 электронов на атом), в то время как для другой ориентации спина имеется только 4,46 электрона на атом. Таким образом, ферромагнитное упорядочение приводит к появлению в $3d$-зоне 0,54 непарного электрона на атом. Поправки на g-фактор этих электронов, на некоторую подстройку заселенности $4s$-зоны и на небольшой магнитный вклад орбитального движения приводят к тому, что при насыщении намагниченности магнитный момент в никеле оказывается равным 0,606 μ_B на атом.

[12] *Stoner E. C.*— Repts. Progr. Phys., 11, 43 (1948).

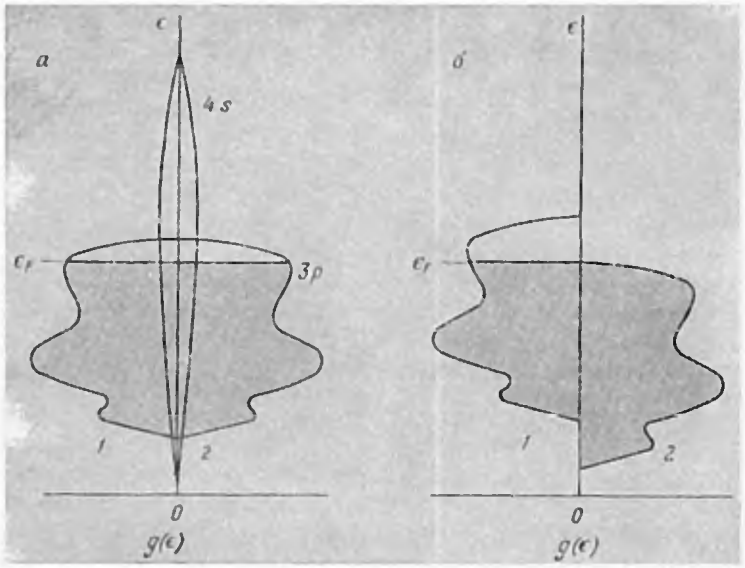


Рис 5.14 *а* — зависимость плотности состояний от энергии для двух направлений спина в 3*d*- и 4*s*-зонах никеля в парамагнитном состоянии; вплоть до энергии Ферми ε_F 3*d*-зона содержит 9,46 электрона на атом, а 4*s*-зона — 0,54 электрона на атом. *б* — в ферромагнитном состоянии 3*d*-зона расщепляется молекулярным полем; заполнение состояний со спином вниз (*2*) превышает заполнение состояний с противоположным направлением спина (*1*).

В самом простом случае предполагается, что при ферромагнитном упорядочении взаимодействие магнитного атома с кристаллом может быть описано некоторым эффективным полем, или *молекулярным полем*. Это приближение было предложено Пьером Вейссом [13]. Он установил, что величина эффективного молекулярного поля H_e пропорциональна магнитному моменту на единицу объема, т. е.

$$H_e = \gamma M. \tag{5.87}$$

Коэффициент пропорциональности γ обычно называют коэффициентом Вейсса. Величина γ является характеристикой данного сорта магнитных атомов; она зависит от кристаллической структуры и не зависит от температуры.

Если твердое тело находится в ферромагнитном состоянии, M и H_e велики в отсутствие всякого внешнего поля. При температурах выше точки Кюри T_c для создания намагниченности необходимо приложить внешнее поле H, но теперь величина

[13] *Weiss P.*— J. Phys. Rad, **4**, 661 (1907).

намагниченности определяется и молекулярным полем, которое она сама и создает. Полную намагниченность в слабом приложенном поле можно записать в виде

$$M = \chi_0 (H + H_e), \qquad T > T_c, \tag{5.88}$$

где $\chi_0 = (C/T)$ — магнитная восприимчивость в слабом поле, которая подчиняется закону Кюри [выражение (5.73)] для парамагнитного материала. Однако χ_0 будет меньше *наблюдаемой* восприимчивости $\chi_m = (M/H)$ — намагниченности, отнесенной к единичному *приложенному* полю. Подставив выражение (5.87) в (5.88), полную намагниченность можно выразить как

$$M = \chi_m H = \chi_0 H (1 - \gamma\chi_0)^{-1}, \qquad T > T_c. \tag{5.89}$$

Тогда наблюдаемая магнитная восприимчивость

$$\chi_m = \frac{\chi_0}{1 - \gamma\chi_0} = \frac{C}{T - \gamma C} = \frac{C}{T - T_c}, \qquad T > T_c = \gamma C \tag{5.90}$$

и теория молекулярного поля приводит прямо к закону *Кюри—Вейсса* для магнитной восприимчивости выше точки Кюри. [Ранее этот закон был представлен выражением (5.75) без всякого объяснения.] Теперь можно получить коэффициент эффективного поля Вейсса для ферромагнитного материала, если известны постоянная Кюри и температура Кюри для его магнитных свойств при высоких температурах. Типичные значения $\gamma = (T_c/C) \sim 10^4$ свидетельствуют, что молекулярное поле имеет величину $H_e \sim 10^{10}$ А/м [$\mu_0 H_e \sim 10^4$ Тл] при низких температурах для таких ферромагнитных материалов, как элементарные ферромагнетики, представленные в табл. 5.4.

Приближение молекулярного поля Вейсса также позволяет получить выражение для температурной зависимости спонтанной намагниченности M_s в ферромагнитном состоянии. Пусть имеется N магнитных атомов в единице объема, причем каждый из них обладает магнитным моментом μ_1. Если ферромагнетизм рассматривать как парамагнетизм со значительным молекулярным полем, которое обеспечивает самоупорядочение, то намагниченность для магнитной индукции B может быть записана через функцию Бриллюэна [выражение (5.81)] как

$$M = N\mu_1 B_J (\mu_1 B / k_0 T). \tag{5.91}$$

При температуре ниже T_c в отсутствие внешнего поля $B = \mu_0 H_e = \gamma\mu_0 M_s$ и спонтанная намагниченность M_s должна удовлетворять условию

$$M_s = N\mu_1 B_J (\gamma\mu_0\mu_1 M_s / k_0 T). \tag{5.92}$$

Для любой температуры от абсолютного нуля до T_c и функции Бриллюэна, определяемой величиной J, имеется только одно значение M_s, которое удовлетворяет уравнению (5.92). Поскольку

любая функция Бриллюэна при абсолютном нуле выходит на насыщение, то $M_{s0}=N\mu_1$ и

$$M_s/M_{s0}=B_J(\gamma\mu_0\mu_1 M_s/k_0T). \tag{5.93}$$

Решение для (M_s/M_{s0}) как функции (T/T_c) может быть получено графически, что и предлагается сделать в задаче 5.9. Уменьшение (M_s/M_{s0}) с ростом температуры показано на рис. 5.15 на примере никеля. На этом же рисунке сплошной кривой представлено уравнение (5.93) для $J=S=1/2$. Подгонка позволяет достаточно хорошо совместить кривую с экспериментальными точками, и поэтому дальнейшее обсуждение настоящей модели представляется интересным.

С самого начала Вейссу было ясно, что его модель «молекулярного поля» нельзя построить на основе *магнитостатического* взаимодействия постоянных магнитных диполей [14], так как предсказываемая в этом случае величина поля оказывается совершенно нереальной. Исходя из значения M_s, он сделал вывод, что каждый диполь в ферромагнетике обладает магнитным моментом μ_1, величина которого сравнима с тем, что теперь мы называем магнетоном Бора. Таким образом, он мог рассчитать поле $\mu_1/r_{ab}{}^3$, создаваемое одним магнитным диполем, в точке r_{ab}, где находится другой диполь. Очевидно, на каждый из упорядоченно выстроенных диполей должно воздействовать эффективное поле

$$H_e=\gamma_1\mu_1 r_{\text{ав}}^{-3}=\gamma_2\mu_1 N, \tag{5.94}$$

где N — число магнитных атомов в единице объема, а γ_1 и γ_2 — безразмерные множители. С помощью аргументов, подобных тем, что были использованы при получении эффективного поля в поляризованном диэлектрике, можно показать, что величины γ_1 и γ_2 несколько превышают единицу. Таким образом, Вейсс смог установить, что эффективное поле H_e не должно превышать 10^6 А/м. При этом энергия связи для каждого диполя оказывается всего лишь $\mu_0\mu_1 H_e\approx 10^{-23}$ Дж (10^{-4} эВ), и столь слабо связанная совокупность диполей должна разрушиться при температуре Кюри $T_c\approx(\mu_0\mu_1 H_e/k_0)\approx 1\,K$. Это зна-

[14] Пожалуй настало время подчеркнуть основные различия между процессами упорядочения в *сегнетоэлектрических* и *ферромагнитных* твердых телах. Действительно, представляется очень странным, что причины возникновения двух таких своеобразных явлений совершенно различны. В сегнетоэлектриках диполи являются *индуцированными* диполями и энергия, связывающая их в упорядоченном состоянии, определяется именно диполь-дипольным взаимодействием. Напротив, магнитные диполи ферромагнетика являются *постоянными* диполями, и напряженность поля для прямого магнитостатического диполь-дипольного взаимодействия на четыре порядка меньше, чем для обменного механизма, благодаря которому и возникает ферромагнитное состояние.

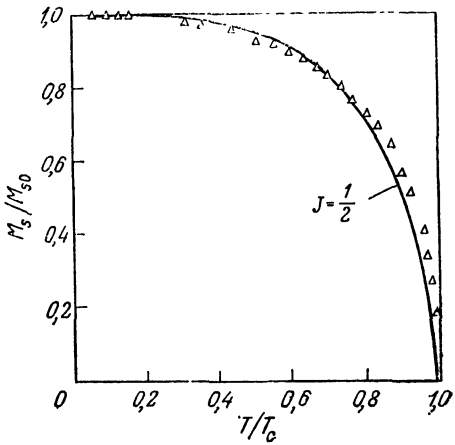


Рис. 5.15. Температурная зависимость спонтанной намагниченности M_s (как доли спонтанной намагниченности при насыщении при абсолютном нуле) для ферромагнетика. Кривая соответствует уравнению (5.93) при $J=1/2$. Точки представляют экспериментальные данные для металлического никеля, у которого $T_c=631$ К, $M_{s0}=5{,}11\cdot10^5$ А/м ($\mu_0 M_{s0}=0{,}643$ Тл) [*Weiss P., Forrer R.*— Ann. Phys., **15**, 153 (1926)].

чение никак не согласуется с температурами Кюри, превышающими для железа и кобальта 1000 К, как видно из табл. 5.4.

Для Вейсса это было очень разочаровывающее заключение, в особенности потому, что его модель вполне удовлетворительно объясняла ход спонтанной намагниченности при температурах ниже T_c (рис. 5.15) и температурную зависимость магнитной восприимчивости выше этой температуры [выражение (5.90)]. Однако типичные значения температуры Кюри около 1000 К требуют энергии упорядочения 0,1 эВ или около того (что согласуется с «эффективным полем» $H_e\sim10^{10}$ А/м и «фактором эффективного поля» или *коэффициентом Вейсса* $\gamma\sim10^4$). Логическое объяснение очень сильного взаимодействия, благодаря которому возникает ферромагнитное состояние, не могло быть найдено до тех пор, пока не наступил золотой век квантовой механики.

Это объяснение было предложено Гейзенбергом [15], который показал, что сильную энергетическую взаимосвязь соседних магнитных атомов можно объяснить как результат электростатического взаимодействия при учете принципа Паули. Гейзенберг отметил, что некоторое перекрытие электронных облаков соседних атомов оказывает влияние на характер размещения электронов по различным спиновым состояниям в частично заполненных внутренних оболочках. Если два соседних магнит-

[15] *Heisenberg W.*— Z. Physik, **49**, 619 (1928). Это — оригинальная работа по ферромагнитному обмену. Обменный механизм для непроводящих ферромагнетиков обсуждается в книге: *Anderson P. W.* Magnetism, v. 1, eds. G. T. Rado, H. Suhl, Academic Press, 1963.

ных атома имеют спины $\mathbf{S}_i$ и $\mathbf{S}_j$, то их потенциальную энергию можно выразить в виде

$$U = \text{const} - 2J_e\, \mathbf{S}_i \cdot \mathbf{S}_j, \tag{5.95}$$

где J_e — *обменный интеграл*, вычисленный для потенциала электростатического взаимодействия между этими двумя атомами. Второй член в правой части выражения (5.95) известен как *обменная энергия*; принято говорить, что она является следствием действия *обменных сил*.

Способ записи обменной энергии в выражении (5.95) показывает, что спин одного атома непосредственно связан со спином соседнего атома и что это взаимодействие не является случайным. Из квантовой механики следует, что обменный интеграл является *отрицательной* величиной для *малых* расстояний r_{ab} между соседними магнитными атомами. Это делает конфигурацию с параллельными спинами энергетически невыгодной (именно такой результат представлен на рис. 1.5 для молекулы водорода, который мы получили в модели Гайтлера—Лондона; σ_u-конфигурация с параллельными спинами энергетически существенно невыгодна по сравнению с σ_g-конфигурацией с антипараллельными спинами). Однако когда r_{ab} больше некоторого критического расстояния, обменный интеграл может быть положительным (рис. 5.16), и тогда параллельное выстраивание спинов оказывается энергетически выгодным. Таким образом, ферромагнетизм связан с ограничением на минимум расстояния между соседними магнитными атомами, что и объясняет возникновение ферромагнетизма в Fe, Co, Ni и его отсутствие у предшествующих переходных элементов периодической системы. Для возникновения ферромагнетизма требуется также, чтобы расстояние между соседними магнитными атомами не было слишком велико, поскольку, как схематически представлено на рис. 5.16, обменный интеграл при очень больших r_{ab} убывает до нуля.

Итак, состояния в атомах с непарными электронными спинами заполняются в соответствии с принципом Паули таким образом, что *электростатическая* энергия оказывается минимальной. Именно поэтому происходит параллельное упорядочение спинов (и возникает спонтанная намагниченность) при соответствующем расстоянии r_{ab}. Отсюда следует, что ферромагнетизм — это в каком-то смысле результат случайного совпадения. Такой элемент как марганец, например, обладает большим непарным спином на атом, но в элементарной форме не является ферромагнетиком из-за малой величины межатомного расстояния r_{ab}. Однако в некоторых соединениях (таких, как MnSb и MnAs) расстояния между атомами марганца возра-

стают настолько, что обменный интеграл становится положительным.

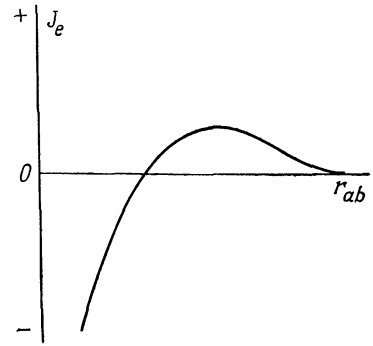


Рис. 5.16. Зависимость обменного интеграла от расстояния r_{ab} между двумя магнитными атомами.

Предложенная Гейзенбергом картина обменных сил, которые приводят к упорядочению спинов, не перечеркивает результаты теории молекулярного поля для ферромагнетизма, изложенные при обсуждении выражений (5.87)—(5.93). Наоборот, обменные силы можно считать той физической причиной, которая обеспечивает требуемую для упорядочения величину поля H_e. Физическая природа ферромагнетизма до сих пор часто обсуждается на языке теории молекулярного поля Вейсса, хотя уже разработаны более точные теории эффективного поля [16]. В таких теориях делаются попытки объяснить отклонение спонтанной намагниченности реальных твердых тел от тех зависимостей, которые допускаются уравнением (5.93). Более сложные теории объясняют также, почему парамагнитная восприимчивость выше T_c для любого реального твердого тела всегда отклоняется от закона Кюри—Вейсса [выражение (5.90)].

Одна такая теория была предложена П. Р. Вейссом [17] (его следует отличать от автора первой теории молекулярного поля — П. Вейсса). П. Р. Вейсс отказался от идеи о том, что молекулярное поле должно быть пропорциональным средней намагниченности. Он предположил, что каждый момент стремится расположиться параллельно соседним спинам, если их направление заметно отклоняется от направления средней намагниченности. Таким образом, согласно Вейссу, при температурах, близких к T_c, может существовать много больших скоплений магнитных атомов со спинами, выстроенными в одном направлении в пределах каждого скопления. Такие скопления по мере возрастания температуры должны распадаться на отдельные парамагнитные атомы, но при температурах, не намного превышающих T_c, отклик на приложенное магнитное поле должен быть меньше той величины, которая следует из закона Кюри—Вейсса [выражение (5.90)]. Как видно из

[16] В книге: *Смарт Дж.* Эффективное поле в теории магнетизма. — М.: Мир, 1968 обсуждаются эти более сложные теории, включая модель обменно-связанных пар Огучи и кластерную модель Бете — Пайерлса — Вейсса.

[17] *Weiss P. R.* — Phys. Rev., **74**, 1493 (1948); см. также примечание 16.

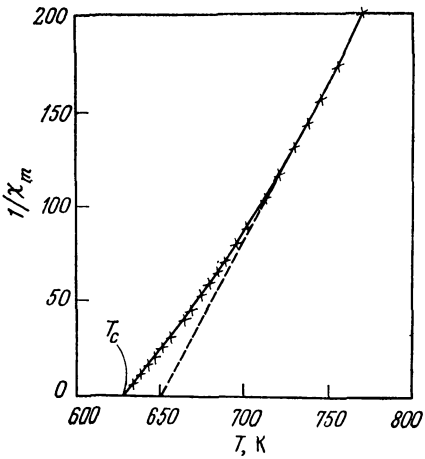


Рис. 5.17. Температурная зависимость обратной величины парамагнитной восприимчивости для никеля при температурах выше точки Кюри [*Weiss P., Forrer R.*— Ann. Physique, **15**, 153 (1926)]. Наблюдаемое на рисунке отклонение от закона Кюри — Вейсса характерно и для других ферромагнетиков.

рис. 5.17, для реального твердого тела возрастание $(1/\chi_m)$ с увеличением температуры отклоняется от закона Кюри—Вейсса именно так, как и следует ожидать, если образование скоплений оказывается существенным.

Спиновые волны

Предполагается, что при абсолютном нуле спины в ферромагнетике полностью упорядочены. Такое состояние спиновой системы мы можем рассматривать как основное; тогда состояние, при котором один спин или более направлены в противоположную сторону, будет возбужденным.

В рамках гейзенберговских представлений об эффективном поле, связывающем зарядовую конфигурацию одного магнитного атома с конфигурацией его соседей, переворот одного спина требует дополнительной энергии $\Delta U \approx 8J_e S^2$. Однако необходимая дополнительная энергия много меньше, если в кристалле возбуждается *спиновая волна*. Этим термином обозначают частичное возбуждение некоторого числа спинов, которое распространяется по системе спинов подобно волне. *Магнон* является квантом энергии спиновой волны, а возбуждение каждого магнона уменьшает полную намагниченность на один спин. Законы дисперсии для магнонов в спиновой системе ферромагнитного материала при низких температурах грубо можно сравнить с законами дисперсии для фононов при описании смещений атомов на основе волновых представлений.

Понятие спиновых волн было введено Ф. Блохом в 1930 г.[18] Используя представление о спиновых волнах, он показал, что

[18] *Bloch F.*— Z. Physik, **61**, 206 (1930).

спонтанная намагниченность при возрастании температуры должна уменьшаться по закону

$$M_s = M_{s0}\,[1 - A\,(k_0 T/J_e)^{3/2}], \qquad T \ll T_c, \tag{5.96}$$

где M_{s0} соответствует насыщению спонтанной намагниченности при абсолютном нуле, а числовой коэффициент A определяется кристаллической структурой. Выражение (5.96) называют *законом $T^{3/2}$ Блоха,* и оно хорошо описывает уменьшение M_s в области низких температур (см. задачу 5.10).

Несмотря на этот успех, представления о спиновых волнах для анализа явлений в ферромагнетиках практически не использовались, поскольку отсутствовало прямое экспериментальное наблюдение спиновых волн. Эти представления получили существенно большее распространение с тех пор, как было обнаружено резонансное поглощение на спиновых волнах в тонких ферромагнитных пленках. В одном из таких экспериментов, выполненных Сивеем и Танненвальдом [19], микроволновое (СВЧ) магнитное поле прилагалось параллельно поверхности тонкой ферромагнитной пленки, а статическое магнитное поле — перпендикулярно этой поверхности. В процессе изменения статического магнитного поля было выявлено несколько значений его напряженности, при которых наблюдалось резкое возрастание поглощения микроволн. Это поглощение связано с возбуждением таких стоячих волн, для которых на толщине пленки укладывается целое число длин волн. Стоячие волны образуются из спиновых волн с такой конфигурацией спинов на поверхностях пленки, которая определяется магнитной анизотропией плоскости поверхности [20].

В другом эксперименте со спиновыми волнами энергия магнонов преобразуется в энергию фононов, когда скорости спиновых и упругих волн совпадают. Это достигается путем подбора подходящего магнитного поля. В ферромагнетике, в котором существует неоднородное магнитное поле (например, размагничивающее поле в образце, имеющем форму сплошного цилиндра), спиновые волны могут быть возбуждены в той его части, которая находится под действием более сильного поля. Далее при передаче энергии в часть образца с более слабым полем, в которой фононы имеют бóльшую скорость, чем магноны, спиновые волны преобразуются в колебания решетки. Успешно такие эксперименты могут быть проведены в ферро-

[19] Постановка эксперимента для наблюдения резонансного поглощения на спиновых волнах была предложена в работе: Kittel C.— Phys. Rev., **110**, 1295 (1958), а первые экспериментальные результаты были получены в работе: *Seavey M. H., Jr., Tannenwald P. E.*— Phys. Rev. Lett., **1**, 168 (1958).

[20] «Пиннинг» (закрепление) ориентаций спинов на поверхностях пленки обсуждается в работе: *Wigen P. E. et al.*— J. Appl. Phys., **34**, 1137 (1963).

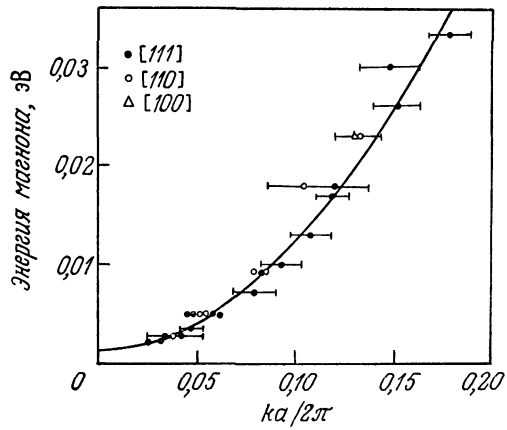


Рис. 5.18. Зависимость энергии магнона от безразмерного волнового вектора спиновой волны для сплава, состоящего из 92 % кобальта и 8 % железа. Данные получены при комнатной температуре в экспериментах по неупругому рассеянию нейтронов [*Sinclair R. N., Brockhouse B. N.*— Phys. Rev., **120**, 1638 (1960)].

магнитных материалах, которые обладают малыми акустическими потерями [21].

В гл. 2 рассматривались законы дисперсии для квантованных колебаний решетки и было показано, каким образом из экспериментов по неупругому рассеянию нейтронов можно получить соотношение между энергией фонона и его волновым вектором. При этом равенства (2.18) и (2.19) были использованы как законы сохранения для описания процессов рождения и уничтожения фононов. Совершенно аналогично неупругое рассеяние медленных нейтронов может быть использовано для изучения процессов рождения и уничтожения магнонов в ферромагнетике. Полученная таким образом дисперсионная кривая представлена на рис. 5.18. Как показано на этом рисунке, энергия магнонов в отличие от энергии акустических фононов растет практически пропорционально квадрату приведенного волнового вектора. Такая зависимость энергии магнона от волнового вектора согласуется с законом $T^{3/2}$ Блоха, что и предлагается проверить при решении задачи 5.10.

Рис. 5.18 также показывает, что магноны с бесконечной длиной волны обладают хотя и малой, но конечной энергией.

[21] Исследования взаимодействий спиновых и акустических волн в железо-иттриевом гранате (ЖИГ) и временные задержки сигнала «спинового эха» описаны, например, в работе: *Damon R. W., van de Vaart H.*— Proc. IEEE, 53, 348 (1965). В действительности ЖИГ является ферримагнитным материалом (этот термин поясняется в конце данного раздела).

Для результатов, приведенных на рис. 5.18, это, по крайней мере отчасти, обусловлено ограничениями в разрешающей способности прибора, однако можно ожидать, что минимальная энергия магнона будет отлична от нуля для любого ферромагнетика из-за дальнодействия эффектов анизотропии.

Магнитные домены

Вполне вероятно, что образец ферромагнетика в целом окажется *ненамагниченным*. Такое может иметь место при случайной ориентации *магнитных доменов*, несмотря на то что каждый домен будет обладать значительной намагниченностью M_s. Каждый домен отделен от соседнего домена доменной стенкой или *стенкой Блоха* (блоховской границей), в пределах толщины которой направление спинов плавно поворачивается, как показано на рис. 5.19.

Число доменов внутри кристалла, их форма и взаимосвязи определяются кристаллографическими направлениями, вдоль которых могут выстраиваться спины, а также требованием минимизации энергии, которая складывается из энергии доменных стенок и энергии размагничивания образца. На рис. 5.20 показано расположение доменов относительно друг друга для

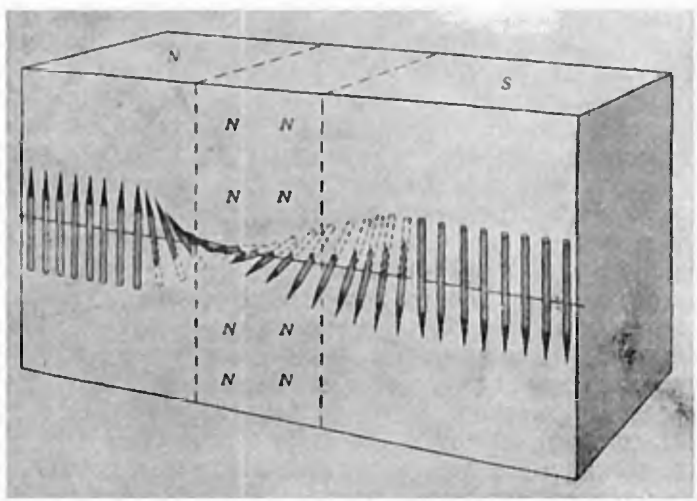


Рис. 5.19. Поворот направления спина в пределах толщины блоховской стенки, разделяющей магнитные домены [*Киттель Ч.* Введение в физику твердого тела. — М.: Мир, 1963]. Толщина такого слоя в железе составляет примерно 300 постоянных решетки.

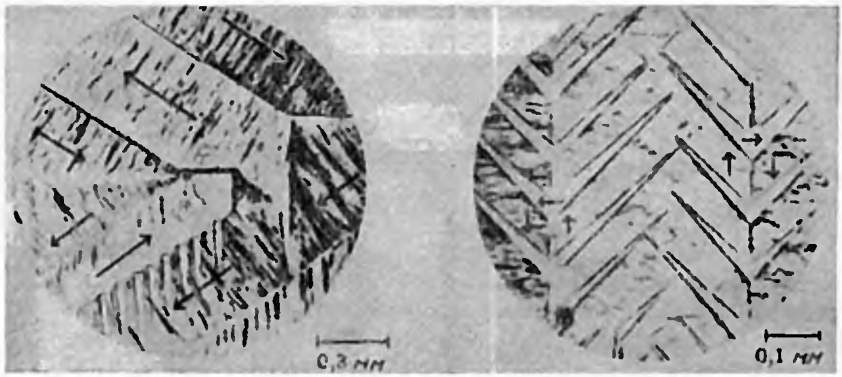


Рис. 5.20. Два примера магнитной доменной структуры, видимой под микроскопом [*Chikazumi S.*— Physics of Magnetism, Krieger, 1978]. Справа — структура железа, содержащего 4 % кремния; направления легкого намагничивания совпадают с направлениями типа [100], как видно для поверхности, которая практически соответствует плоскости (100), домены разделены 180- и 90-градусными стенками. Слева — домены в пермаллое (78 % Ni, 22 % Fe), для которого направления легкого намагничивания совпадают с направлениями типа [111]; возникающие при этом домены разделены 180-, 109- и 71-градусными стенками.

двух типов ферромагнетиков; в одном из них направления легкого намагничивания совпадают с ребрами куба, а в другом — с объемными диагоналями куба.

На рис. 5.21 состояние явно «ненамагниченного» образца представлено началом координат. При включении внешнего магнитного поля блоховские границы движутся таким образом, что происходит рост доменов, в которых M_s достаточно близко к направлению H. Это движение доменных стенок может быть обратимым, если приложенное поле невелико, но оно приобретает необратимый характер в случае сильного поля. Если произошло такое необратимое движение, то образец приобретает остаточную намагниченность даже если поле убрано.

На рис. 5.21 показано увеличение намагниченности практически до насыщения, которое обеспечивается не только движением блоховских границ, но и *поворотами доменов* в более сильных полях. Периодические изменения направления приложенного поля (положительное или отрицательное направление) приводят к хорошо известной *петле гистерезиса*, устанавливающей связь между магнитной индукцией и приложенным полем. Площадь, ограниченная этой кривой, представляет собой работу на единицу объема, выполненную при движении вдоль контура петли.

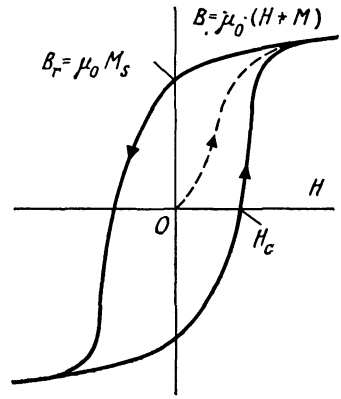


Рис. 5.21. Схематическое изображение петли намагничивания для многодоменного образца ферромагнетика. H_c — *коэрцитивная сила*, B_r — *остаточная намагниченность*. Пунктиром представлено намагничивание предварительно намагниченного образца (со случайной ориентацией доменов). Стрелки на сплошной кривой показывают ход петли гистерезиса.

Технические аспекты ферромагнетизма непосредственно связаны с петлями намагниченности для реальных твердых тел; с «магнитожесткими» материалами, обладающими большой коэрцитивной силой, с материалами, обладающими большой магнитной проницаемостью и т. д. С необъятной литературой по этим вопросам можно частично познакомиться по библиографии, приведенной в конце настоящей главы.

Антиферромагнетизм и ферримагнетизм

Модель магнитного упорядочения, при котором твердое тело не проявляет значительной намагниченности из-за того, что соседние диполи *антипараллельны* по отношению друг к другу, была предложена Неелем [22]. Этот тип упорядочения был назван *антиферромагнетизмом*, и в настоящее время выявлено множество твердых тел, обладающих подобными свойствами [23].

Состоянию с минимальной энергией двух соседних магнитных атомов, для реализации которого требуется антипараллельная ориентация их спинов, в модели Гейзенберга, очевидно, соответствует отрицательное значение обменного интеграла J_e. Однако существование отрицательного значения J_e еще не является достаточным условием для возникновения антиферромагнитного упорядочения в реальном трехмерном кристалле. Кристалл должен также обладать таким трехмерным расположением магнитных ионов, которое способствует образованию *магнитных подрешеток* с противоположно направленными намаг-

[22] Теория антиферромагнетизма, предложенная Неелем [*Néel L.* — Ann. Phys., 17, 5 (1932)], была видоизменена Ван-Флеком [*van Vleck J. H.* — J. Chem. Phys., 9, 85 (1941)] и в таком виде применяется в настоящее время.

[23] Информация о многих антиферромагнитных телах собрана в таблицах в источниках, указанных в примечании 11.

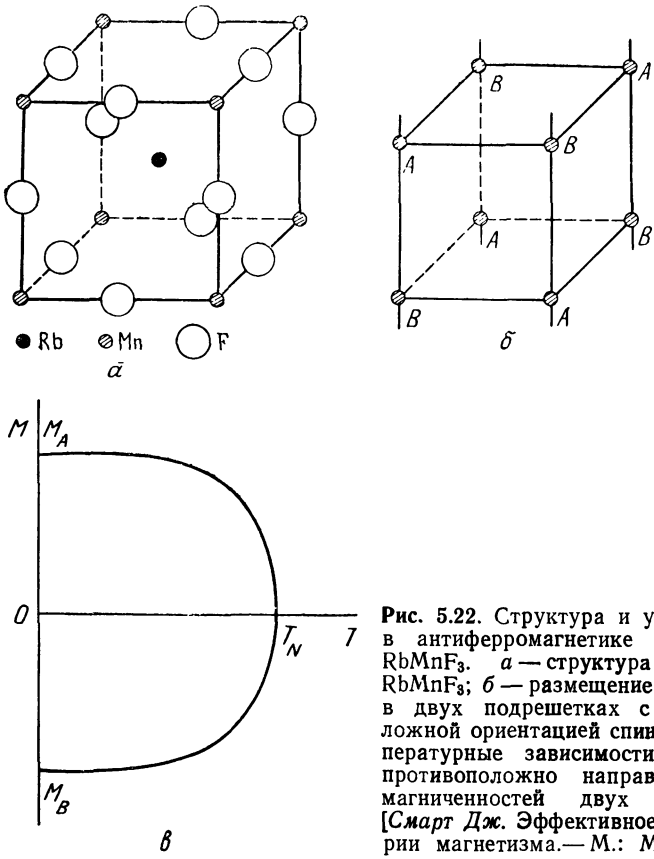


Рис. 5.22. Структура и упорядочение в антиферромагнетике на примере $RbMnF_3$. *а* — структура кристалла $RbMnF_3$; *б* — размещение ионов Mn^{2+} в двух подрешетках с противоположной ориентацией спинов; *в* — температурные зависимости равных и противоположно направленных намагниченностей двух подрешеток [*Смарт Дж.* Эффективное поле в теории магнетизма.— М.: Мир, 1968].

ниченностями. При описании некоторых кристаллических структур оказывается удобным представить, что магнитные ионы расположены в четырех и более подрешетках. Однако можно показать, что для общего представления об антиферромагнетизме достаточно ограничиться простейшим случаем, когда имеются всего две противоположно направленные магнитные подрешетки. Проиллюстрировать это можно на примере ионов Mn^{2+} в соединении $RbMnF_3$ (рис. 5.22).

Кристаллическая решетка $RbMnF_3$ позволяет разделить всю систему ионов Mn^{2+} на подрешетку *A*, в которой каждый ион окружен шестью ближайшими *магнитными* соседями, принадлежащими подрешетке *B*, и подрешетку *B*, в которой каждый ион также имеет шесть ближайших магнитных соседей, но из подрешетки *A*. Таким образом, если между соседними магнитными атомами обменное взаимодействие оказывается отрица-

тельным и все обменные интегралы J_e для более удаленных соседей можно считать равными нулю, магнитные моменты подрешеток A и B спонтанно ориентируются во взаимно противоположных направлениях. Антиферромагнитное упорядочение должно быть полным при абсолютном нуле, но оно начинает нарушаться при повышении температуры и исчезает полностью при температуре Нееля T_N.

Чтобы описать влияние обменного взаимодействия на ориентацию спина магнитного иона, можно опять воспользоваться представлениями теории молекулярного поля. Молекулярное поле, действие которого испытывает на себе ион Mn^{2+}, принадлежащий подрешетке A (рис. 5.22), по аналогии с соотношением (5.87) для ферромагнетика можно описать выражением

$$H_A = -\gamma M_B. \tag{5.97}$$

Тогда по аналогии с уравнением (5.88) намагниченность подрешетки A будет равна

$$M_A = (C/2T)[H - \gamma M_B], \tag{5.98}$$

где H — любое нелокальное магнитное поле, а $C/2$ — константа Кюри для половины магнитных ионов кристалла. Точно так же намагниченность подрешетки B будет равна

$$M_B = (C/2T)[H - \gamma M_A]. \tag{5.99}$$

Для совместного решения уравнений (5.98) и (5.99) необходимо, чтобы M_A и M_B не были равны нулю в отсутствие внешнего поля при температуре, равной или меньшей $T_N = (\gamma C/2)$.

Наблюдаемая величина магнитной восприимчивости может быть определена из уравнений (5.98) и (5.99) таким же образом, как это было сделано выше при получении соотношения (5.90) для ферромагнитных твердых тел.

Из этих двух уравнений находим, что

$$\chi_m = \frac{M_A + M_B}{H} = \frac{C}{T + T_N}, \qquad T \geqslant T_N. \tag{5.100}$$

Приведенное выражение недостаточно хорошо согласуется с экспериментальными данными. В действительности поведение антиферромагнетика при высоких температурах подчиняется следующей формуле:

$$\chi_m = \frac{C}{T + \theta}, \qquad T \geqslant T_N, \tag{5.101}$$

а отношение θ/T_N обычно больше единицы. Это противоречие можно объяснить в рамках более сложных моделей магнитного взаимодействия, которые учитывают влияние не только

ближайших, но и более удаленных магнитных атомов. Проверить это мы предлагаем читателю в задаче 5.11.

Магнитная восприимчивость антиферромагнетиков уменьшается с понижением температуры ниже температуры Нееля, как следует из рис. 5.11, *б*. Причина состоит в том, что обменные взаимодействия «запирают» магнитные моменты в состоянии антипараллельного упорядочения, что подавляет реакцию на приложенное магнитное поле. Магнитная восприимчивость монокристаллического образца ниже температуры Нееля обычно имеет значительную анизотропию, поскольку много легче слегка «наклонить» все спины приложением магнитного поля, перпендикулярного цепочкам противоположно направленных спинов, чем вызвать «опрокидывание спинов» полем, направленным вдоль этих цепочек.

В 1948 г. Неель[24] рассмотрел, что произошло бы, если магнитные ионы в кристалле можно было бы разделить на две группы или две подрешетки с *неравными* магнитными моментами. Если молекулярное поле, действие которого испытывает на себе ион подрешетки A со стороны магнитных соседей из подрешетки B (и наоборот), определяется отрицательным обменным интегралом, то такое твердое тело является *ферримагнетиком*; при низких температурах оно обладает спонтанной намагниченностью, определяемой разностью намагниченностей двух подрешеток.

Предложенное Неелем понятие ферримагнитного упорядочения было сразу использовано для объяснения магнитных свойств целой группы твердых тел — *ферритов*. Состав ферритов имеет общую формулу $MO \cdot Fe_2O_3$, где M — двухвалентный металлический катион, например Fe, Ni, Co, Cu, Mg, Zn или Cd. Таким образом, группа ферритов включает в себя и Fe_3O_4 — природный магнетит, который использовали древние мореплаватели. Магнитный момент магнетита (только $4\mu_B$ на молекулу) много меньше той величины, которую можно было ожидать при обычном ферромагнитном упорядочении всех спинов. Неель показал, что такой магнитный момент обеспечивается одним лишь ионом Fe^{2+} в молекуле, в то время как два иона Fe^{3+} каждой молекулы занимают *неэквивалентные кристаллографические положения* и принадлежат двум противоположно направленным магнитным подрешеткам.

Спонтанная намагниченность ферримагнетика уменьшается от некоторого максимального значения при очень низких температурах до нуля при верхнем пределе по температуре T_f, что в общих чертах подобно поведению ферромагнитного материала. При этом соображения, изложенные при выводе уравне-

[24] *Néel L.*— Ann. Physique, 3, 137 (1948).

ния (5.93), сохраняют силу и для данной ситуации. При температурах, превышающих T_f, поведение магнитной восприимчивости оказывается более сложным, чем следует из закона Кюри—Вейсса, что представлено пунктирной кривой на рис. 5.11, *г*. Неель[24] показал, что обратную величину магнитной восприимчивости можно представить выражением (см. задачу 5.12):

$$(1/\chi_m) = \left(\frac{T-\theta}{C}\right) - \left(\frac{W}{T-\theta_0}\right), \qquad T > T_f, \tag{5.102}$$

где C — как и ранее, константа Кюри, W и θ_0 — константы, величины которых определяются различными значениями коэффициентов молекулярного поля (или обменных интегралов), а θ — параметр, численное значение которого заметно меньше T_f.

Существует заметный практический интерес к магнитным свойствам некоторых кристаллов из семейства ферритов, которые кристаллизуются в структуре шпинели. Эти материалы являются диэлектриками, а именно это свойство магнитных материалов требуется для многих практических целей. Созданы ферриты, обладающие практически прямоугольной петлей гистерезиса; такие тороидальные ферритовые сердечники стали основой запоминающих устройств в большинстве компьютеров, созданных в 1960-х и начале 1970-х годов.

Другой группой ферримагнетиков, представляющей значительный интерес для практических целей, являются гранаты на основе железа — вещества состава $M_3Fe_5O_{12}$, где M — катион трехвалентного металла, такого, как иттрий или одного из редкоземельных элементов. Гранаты использовались в экспериментах по магнитному резонансу и при исследованиях взаимодействий спиновых волн.

Термины *ферромагнетизм, ферримагнетизм* и *антиферромагнетизм* описывают лишь простейшие из возможных типов спонтанного магнитного упорядочения в кристаллических твердых телах. Имеются твердые тела, в которых конфигурация с минимальной энергией требует геликоидального (винтообразного) упорядочения спинов, и другие твердые тела, в которых спины подрешеток скорее наклонены относительно друг друга, чем просто взаимно противоположны. Эти экзотические системы подробно рассматриваются в литературе, приведенной в конце этой главы.

5.3. Магнитный резонанс

Магнитный момент заряженной, вращающейся относительно собственной оси частицы (электрона или атомного ядра), взаимодействует с макроскопической магнитной индукцией **B**, и это

приводит к расщеплению энергетического уровня на два или большее число состояний. Переходы между этими состояниями, подчиняющиеся соответствующим правилам отбора, могут быть стимулированы поглощением фотонов. В тех случаях, когда энергия радиочастотного кванта соответствует расстоянию между двумя индуцированными магнитным полем энергетическими состояниями электрона или ядра, наблюдается *магнитный резонанс.*

Концепция электронного спина и связанного с ним магнитного момента была предложена Гаудсмитом и Уленбеком в 1925 г. для объяснения пространственного квантования, обнаруженного в 1922 г. в опытах Штерна и Герлаха, а также для объяснения сложной структуры атомных спектров и «аномального» эффекта Зеемана. В 1924 г. Паули высказал мысль о том, что сверхтонкая структура атомных спектров возникает в результате взаимодействия между находящимися на орбите электронами и небольшим магнитным моментом ядра, связанным с его вращением вокруг собственной оси.

Мы можем получить грубую оценку величины магнитных моментов, если мысленно представим некоторый сферический объект, имеющий массу M, которая равномерно распределена по его объему, и заряд Q, равномерно распределенный по его поверхности. Если этот объект вращается вокруг собственной оси таким образом, что его момент количества движения равен L_s, то циркуляция заряда создает магнитный момент μ, который оказывается соосным с L_s и прямо пропорционален ему. Нетрудно проверить (см. задачу 5.13), что

$$\mu = \gamma L_s = (5Q/6M)L_s \tag{5.103}$$

и не зависит от радиуса сферы. Коэффициент пропорциональности γ называется *гиромагнитным отношением* вращающегося вокруг собственной оси объекта. Если принять во внимание квантование момента количества движения $L_s = [\hbar^2 J(J+1)]^{1/2}$, то магнитный момент равен

$$\mu = (5Q\hbar/6M)\sqrt{J(J+1)}. \tag{5.104}$$

В действительности магнитные моменты электрона или атомного ядра из-за эффектов квантовой электродинамики, а также из-за сложного характера распределения массы и заряда отличаются от тех величин, которые следуют из выражения (5.104). Магнетон Бора и ядерный магнетон

$$\mu_B = \frac{e\hbar}{2m} = 9{,}274 \cdot 10^{-24} \text{ Дж/Тл},$$

$$\mu_n = \frac{e\hbar}{2M_p} = 5{,}051 \cdot 10^{-27} \text{ Дж/Тл} \tag{5.105}$$

уже упоминались в начале разд. 5.2 как единицы, в которых обычно выражаются магнитные моменты вращающихся вокруг собственной оси электронов и атомных ядер.

Простая картина электронного парамагнитного резонанса

Спектры ЭПР неспаренных электронов в твердых телах часто оказываются очень сложными, и в действительности вклад ЭПР в расширение наших знаний в области физики твердого состояния определяется именно деталями спектров ЭПР. Однако сначала обсудим основное условие для ЭПР-поглощения. Для этого рассмотрим поведение спинов изолированных электронов в однородной среде.

Электрон обладает моментом количества движения $L_s = \hbar\sqrt{3/4}$ при двух значениях проекции спина $m_s = \pm 1/2$. На рис. 5.23, *а* показаны два возможных угла наклона между направлением спинового момента и магнитным полем **B**. Соответствующая проекция магнитного момента на направление **B** обычно записывается в виде

$$\mu = -g\mu_B m_s. \tag{5.106}$$

Здесь безразмерная величина g — фактор спектроскопического расщепления или g-фактор. (Как отмечалось в разд. 5.2, $g = 2{,}0023$ для изолированного свободного электрона, и это дает величину магнитного момента, почти в точности совпадающую с магнетоном Бора. Однако g-фактор для электронов в твердом теле может радикально отличаться от 2).

Неспаренный электрон, обладающий энергией U_0 в отсутствие магнитного поля, при включении поля приобретает дополнительную энергию $\Delta U = -\mu B$. Таким образом, как показано на рис. 5.23, *б*, электрон может находиться в одном из двух энергетических состояний

$$U = U_0 + \Delta U = U_0 - \mu B = U_0 \pm \frac{1}{2} g\mu_B B, \tag{5.107}$$

соответствующих двум возможным значениям $m_s = \pm 1/2$ (со спином верх или со спином вниз).

Для заселенностей N_U и N_L верхнего и нижнего состояний в случае теплового равновесия можно записать

$$(N_U/N_L)_{\text{равн}} = \exp(-g\mu_B B/k_0 T). \tag{5.108}$$

Обозначим через ΔN_0 равновесное значение разности заселенностей $\Delta N_0 = (N_L - N_U)$. Поскольку температура, при которой

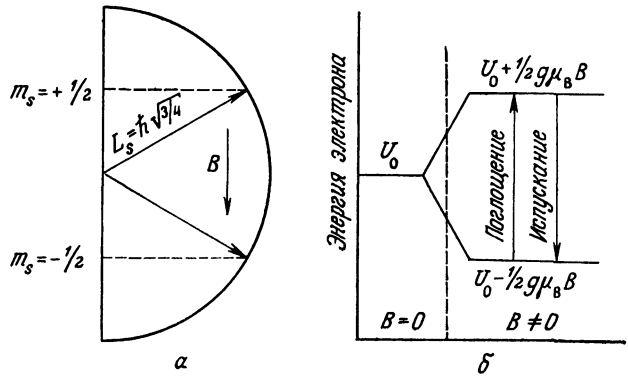


Рис. 5.23. *а* — возможные ориентации спинового момента электрона относительно направления внешнего магнитного поля; проекция магнитного момента пропорциональна m_s. *б* — расщепление энергетического уровня электрона в магнитном поле; переходы между верхним и нижним уровнями удовлетворяют законам сохранения энергии и момента количества движения, если поглощаются или испускаются фотоны с энергией $g\mu_BB$.

выполняются эксперименты по ЭПР, обычно заметно больше, чем $(g\mu_BB/k_0)$,

$$\Delta N_0 \simeq N(g\mu_BB/2k_0T) \tag{5.109}$$

обычно существенно меньше полной плотности спинов $N=(N_L+N_U)$. Тем не менее посредством внешних воздействий можно достичь значений ΔN, меньших ΔN_0. В экспериментах по ЭПР это достигается за счет резонансного поглощения радиочастотных фотонов с энергией

$$h\nu = \hbar\omega = g\mu_BB, \tag{5.110}$$

поскольку эти фотоны могут удовлетворять законам сохранения энергии и момента количества движения при переходах между двумя уровнями, соответствующими равенству (5.107).

Как показано на рис. 5.24 и последующих рисунках, ЭПР и ЯМР обычно наблюдаются на фиксированной радиочастоте ν, а магнитное поле B изменяется в некотором интервале, внутри которого находится и его резонансное значение. Комбинация B (Тл) и ν (ГГц), удовлетворяющих условию резонанса, выражается соотношениями:

$$\nu = 14{,}00\ gB,$$
$$g = 0{,}07145\,(\nu/B). \tag{5.111}$$

Таким образом, для свободного электрона с $g=2{,}0023$ ЭПР можно наблюдать на микроволновой частоте ~ 10 ГГц (длина

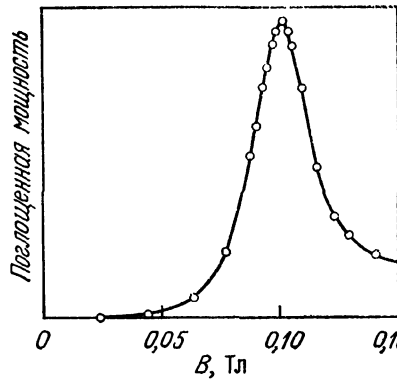


Рис. 5.24. Электронный парамагнитный резонанс на неспаренных спинах ионов Mn^{++} в $MnSO_4$ при комнатной температуре на частоте $\nu=2{,}75$ ГГц [*Zavoisky E.*— J. Phys. USSR, **10**, 197 (1964)]. Как обычно при выполнении экспериментов по ЭПР, частота ν фиксирована, а поле **B** изменяется и проходит через значение, соответствующее резонансу. Положение пика поглощения согласуется с уравнением (5.108) для $g=1{,}95$.

волны ~3 см) в полях ~0,4 Тл или в меньших полях на более низких частотах.

Первые наблюдения ЭПР были выполнены Завойским [25] на радиочастотах в несколько МГц и в полях менее 0,01 Тл. При этих условиях резонанс едва разрешим. Впоследствии Завойским было получено много лучшее разрешение при более высоких ν и B. В качестве примера на рис. 5.24 приведены его более поздние результаты. Преимущества высоких частот и сильных магнитных полей при наблюдении ЭПР позднее были продемонстрированы в ранних работах, выполненных в Соединенных Штатах [26] и Великобритании [27]. В настоящее время для наблюдения ЭПР почти всегда используются частоты микроволнового диапазона.

Применение сильных полей и высоких частот не только улучшает разрешение, но также повышает вероятность поглощения. Соотношение (5.109) наводит на мысль, что для наблюдения резонанса желательно использовать низкие температуры и высокие частоты, так как

$$\Delta N_0 \simeq N(g\mu_B B/2k_0T) = N(h\nu/2k_0T). \qquad (5.112)$$

Здесь знак приближения напоминает о том, что выражение (5.109) было получено для температур, которые велики по сравнению с $(h\nu/2k_0)$. Обычно это условие выполняется, поскольку

[25] *Zavoisky E.*— J. Phys. USSR. **9**, 211, 245 (1945).
[26] *Cummerow R. L., Halliday D.*— Phys. Rev., **70**, 433 (1946).
[27] *Bagguley D. M. S., Griffiths J. H. E.*— Nature, **160**, 532 (1947).

$(h\nu/2k_0)\simeq 0{,}2$ К при $\nu\simeq 10$ ГГц. Хотя соотношение (5.112) указывает, что ΔN_0 пропорционально ν, в ряде экспериментальных ситуаций чувствительность ЭПР оказывается почти пропорциональной ν^2. Современные высокочувствительные спектрометры ЭПР могут детектировать и исследовать даже 10^{11} электронных спинов.

Спин-решеточная и спин-спиновая релаксация

Скорости переходов электронов вверх и вниз, которые индуцируются фотонами, пропорциональны соответственно N_L и N_U. Таким образом поглощение микроволновой мощности зависит от разности $\Delta N=(N_L-N_U)$, которая в экспериментах по ЭПР всегда положительна. Идеальным следует считать такие условия эксперимента, когда падающая микроволновая мощность настолько низкая, что ΔN остается практически равновесной (см. уравнение (5.112)) во всем интервале значений B, в котором происходит резонансное поглощение. Насколько велика может быть подводимая мощность, зависит от тех механизмов, которые возвращают спины в состояния с более низкой энергией, что, конечно, сопровождается передачей момента количества движения, равного $\hbar$.

Процесс так называемой *прямой спин-решеточной релаксации* зависит от спин-орбитального взаимодействия электронного спина с колебательным спектром кристаллической решетки твердого тела. В этом процессе электрон переходит в спиновое состояние с более низкой энергией и возникают длинноволновые акустические фононы[28]. Очевидно, что постоянная времени такого прямого процесса изменяется приблизительно как $(1/T)$.

В некоторых твердых телах при более высоких температурах с прямым процессом эффективно конкурирует *рамановский процесс спин-решеточной релаксации.* В этом процессе фононы кристаллической решетки неупруго рассеиваются атомами, у которых неспаренный электрон находится в верхнем энергетическом состоянии. Постоянная времени для рамановского процесса изменяется как $(1/T)^5$.

Комбинация описанных выше процессов спин-решеточной релаксации характеризуется постоянной времени τ_1. Таким образом, если ΔN в результате возмущения отличается от ΔN_0 и для ΔN существует возможность за время t_0 вернуться к равновесному значению, то можно записать, что

$$\Delta N=\Delta N_0\{1-A\exp[-(t-t_0)/\tau_1]\}. \qquad (5.113)$$

[28] Если $\nu\sim 10$ ГГц, генерируются фононы с энергией $\sim 4\cdot 10^{-5}$ эВ и длиной волны в несколько сотен межатомных расстояний.

Тогда τ_1 может быть измерено по затуханию отклика коллектива спинов на резкое выключение статического магнитного поля или изменение его направления на обратное (τ_1 можно измерить и в экспериментах по ЭПР). Большое число исследований, посвященных вопросам *парамагнитной релаксации*[29], было выполнено еще до того, как были достигнуты успехи в развитии ЭПР.

В случае ЭПР τ_1 связано со скоростью поглощения и рассеяния микроволновой мощности системой спинов без уменьшения ΔN до нуля. Можно ожидать, что ширина любой линии поглощения ЭПР также должна определяться величиной τ_1. В дальнейшем мы увидим, что это предположение соответствует действительности не всегда, а лишь в некоторых случаях.

Другая важная характеристика релаксации — время τ_2 — определяется вероятностью *спин-спиновой релаксации* и связана с влиянием (т. е. вкладом в локальное магнитное поле), оказываемым одним магнитным атомом на другие. Величина $1/\tau_2$ является константой скорости для потери фазовой когерентности коллектива спинов и приложенного электромагнитного поля. Когда коллектив спинов теряет фазовую когерентность, возникают повышенные флуктуации магнитного поля, через которые спины влияют друг на друга. В результате условия резонанса для различных спинов выполняются при различных значениях B. Это обстоятельство может быть основной причиной того, что ширина линии резонансного поглощения оказывается конечной. Приведенное время релаксации

$$\tau_r = \tau_1\tau_2/(\tau_1 + \tau_2) \qquad (5.114)$$

определяется более коротким временем релаксации τ_1 или τ_2.

На рис. 5.25 приведены значения времени релаксации τ_r, которые определены из ширины линии ЭПР на электронах в малых частицах металлического натрия. Хорошее совпадение экспериментальных точек с зависимостью вида $\tau_r = 3 \cdot 10^{-8}/T$ наводит на мысль о том, что τ_r определяется механизмом *прямой* спин-решеточной релаксации во всем исследованном интервале температур.

Естественно ожидать, что одно лишь время τ_1 будет контролировать ширину линии в тех случаях, когда малые количества атомов магнетика растворены в немагнитной матрице. При этом спин-спиновыми корреляциями можно пренебречь из-за

[29] Значительная часть ранних работ по парамагнитной релаксации была выполнена в Нидерландах. Обзор этих работ выполнен Гортером в книге: *Gorter C. J.* Paramagnetic Relaxation, Elsevier, 1947. [Имеется перевод: *Гортер К.* Парамагнитная релаксация.— М.: ИЛ, 1949.]; см. также: Physica, **14**, 504 (1948).

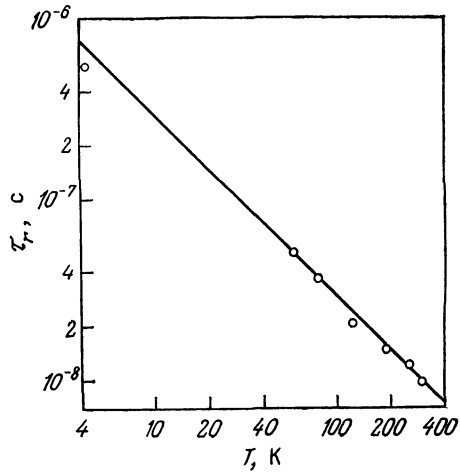


Рис. 5.25. Температурная зависимость времени релаксации для электронов в микроскопических частицах металлического натрия. Частицы имеют диаметр 5 мкм и меньше, поэтому их радиус оказывается меньше толщины скин-слоя на радиочастотах. Значения τ_r определены по ширине линии ЭПР для каждой температуры измерения. Одна из линий ЭПР (полученная при температуре 296 К) приведена на рис. 5.27 [*Feher G., Kip A. F.*— Phys. Rev., 98, 337 (1955)].

больших расстояний между атомами магнетика. Возрастание концентрации атомов магнетика может привести к тому, что τ_2 станет меньше τ_1, а это будет сопровождаться уширением линии ЭПР. Такое уширение линии ЭПР наблюдается для многих систем атомов магнетика.

Однако ширина линии ЭПР для некоторых систем взаимодействующих спинов на соседних атомах определяется не только механизмами релаксации. Когда энергия электростати-

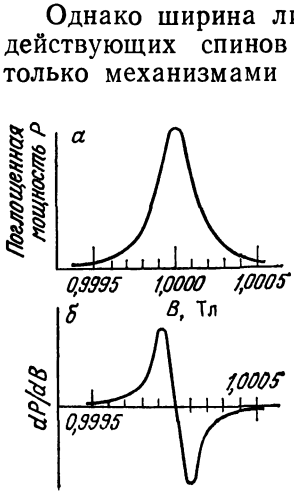


Рис. 5.26. *а* — линия поглощения ЭПР, *б* — производная линии поглощения для DPPH (дифенилпикрилгидразид) при комнатной температуре и ν=28 ГГц. Очень малая ширина линии $\Delta B_{1/2} = 3 \cdot 10^{-4}$ Тл обусловлена обменным сужением. DPPH часто используется для калибровки спектрометров ЭПР, поскольку для наблюдаемой у него линии g= =2,0036, что очень близко к g-фактору свободного электрона.

ческого взаимодействия электронов (представленная величиной J_e в выражении (5.95)) велика по сравнению с энергией магнитного диполь-дипольного взаимодействия, наблюдается сужение линии ЭПР. В таких случаях говорят, что малая ширина линии является следствием *обменного сужения*[30]. Линия ЭПР для образцов DPPH оказывается очень узкой (см. рис. 5.26) именно по этой причине.

Ширина и форма линии ЭПР

Положение линии ЭПР (отношение B к ν, или эффективный g-фактор), которое можно определить из экспериментов, несет в себе полезную информацию, однако обычно желательна информация большéго объема. Сведения об окружении спинов могут быть получены из анализа ширины и формы резонансной линии (а также ее тонкой и сверхтонкой структур, которые мы рассмотрим несколько позже).

При измерениях ЭПР обычно производится развертка поля B (при постоянной частоте ν) в интересующей нас области. Можно получить зависимость поглощаемой мощности P от B, однако бóльшая чувствительность может быть достигнута при использовании метода фазочувствительного детектора, который позволяет получить производную (dP/dB). На рис. 5.26 представлены поглощаемая мощность P и ее производная (dP/dB) для простой линии поглощения. Любая из этих зависимостей может быть использована для определения положения B_r максимума поглощения и ширины линии $\Delta B_{1/2}$ на ее полувысоте.

Динамика стимулированных переходов вверх и вниз для системы спинов, когда B линейно изменяется во времени и проходит через резонансное значение, была рассмотрена Блохом[31]. Его работа касалась ЯМР, однако она в равной степени ценна и для данной ситуации. Результаты Блоха говорят о том, что если уровень падающей мощности достаточно мал, а скорость развертки достаточно велика, так что ΔN незначительно отличается от ΔN_0 по всей области развертки магнитного поля, то можно предсказать лоренцеву форму линии:

$$P = P_{\text{макс}}[1 + (B - B_r)^2/\Gamma^2]^{-1}. \qquad (5.115)$$

Максимум P или ноль (dP/dB) имеют место при $B = B_r$. Ширина линии поглощения на полувысоте $\Delta B_{1/2} = 2\Gamma$, а расстояние между экстремумами (peak-to-peak) производной dP/dB

[30] *Van Vleck J. H.*— Phys. Rev., **74**, 1168 (1948).
[31] *Bloch F.*— Phys. Rev., **70**, 460 (1946).

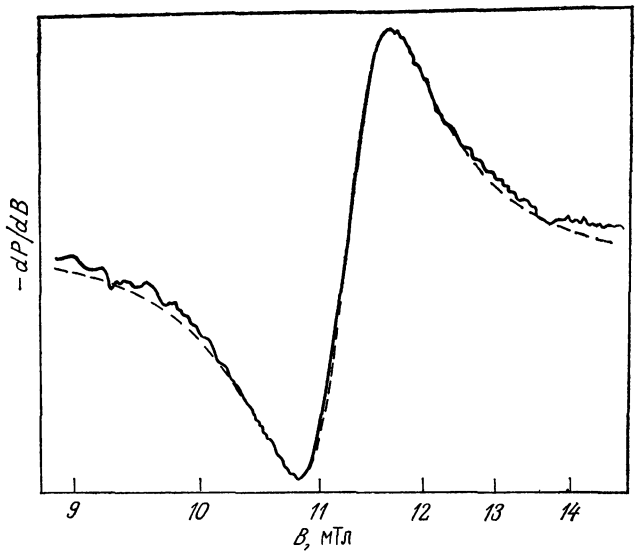


Рис. 5.27. Производная линии поглощения ЭПР малых частиц натрия (диаметр менее 5 мкм), полученная при 296 К и ν=320 МГц [*Feher G., Kip A.*— Phys. Rev., **98**, 33, 1955]. Штриховая кривая имеет лоренцеву форму и представляет производную выражения (5.115). Проверке формы линии и вычислению времени релаксации посвящена задача 5.14.

равно $\Delta B_{pp} = (2/\sqrt{3})\,\Gamma = \Delta B_{1/2}/\sqrt{3}$. В задаче 5.14 предлагается ширину линии в теслах преобразовать в ширину линии в частоте, которую часто можно отождествлять с $(1/\tau_r)$. Выражение для времени спиновой релаксации можно записать в виде

$$\tau_r \simeq 1{,}1 \cdot 10^{-11}/g\Gamma \text{ с}, \qquad (5.116)$$

если Γ выражена в теслах[32].

Когда ΔN в процессе наблюдения резонанса заметно отличается от ΔN_0, кривая поглощения по форме приближается к гауссовой:

$$P = P_{\text{макс}} \exp\left[-(B - B_r)^2/\Gamma^2 \ln 2\right]. \qquad (5.117)$$

Это выражение записано таким образом, что ширина на полувысоте, как и ранее, $\Delta B_{1/2} = 2\Gamma$, однако связь между Γ и τ_r отличается от зависимости (5.116), и ΔB_{pp} теперь представляет большую долю от $\Delta B_{1/2}$.

[32] Здесь уместно напомнить читателю, что практически во всей литературе по ЭПР и ЯМР используется система единиц СГС; при этом для обозначения напряженности поля не делается различия между терминами «гаусс» и «эрстед».

На рис. 5.27 производная линии поглощения ЭПР сопоставляется с кривой лоренцевой формы, а в задаче 5.14 читателю предлагается убедиться в том, что форма экспериментальной кривой отлична от гауссовой. Время релаксации, полученное из ширины линии на рис. 5.27, можно сравнить с данными рис. 5.25 для того же образца металлического натрия.

Детали спектров ЭПР

Мы уже рассмотрели самый простой спектр ЭПР, и теперь нам предстоит познакомиться с деталями спектров ЭПР, которые выявляются в большинстве экспериментов. Некоторые из этих деталей оказываются полезными для использования ЭПР с целью изучения структуры твердого тела, а также положения дефекта в кристаллической решетке и его симметрии.

Тонкая структура для $S>1/2$

Спектр ЭПР становится более сложным, если магнитный атом или ион имеет более одного неспаренного электрона, и поэтому его полное спиновое квантовое число оказывается больше, чем 1/2. Такая ситуация имеет место, например, в случае ионов Mn^{++}, спектр ЭПР для которых представлен на рис. 5.24. Каждый ион Mn^{++} имеет пять неспаренных $3d$-электронов; считается, что орбитальная часть момента количества движения «заморожена» для этих электронов. К такому заключению мы пришли при обсуждении рис. 5.13, когда убедились в том, что $S=5/2$ для иона Mn^{++}. Существует $(2S+1)$ возможных ориентаций спина атома или иона в магнитном поле, которым соответствуют значения магнитного квантового числа $m_s=S$, $(S-1)$, ..., $(1-S)$, $-S$. Теперь в магнитном поле каждой из величин m_s соответствует свое значение энергии, а двухуровневая картина рис. 5.23, *б* должна быть обобщена на случай $(2S+1)$ уровней (например, на рис. 5.28, *а* это сделано для $S=5/2$). Обычно $(2S+1)$ уровней для данного значения B разделены неодинаковыми энергетическими зазорами, причем, как показано на рис. 5.28, для $B=0$ имеет место вырождение, которое частично снимается вследствие асимметрии поля в кристалле.

Для переходов между этими уровнями, которые стимулированы фотонами, справедливо обычное правило отбора[33] $\Delta m_s=$

[33] Это правило отбора является следствием того, что момент количества движения фотона равен $\hbar$. Переходы с $\Delta m_s=\pm 2$ иногда наблюдаются, например для триплетного состояния нафталина [*Van der Waals J. H., de Groot M. S.*— Molecular Phys. 2, 333 (1959)].

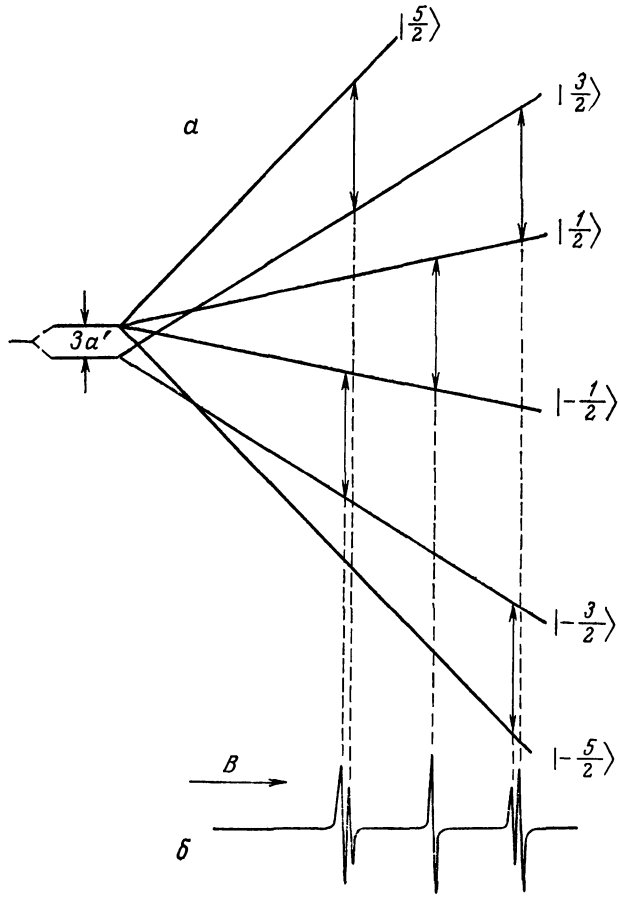


Рис. 5.28. *а* — изменение положения энергетических уровней при изменении поля **B** для иона с электронной конфигурацией $3d^5$ в кристаллическом поле октаэдрической симметрии; поле **B** параллельно главной оси октаэдра. *б* — разрешаемый спектр ЭПР для случая, когда $h\nu \gg 3a'$ [*Wertz J. E., Bolton J. R.* Electron Spin Resonance, McGraw-Hill, 1972].

$= \pm 1$. При данной микроволновой частоте это правило соответствует нанесенным на рис. 5.28, *а* вертикальным стрелкам для пяти различных значений B. Возникающий спектр ЭПР, который на рис. 5.28, *б*, представлен в виде производной линии поглощения, имеет *тонкую структуру*, состоящую из пяти линий, а в общем случае он должен иметь $2S$ линий. Если ширины этих линий превышают расстояния между отдельными линиями, должна наблюдаться одна широкая линия.

Ферромагнитный и ферримагнитный резонанс

Твердое тело, для которого обменный интеграл J_e [см. выражение (5.95)] положительный, в ферромагнитном состоянии при температурах ниже точки Кюри $T_c \sim (J_e/k_0)$ обладает объемной спонтанной намагниченностью. При температуре, значительно более низкой, чем T_c, возбуждается сравнительно малое количество спиновых волн и в сущности все неспаренные спины в образце оказываются связанными воедино. Эту ситуацию можно рассматривать как предельный случай системы связанных спинов (что для малого количества спинов одного атома уже обсуждалось в предыдущем разделе). Магнитный момент ферромагнитного образца является суммой магнитных моментов всех его магнитных атомов, и его составляющая, параллельная некоторому внешнему магнитному полю H_0, может быть представлена очень большим квантовым числом m_s.

Полный магнитный момент ферромагнитного образца прецессирует вокруг направления $\mathbf{H}_0$ с угловой частотой $\omega = 2\pi\nu$, зависящей от $\mathbf{H}_0$. *Ферромагнитный резонанс*[34] возникает, когда имеет место сильное поглощение микроволновых квантов, частота которых совпадает с частотой прецессии. Поглощение фотона изменяет m_s в соответствии с обычным правилом отбора $\Delta m_s = \pm 1$.

В предыдущем разделе использовалась величина $\mathbf{H}_0$, поскольку $\mathbf{H}_0$ или $\mathbf{B}_0 = \mu_0 \mathbf{H}_0$ могут быть измерены вне образца. Однако внутри образца существует иное эффективное магнитное поле. Вместе с тем, нет оснований считать, что внутреннее эффективное поле включает в себя очень сильное «молекулярное поле» Вейсса $\mathbf{H}_e$ [см. выражение (5.87)], в противном случае частоты ферромагнитного резонанса были бы порядка 10^{13} Гц даже в отсутствие внешнего поля. Поскольку этого не наблюдается экспериментально, $\mathbf{H}_e$ не должно быть эффективным. Причина заключается в том, что используемые радиочастотные кванты обладают длиной волны, которая велика по сравнению с размерами образца и взаимодействует в фазе со всеми спинами сразу. В этом случае отсутствует какой-либо угол между любым спином и его соседями и $\mathbf{H}_e$ не оказывает влияние на вращающий момент.

Какое же эффективное поле следует использовать в условии резонанса? Только для сферического образца соотношение $h\nu = g\mu_B \mathbf{B}_0$ дает такие значения g-фактора, которые совпадают с g-фактором для свободного электрона. Как показал Киттель [35], расхождение между приложенным и эффективным по-

[34] Первое сообщение о ферромагнитном резонансе принадлежит Гриффитсу [*Griffiths J. H. E.*— Nature, 158, 670 (1946)].

[35] *Kittel C.*— Phys. Rev., 71, 270 (1947).

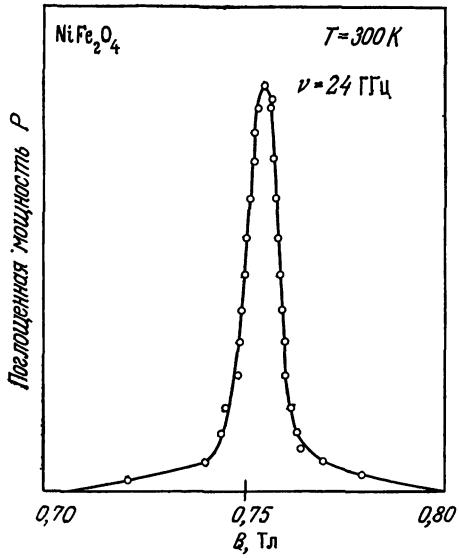


Рис. 5.29. Линия ферримагнитного резонанса для никелевого феррита. Измерения выполнены на сферическом образце диаметром 0,38 мм при комнатной температуре и частоте 24 ГГц. *Эффективное поле* в образце, скорректированное на малый размагничивающий фактор образца, форма которого несколько отлична от сферической, $B_e=0{,}787$ Тл для $g=2{,}18$ [*Yager W. A. et al.*— Phys. Rev., 80, 744 (1950)].

лями для подавляющего большинства образцов различной геометрической формы связано с сильным влиянием размагничивания в ферромагнетиках. В общем случае

$$h\nu = g\mu_B [(B_0 + \alpha\mu_0 M)(B_0 + \beta\mu_0 M)]^{1/2}, \qquad (5.118)$$

где величины α и β зависят от ориентации поля и формы образца; $\alpha=\beta=0$ для образца простой сферической формы.

Ферромагнитный резонанс имеет тенденцию давать значения g-фактора, несколько превышающие g-фактор для свободного электрона, если для обработки экспериментальных результатов использовать выражение (5.118) с величинами α и β, соответствующими образцу и геометрии эксперимента. Таким образом были получены значения $g=2{,}10$; 2,18 и 2,21 соответственно для железа, кобальта и никеля — ферромагнетиков группы $3d$-переходных элементов. Резонансное поглощение наблюдалось и для *ферримагнитных* соединений, при этом для железо-иттриевого граната ($Y_3Fe_5O_{12}$) было найдено $g=2{,}11$, а для никелевого феррита ($NiFe_2O_4$) $g=2{,}18$.

Ширина линии ферромагнитного или ферримагнитного резонанса зависит сложным образом от геометрии образца, темпе-

ратуры и особенностей материала. На рис. 5.29 представлена сравнительно широкая резонансная линия для ферримагнетика $NiFe_2O_4$, однако имеются сообщения, что для некоторых твердых тел наблюдаются узкие линии с шириной порядка 10^{-4} Тл. Теоретические представления о ширине линии в ферромагнетиках разработаны еще не полностью[36].

Высокая электропроводность металлических ферромагнетиков — чистых элементов и сплавов — создает экспериментальные сложности, которые связаны с очень малой глубиной проникновения (толщиной скин-слоя) микроволнового излучения. Эти сложности не возникают в случае непроводящих ферромагнитных и ферримагнитных соединений. В качестве примера на рис. 5.29 показана кривая поглощения ферримагнитного резонанса.

Антиферромагнитный резонанс

Рассмотрим твердое тело, которое является парамагнетиком при высоких температурах, а при температурах ниже точки Нееля T_N его спины испытывают антиферромагнитное упорядочение. В ранних исследованиях ЭПР на таких твердых телах при температурах выше T_N наблюдали резонансное поглощение, однако при температурах ниже T_N резонансное поглощение обнаружить не удавалось. Эта загадка была разрешена, когда Кеффер и Киттель[37] показали, что для *антиферромагнитного резонанса* требуются кванты, обладающие значительно большей энергией (ν в интервале от 10^{11} до 10^{12} Гц).

Кеффер и Киттель приняли во внимание следующие два фактора:

а) В антиферромагнетике вследствие магнитной анизотропии запасена большая энергия. Она может быть представлена параметром B_A.

б) Каждая магнитная подрешетка испытывает на себе действие молекулярного поля (обменного поля) со стороны другой подрешетки. При этом, в отличие от ферромагнитного материала, обменное поле обеспечивает ненулевой (на самом деле подавляющий) вклад B_E в эффективное поле.

Как следствие всего сказанного, условие антиферромагнитного резонанса принимает вид

$$h\nu = g\mu_B \{[B_A(2B_E + B_A)]^{1/2} \pm (1-\alpha) B_{\text{ext}}\}, \qquad (5.119)$$

где B_{ext} — внешнее магнитное поле. Величина α связана с размагничивающим фактором, соответствующим внешнему полю;

[36] См., например, *Prange R. E., Korenman V.*— J. Magn. Resonance, 6, 274 (1972).

[37] *Keffer F., Kittel C.*— Phys. Rev., 85, 329 (1952).

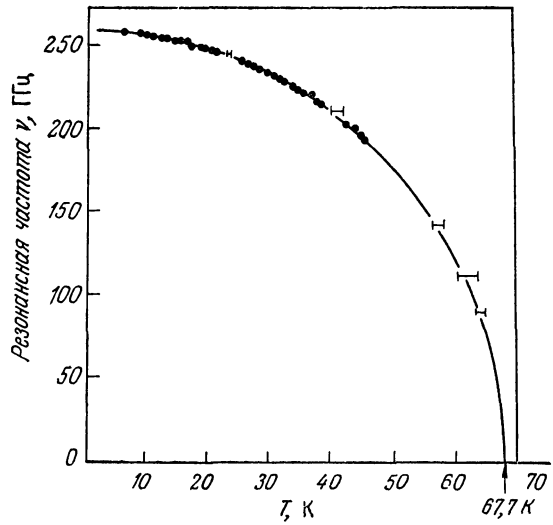


Рис. 5.30. Температурная зависимость частоты антиферромагнитного резонанса для MnF_2 в отсутствие внешнего магнитного поля. Экспериментальные точки, в действительности полученные с использованием некоторого внешнего магнитного поля, затем были пересчитаны с помощью уравнения (5.119) в частоты, соответствующие нулевому полю. [*Johnson F. M., Nethercot A H.*— Phys. Rev., **114**, 705 (1959).]

она уменьшается до нуля, когда температура поднимается до T_N.

На рис. 5.30 показано, как для кристаллов MnF_2 резонансная частота, соответствующая нулевому внешнему магнитному полю, изменяется с ростом температуры. Частота для самой низкой температуры согласуется с $B_A=0{,}88$ Тл и $B_E=56$ Тл, так что величина $[B_A(2B_E+B_A)]^{1/2}$ составляет почти 10 Тл. Это означает, что для изменения резонансной частоты внешним магнитным полем необходим очень мощный магнит. В качестве таких магнитов могут быть использованы сверхпроводящие соленоиды. Тогда сильное магнитное поле B_{ext}, приложенное в соответствующем направлении, может обеспечить уменьшение резонансной частоты до частот обычного микроволнового диапазона [38] [при этом используются преимущества, появляющиеся при выборе отрицательного знака в уравнении (5.119)]. Однако такие условия отсутствовали в раннем эксперименте, результаты которого приведены на рис. 5.30. Эта пионерская работа по антиферромагнитному резонансу была выполнена в миллиметровом диапазоне длин волн. Прохождение линии резонанса для каждой температуры осуществлялось

[38] *de Wijn H. W. et al.*— Phys. Rev., 8, 299 (1973).

с помощью развертки небольшим магнитным полем (до 0,5 Тл). Полученные экспериментальные данные затем были с помощью формулы (5.119) пересчитаны в эквивалентные резонансные частоты при нулевом внешнем поле, которые содержат информацию об изменении с температурой B_E и B_A. Кривая рис. 5.30 связана с (но не идентична) кривой функции Бриллюэна для намагниченности каждой из противоположно направленных магнитных подрешеток. И намагниченность и частота антиферромагнитного резонанса в нулевом поле стремятся к нулю по мере приближения снизу к $T_N = 67{,}7$ К.

ЭПР в электронно-ядерной системе

Вопросы, обсуждаемые в настоящем разделе, вовсе не связаны с корреляциями спинов в большой системе атомов. Мы увидим, что сложный характер резонансных спектров может быть связан с взаимодействием между спинами электронов атома и спином его ядра.

Угловой момент ядра запишем в виде $L_s = [\hbar^2 I(I+1)]^{1/2}$, где I — спиновое квантовое число. Различные изотопы имеют разные значения I, и некоторые из них приведены в табл. 5.5 в качестве примера. Нулевым спином ($I=0$) обладают ядра, содержащие как четное число протонов, так и четное число нейтронов, среди них такие широко распространенные ядра, как ^{12}C, ^{16}O, ^{24}Mg, ^{28}Si, ^{32}S.

Любое ядро с отличным от нуля спином обладает также магнитным моментом μ; значения магнитных моментов для некоторых атомных ядер приведены в табл. 5.5. Обычно ядерные магнитные моменты выражаются как кратное ядерного магнетона μ_n [см. соотношение (5.105)]; во всех случаях это кратное не является целым числом. Ядерный g-фактор $g_N = (\mu/I\mu_n)$ используется, когда проекцию ядерного магнитного момента на направление магнитного поля записывают в виде $g_N \mu_n m_I$. При этом магнитное квантовое число m_I может принимать любое из $(2I+1)$ значений: I, $(I-1), \ldots, (1-I), -I$. Существование ядерных магнитных моментов приводит к нескольким следствиям, которые мы должны принять во внимание.

Первое и наиболее очевидное из них состоит в том, что ядерный спин имеет $(2I+1)$ возможных значений угла наклона по отношению к направлению приложенного поля. Ядро, обладающее энергией u_0 в отсутствие поля, будет в магнитном поле характеризоваться одним из значений энергии $u = u_0 + \Delta u = u_0 + g_N \mu_n m_I B$. Магнитное квантовое число может меняться на единицу ($\Delta m_I = \pm 1$), что соответствует изменению момента количества движения на $\hbar$, если поглощается квант с энергией

Таблица 5.5. **Характеристики атомных ядер некоторых изотопов***

Изотоп	Спиновое квантовое число I	Ядерный магнитный момент μ (в единицах μ_n)	Ядерный g-фактор $g_N = (\mu/I\mu_n)$	Электрический квадрупольный момент Q (в единицах 10^{-28} м^2)
$^1n^0$	1/2	—1,913	—3,826	0
^1H	1/2	2,793	5,586	0
^2H	1	0,857	0,857	0,003
^7Li	3/2	3,256	2,171	—0,042
^9Be	3/2	—1,177	—0,785	0,049
^{14}N	1	0,404	0,404	0,071
^{19}F	1/2	2,627	5,254	0
^{23}Na	3/2	2,216	1,473	0,1
^{25}Mg	5/2	—0,855	—0,342	0,22
^{27}Al	5/2	3,639	1,456	0,149
^{31}P	1/2	1,131	2,262	0
^{35}Cl	3/2	0,821	0,547	—0,080
^{45}Sc	7/2	4,749	1,357	0,22
^{50}V	6	3,341	0,557	—
^{55}Mn	5/2	3,461	1,384	0,3

* Существенно более полная таблица дана в монографии: *Andrew E. R.* Nuclear Magnetic Resonance, Cambridge, 1958. Для любого ядра с четным числом протонов и четным числом нейтронов $I = 0$ и $\mu = 0$.

$\hbar\omega = h\nu = g_N\mu_n m_I B$. Все сказанное напоминает введение к изложению материала по ЭПР, но это и естественно, поскольку ниже мы будем рассматривать ЯМР. Здесь лишь отметим, что поскольку $g_N\mu_n$ в 10^3—10^4 раз меньше аналогичного произведения для электронов $g\mu_B$, частота квантов для ЯМР много меньше частоты квантов, используемых в экспериментах по ЭПР.

Существование ядерного магнитного момента оказывает влияние на величину энергии неспаренных электронов того же атома посредством двух механизмов, которые известны как *фермиевское контактное взаимодействие* и *дипольное взаимодействие*. Мы кратко рассмотрим эти механизмы. Их влияние удобно описывать некоторыми эффективными полями ΔB_{fc} и ΔB_{di}, которые накладываются на внешнее магнитное поле. Результирующее суммарное поле определяет поведение любого неспаренного электрона в атоме. Величины ΔB_{fc} и ΔB_{di} зависят от m_I и могут принимать максимальные значения даже при отсутствии внешнего поля. Таким образом, энергия системы неспаренных электронов, принадлежащих атомам, может в нулевом внешнем поле иметь $(2I+1)$ значений. Включение внешнего магнитного поля полностью снимает вырождение и появляется $(2I+1)(2S+1)$ состояний с различными значениями энергии. В экспериментах по ЭПР резонансное поглощение

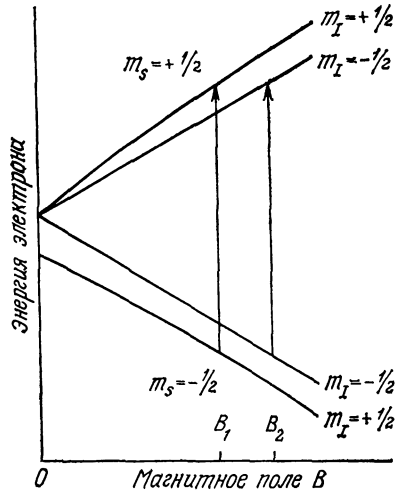


Рис. 5.31. Зависимость энергии электрона от магнитного поля для системы, состоящей из электронного спина $S=1/2$ и ядерного спина $I=1/2$. Стрелками обозначены два перехода с участием квантов одной и той же энергии, удовлетворяющие правилу отбора $\Delta m_s=\pm 1$, $\Delta m_I=0$. Эта задача впервые была решена Брейтом и Раби [*Breit G., Rabi I.*— Phys Rev., 38, 2082 (1931)] применительно к атомной спектроскопии, и возникающие при этом уровни иногда называют уровнями Брейта — Раби. Следует отметить, что в слабых полях наблюдается заметная кривизна представленных зависимостей, которая обусловлена смешиванием различных m_s-, m_I-состояний.

квантов с фиксированным значением энергии может иметь место для $2S(2I+1)$ различных значений напряженности поля при условии выполнения следующих правил: $\Delta m_s=\pm 1$, $\Delta m_I=0$. Второе из этих правил отбора является необходимым поскольку, несмотря на то что поворот ядерного спина одновременно с поворотом электронного спина требует очень небольшой дополнительной энергии, такой процесс не согласуется с величиной момента количества движения $\hbar$ поглощаемого фотона (существует исключение для переходов $\Delta m_s=+1$, $\Delta m_I=-2$, которые иногда называют «запрещенными» переходами).

На рис. 5.31 показана зависимость энергии электрона от магнитного поля для простейшей возможной комбинации электронного и ядерного спинов: $S=1/2$ и $I=1/2$. В книгах, приведенных в библиографии в конце настоящей главы, решены задачи для более сложных комбинаций спинов. Как следует из нашего упрощенного рассмотрения, даже в случае, представленном на рис. 5.31, вместо одной линии ЭПР должны появиться две линии, соответствующие разрешенным переходам для квантов данной энергии. Разность (B_2-B_1) магнитных полей

для этих двух резонансных линий принято называть *сверхтонким расщеплением.* Таким названием подчеркивается, что сверхтонкое расщепление всегда должно быть очень малым, меньше, чем расщепление, упоминавшееся при обсуждении рис. 5.28 как «тонкая структура». Однако сверхтонкое расщепление, будучи мало для некоторых атомов (10^{-3} Тл или меньше), для других оказывается весьма значительным. Ядро атома водорода обеспечивает наибольшее сверхтонкое расщепление, равное 0,0508 Тл.

Спиновый гамильтониан

В большей части настоящей книги мы старались обойтись без применения формализма квантовой механики. Однако без упоминания о спиновом гамильтониане $\mathscr{H}$ трудно просуммировать взаимодействия электронных и ядерных спинов. Волновая функция ψ электрона (или коллектива неспаренных электронов атома) должна удовлетворять уравнению Шредингера с соответствующим оператором Гамильтона, а собственные значения этого уравнения являются разрешенными значениями энергии. Интересуясь лишь спиновыми взаимодействиями, мы можем опустить большие члены гамильтониана, представляющие орбитальное движение и спин-орбитальное расщепление, и оставить только члены, соответствующие электронному спину, спину ядра и их взаимодействию. Таким образом,

$$\mathscr{H} = \mathscr{H}_1 + \mathscr{H}_2 + \mathscr{H}_3 + \mathscr{H}_4 = g\mu_B \mathbf{B}\cdot\mathbf{S} + g_N\mu_n \mathbf{B}\cdot\mathbf{I} + \mathscr{A}\mathbf{I}\cdot\mathbf{S} + \\ + \mathscr{B}[\mathbf{I}\cdot\mathbf{S} - (3/r^2)(\mathbf{I}\cdot\mathbf{r})(\mathbf{S}\cdot\mathbf{r})]. \qquad (5.120)$$

Первый член в правой части уравнения (5.120) описывает взаимодействие спина электрона с внешним полем B, которое приводит к зеемановскому расщеплению уровней, что проиллюстрировано на рис. 5.23 и 5.28. Когда $\mathscr{H}_1$ является *единственным* членом спинового гамильтониана, разрешенные значения энергии представляются набором $g\mu_B B m_s$. Таким образом, результат действия члена $\mathscr{H}_1$ нам известен.

Член $\mathscr{H}_2$ описывает зеемановское расщепление спиновых состояний ядра, и ему соответствует набор энергетических состояний $g_N\mu_n B m_I$, между которыми происходят переходы ЯМР. Сейчас нет необходимости уделять внимание этому слагаемому, и мы вернемся к нему ниже. Пренебрежение им не помешает нам обсудить, каким образом электронный и ядерный спины взаимодействуют друг с другом.

Первым членом, соответствующим электронно-ядерному взаимодействию, является оператор $\mathscr{H}_3 = \mathscr{A}\mathbf{I}\cdot\mathbf{S}$. Он описывает влияние ядерного магнетизма на электрон, который движется с ре-

лятивистской скоростью в непосредственной близости от ядра. Ферми [39] показал, что

$$\mathscr{A} = \left(\frac{2}{3}\right)\mu_0 g_B g_N \mu_n |\psi(0)|^2 \text{ Дж} \tag{5.121}$$

и зависит от вероятности $|\psi(0)|^2$ нахождения электрона в месте расположения ядра. Этот член $\mathscr{A}\mathbf{I}\cdot\mathbf{S}$ описывает *фермиевское контактное* взаимодействие; оно отсутствует, если орбитальная волновая функция имеет узел на ядре. Следовательно, фермиевское контактное взаимодействие оказывается существенным только для *s*-состояний или для гибридизованных состояний с некоторой долей *s*-подобного состояния. В случае чистых *p*-, *d*- или *f*-состояний это взаимодействие можно не учитывать. Также просто можно показать, что член $\mathscr{H}_4$ равен нулю для любого чистого *s*-состояния.

Если рассматривать спиновый гамильтониан, который содержит только члены $\mathscr{H}_1$ и $\mathscr{H}_3$, то в *первом приближении* энергия электрона описывается следующим выражением:

$$U = U_0 + g\mu_B B m_s + \mathscr{A} m_s m_I. \tag{5.122}$$

Его можно переписать в виде

$$U = U_0 + g\mu_B m_s (B + \Delta B_{fc}), \tag{5.123}$$

где ΔB_{fc} — магнитное поле, эквивалентное фермиевскому контактному взаимодействию. Из выражений (5.121)—(5.123) можно получить, что

$$\Delta B_{fc} = (\mathscr{A}/g\mu_B)\, m_I = \left(\frac{2}{3}\right)\mu_0 g_N \mu_n |\psi(0)|^2 m_I. \tag{5.124}$$

Изменение m_I на единицу должно привести к тому, что необходимое для ЭПР внешнее магнитное поле изменится на $(\mathscr{A}/g\mu_B)$. Эта величина называется *сверхтонким расщеплением* и на рис. 5.31 показана как $(B_2—B_1)$. Сверхтонкое расщепление можно легко рассчитать, если известны g_N и $\psi(0)$ (см. задачу 5.16)

На рис. 5.32 и 5.33 показано сверхтонкое расщепление для локализованных электронов в одном полупроводнике и локализованных дырок в другом полупроводнике. Зависимости рис. 5.32 получены для кремния ($I=0$), содержащего донорную примесь замещения ^{31}P — атомы с ядерным спином $I=1/2$. Доноры пятой группы в кремнии отдают четыре электрона внешней оболочки на образование тетраэдрических связей в решетке, а пятый электрон внешней оболочки, если он связан со своим

[39] *Fermi E.*— Z. Physik, 60, 320 (1930).

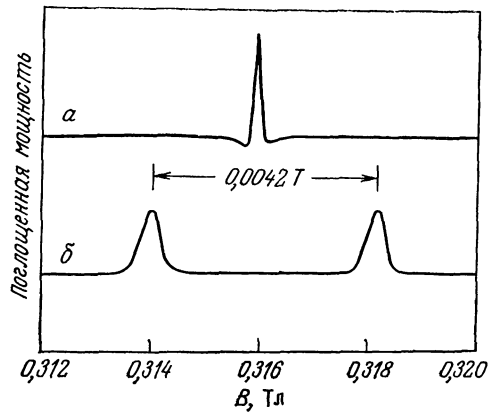


Рис. 5.32. Спектр ЭПР двух образцов кремния, легированных мелкими донорами — фосфором [*Feher G.*— Phys. Rev., **114**, 1219 (1959)]. Спектры записаны при температуре 1,25 К на частоте 8,845 ГГц. В образце *а* содержится $3 \cdot 10^{24}$ м$^{-3}$ атомов фосфора. Энергия активации при таком легировании спадает до нуля подобно тому, как показано на рис. 4.33. Образец *б* является менее легированным, для него $N_d = 7 \cdot 10^{21}$ м$^{-3}$ и электрон каждого донорного атома движется по орбите 1*s*-состояния вокруг своего ядра ^{31}P. Поэтому для образца *б* наблюдается сверхтонкое расщепление, обусловленное фермиевским контактным взаимодействием.

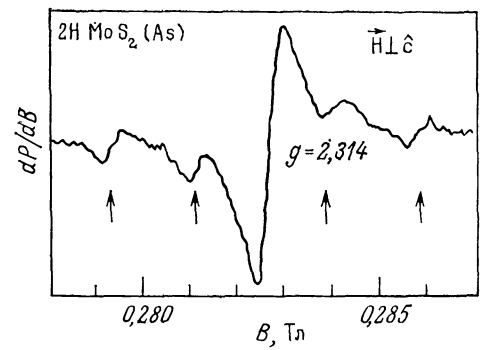


Рис. 5.33. Сверхтонкая структура спектра ЭПР естественного кристалла MoS_2, содержащего акцепторные атомы ^{75}As [*Title R. S., Shafer M. W.*— Phys. Rev., **B8**, 615 (1973)]. Измерения выполнены при 77 К на частоте 9 ГГц. Центральная линия представляет поглощение нелокализованными дырками в отдельных участках кристалла, для которых N_a велико. Четыре линии, расщепленные дипольным взаимодействием (отмечены стрелками), порождаются дырками, локализованными на акцепторах ^{75}As. Связанные дырки описываются *d*-подобными волновыми функциями, на что указывает анизотропия сверхтонкой структуры.

собственным ядром, движется по s-подобной орбите. На линии ЭПР для слабо легированного образца *б* наблюдается сверхтонкий дублет, причем величина расщепления позволяет вычислить $\psi(0)$. Заметим, что сверхтонкие линии отсутствуют в спектре для сильно легированного образца *а*, в котором электроны уже не локализованы на своих ядрах. Для этого образца наблюдается единственная линия, суженная обменным взаимодействием.

Фермиевское контактное взаимодействие для s-состояний является *изотропным* и вращение образца, в котором имеет место такое взаимодействие, не приводит к изменению сверхтонкой структуры. Однако в случае дипольного взаимодействия ядра с электронными p-, d- или f-состояниями сверхтонкая структура изменяется с ориентацией образца. Величина $\mathscr{B}$ в гамильтониане (5.120), которая описывает взаимодействие ядерных и электронных диполей, имеет вид

$$\mathscr{B} = -\mu_0 g \mu_B g_N \mu_n / r^3. \tag{5.125}$$

Качественно этот результат можно представить как взаимодействие магнитного момента электрона с магнитным полем

$$\Delta B_{di} = \mu_0 g_N \mu_n / \langle r^3 \rangle, \tag{5.126}$$

где $\langle r^3 \rangle$ — средний куб расстояния между ядром и электроном.

На рис. 5.33 приведен пример спектра ЭПР со сверхтонким расщеплением, которое обусловлено дипольным взаимодействием вида (5.126). Этот спектр получен для дырок в естественном кристалле MoS_2 из Квинсленда, который содержит акцепторы ^{75}As. Поскольку $I=3/2$ для ядер ^{75}As, сверхтонкая структура состоит из четырех линий. (Этот сигнал накладывается на одну центральную линию «свободных дырок», которая возникает под влиянием нелокализованных дырок в отдельных областях кристалла). Тайтл и Шейфер [40] установили, что изменение сверхтонкого расщепления с ориентацией согласуется с тем, что дырки находятся на d_{z^2}-орбиталях и величина $\langle r^3 \rangle$ может быть оценена, исходя из ΔB_{di}. Можно привести много примеров, когда электроны, которые дают спектр ЭПР, занимают гибридизованные состояния, и сверхтонкая структура оказывается сложной, поскольку в данном случае в гамильтониане присутствуют как $\mathscr{A}$-, так и $\mathscr{B}$-члены, вносящие вклад в энергетический спектр. Эти сложности оказываются привлекательными для экспериментаторов, так как из угловой

[40] *Title R. S., Shafer M. W.*— Phys. Rev., **B8**, 615 (1973).

зависимости сверхтонкого расщепления можно получить информацию о месте расположения атома с неспаренными спинами в кристаллической решетке.

Расщепление в кристаллическом поле

О *тонкой структуре* линии ЭПР, возникающей вследствие анизотропии кристаллического поля, очень кратко упоминалось при обсуждении рис. 5.28, однако этот вопрос заслуживает гораздо большего внимания. Основной объем текущей экспериментальной работы по ЭПР обусловлен эффектами кристаллического поля, так как анизотропия тонкой структуры в спектре ЭПР может быть сопоставлена с положением в кристаллической решетке и взаимным расположением электронов, связанных с атомами (или с вакансиями, в которых должны были бы находиться атомы).

Рассмотрим одно из самых простых возможных проявлений расщепления линии кристаллическим полем. Предположим, что некоторое твердое тело, которое кристаллизуется в кубической решетке, содержит парамагнитные примесные атомы, и окружение этих атомов обладает более низкой симметрией, чем кубическая. Для простоты будем считать, что окружение каждого примесного атома имеет тетрагональную симметрию, и ось симметрии совпадает с одним из направлений типа [100]. Если осью симметрии для данной примеси является ось z, то связанная с ней система электронных спинов имеет различные величины $g_{\parallel}$ и $g_{\perp}$ для направлений, которые соответственно параллельны и перпендикулярны этой оси.

Упрощенный спиновый гамильтониан можно записать в виде

$$\mathcal{H} = g_{\parallel}\mu_B B_z S_z + g_{\perp}\mu_B (B_x S_x + B_y S_y) + D S_z^2, \qquad (5.127)$$

где D — расщепление в отсутствие поля. Для центра с двумя неспаренными электронами $S=1$. Значения энергии U, вытекающие из уравнения (5.127) для $S=1$, могут быть получены как корни кубического уравнения

$$U(U-D)^2 - U(g_{\parallel}^2\mu_B^2 B^2\cos^2\theta) - (U-D)(g_{\perp}^2\mu_B^2 B^2\sin^2\theta) = 0, \qquad (5.128)$$

где θ — угол между полем $\mathbf{B}$ и осью z.

Если теперь вектор $\mathbf{B}$ направить вдоль кристаллографического направления [100], для одной трети всех примесных центров в кристалле ось z будет параллельна полю, а для оставшихся двух третей ось z будет перпендикулярна ему. Наблюдаемый в данном случае спектр ЭПР является суперпозицией «параллельного» и «перпендикулярного» спектров, причем последний имеет в два раза большую амплитуду, чем первый.

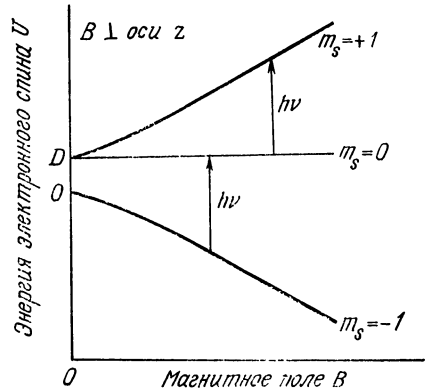


Рис. 5.34. Решения уравнения (5.128) для трех энергетических состояний в случае, когда вектор **B** перпендикулярен тетрагональной оси (оси z) кристаллического поля, оказывающего влияние на систему электронных спинов. Стрелками отмечены два значения поля, при которых имеют место разрешенные переходы для квантов с энергией $h\nu > D$. В задаче 5.17 предлагается получить это решение, а также решение для случая, когда вектор **B** параллелен оси c, а в задаче 5.18 предлагается сопоставить информацию, получаемую из спектров ЭПР для случаев, когда $h\nu$ соответственно больше или меньше D.

(Если вектор **B** не параллелен направлению [100], спектр ЭПР является более сложной суммой решений для θ и $(\pi/2 - \theta)$.)

На рис. 5.34 схематически представлены решения уравнения (5.128) для случая, когда вектор **B** перпендикулярен оси z. Из этого рисунка видно, что «перпендикулярный» компонент спектра ЭПР для $h\nu > D$ состоит из двух линий. Если измерения проводятся на микроволновой частоте много большей, чем (D/h), эти линии достаточно далеко отстоят друг от друга по магнитному полю. В задаче 5.17 предлагается получить решения для «перпендикулярного» и «параллельного» спектров. Из этой задачи следует, что решение для «параллельного» случая дает еще две линии ЭПР, когда $h\nu > D$. Величины D, $g_\parallel$ и $g_\perp$ могут быть определены по положениям в магнитном поле различных линий для радиочастотных квантов данной энергии; это можно сделать как для $h\nu > D$, так и для $h\nu < D$ (см. задачу 5.18).

Парамагнитные примеси в кристаллах часто создают ситуацию, более сложную, чем описываемая уравнением (5.128). Соответствующее эффективное поле может обладать более низкой симметрией, чем тетрагональная, спектр может быть более сложным, если $S > 1$, и, конечно, если ядерный спин не равен нулю, появляется также сверхтонкая структура. На рис. 5.35

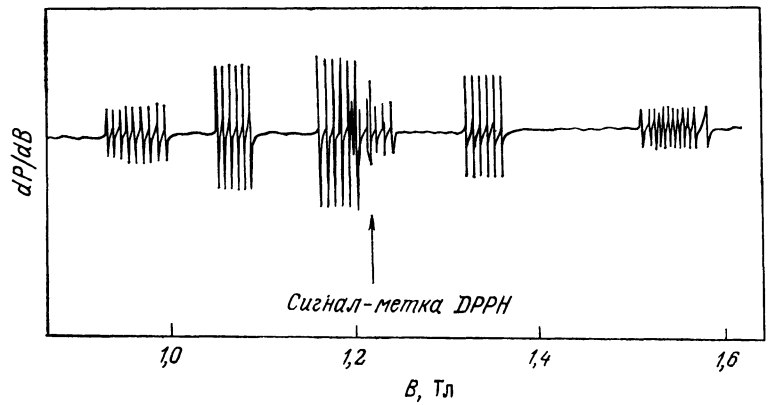


Рис. 5.35. Семь из десяти основных линий спектра ЭПР ионов Mn^{++} в $NH_4Cl \cdot 2H_2O$, измеренных в работе Сида на частоте 35 ГГц [*Seed T. J.*— J. Chem. Phys., **41**, 1486 (1964)]. Каждая линия расщеплена на шесть сверхтонких компонент ядерным спином ^{55}Mn. Одна дополнительная линия со стороны низких полей при 0,56 Тл на рисунке опущена, а две линии со стороны самых высоких полей Сид не мог наблюдать. Положения этих линий были предсказаны в результате анализа, выполненного в работе: *Forman A., van Wyck J. A.*— Canad. J. Phys., **45**, 3381 (1967). Эти линии были экспериментально зарегистрированы в работе: *Kennewell J. A., Pilbrow J. R., Price J. H.*— Phys. Lett., **27A**, 228 (1968) в магнитных полях до 1,85 Тл. Заметим, что интенсивности двух линий, наблюдаемых в полях около 1,2 Тл, приближенно относятся как 2 : 1, как и следует ожидать в случае «перпендикулярного» и «параллельного» спектров.

показана бо́льшая часть спектра ЭПР ионов Mn^{++}, присутствующих в качестве примеси в гидратированном хлориде аммония $NH_4Cl \cdot 2H_2O$. Под действием тетрагонального внутрикристаллического поля для спиновой системы с $S=5/2$ в спектре ЭПР появляются десять линий (пять линий соответствуют ситуации, когда вектор **B** параллелен оси z, а пять других — ситуации, когда вектор **B** перпендикулярен этой оси). Каждая из этих десяти линий расщепляется далее на $(2I+1)=6$ линий сверхтонкой структуры, поскольку для ядер ^{55}Mn спин $S=5/2$. Детальный анализ этого спектра очень сложен, но он позволяет определить g-факторы для различных направлений, компоненты кристаллического поля и коэффициенты для членов сверхтонкого расщепления. Вся эта информация может быть использована для того, чтобы определить положения активных спинов в твердом теле.

Ядерный магнитный резонанс

Для анализа сверхтонкой структуры спектров ЭПР необходимо знать спиновое квантовое число атомного ядра I. Вместе с тем, как отмечалось, магнитный момент ядра

$$\mu = g_N \mu_n \sqrt{I(I+1)}. \tag{5.129}$$

можно выразить через I, ядерный магнетон μ_n [см. соотношения (5.105)] и ядерный g-фактор g_N. В табл. 5.5 даны для некоторых ядер значения I, g_N и максимальные значения проекции магнитного момента $g_N \mu_n I$ на направление магнитного поля.

Член $\mathscr{H}_2$ в гамильтониане (5.120), как уже отмечалось, описывает взаимодействие ядерного магнитного момента с магнитным полем **B**. В магнитном поле **B** энергия ядра принимает одно из значений

$$u = u_0 - g_N \mu_n B m_I, \tag{5.130}$$

где u_0 — энергия ядра в отсутствие магнитного поля, а m_I может иметь одно из значений I, $(I-1), \ldots, (1-I)$, $-I$. Эти значения соответствуют $(2I+1)$ возможным углам между направлениями поля и ядерного магнитного момента.

Переходы с одного из указанных выше уровней на другой могут быть индуцированы фотонами при условии, что $\Delta m_I = \pm 1$ (условие сохранения момента количества движения) и

$$\hbar\omega = h\nu = g_N \mu_n B. \tag{5.131}$$

Это и есть ядерный магнитный резонанс (ЯМР). Для наблюдения ЯМР необходимы существенно более низкие частоты, чем для наблюдения ЭПР, что связано с малыми величинами ядерных магнитных моментов. Если B выразить в теслах, а ν — в МГц,

$$\nu = 7{,}623 g_N B \qquad (\text{МГц}). \tag{5.132}$$

Поэтому для наблюдения ЯМР подходящими оказываются частоты в несколько МГц. Сигнал ЯМР от молекул воды часто используется для калибровки магнитных полей, которая основана на следующем соотношении для протонов: $B(\text{Тл}) = [\nu(\text{МГц})/42{,}58]$.

Наблюдения ЯМР, как и ЭПР, обычно выполняются на фиксированной и высокостабильной радиочастоте, а магнитное поле изменяется и проходит через резонанс. Поле B обычно является суммой постоянного поля большого электромагнита и малого переменного поля $\Delta B \sin(\omega_0 t)$, которое создается вспомогательными катушками. При этом суммарное маг-

нитное поле проходит через действительное резонансное значение (ω_0/π) раз в секунду. Использование фазового детектора позволяет регистрировать либо линию поглощения, либо ее производную (dP/dB).

При тепловом равновесии отношение заселенностей спиновых состояний ядра, для которых $\Delta m_I = 1$, определяется как

$$\exp(-g_N\mu_n B/k_0T) = \exp(-h\nu/k_0T) \simeq [1-(h\nu/k_0T)] \quad (5.133)$$

и оказывается близким к единице, если только измерения не выполняются при очень низкой температуре. Так, для комнатной температуры и $\nu = 60$ МГц величина $(h\nu/k_0T) \simeq 10^{-5}$. Малое различие в заселенностях для различных спиновых состояний означает, что насыщение в системе ядерных спинов может быть легко получено, если подается избыточная радиочастотная мощность.

Усовершенствование опыта Штерна и Герлаха позволило в 30-х годах нашего столетия выполнить первые измерения ядерных магнитных моментов, а в 1937 г. Лазарев и Шубников[8] на молекулярном водороде продемонстрировали наличие парамагнетизма у протонов. Точные измерения ядерных моментов стали возможными после того, как Раби[41] применил на молекулярных пучках технику резонансного поглощения, которая позднее была усовершенствована Рамзеем[42]. Это были первые измерения ЯМР, но ЯМР на твердых телах не удавалось измерить еще в течение ряда лет. Эксперименты на молекулярных пучках давали информацию о величинах I и g_N для различных ядер, которая необходима для развития ядерной физики. Специалистам в области физики твердого тела, занимающимся исследованиями ЯМР, также требуется знание величин I и g_N для ядер в их образцах, но они используют эту информацию лишь как средство для достижения цели.

В 1942 г. Гортером и Броером[43] были предприняты попытки обнаружить резонансное поглощение ядрами в магнитном поле в кристаллах (^{7}Li в LiCl и ^{19}F в KF). Они пытались найти слабую дисперсию диэлектрической постоянной при прохождении условий резонанса. Их попытки оказались безуспешными, а резонансное поглощение радиочастотной мощности в магнитном поле было продемонстрировано четыре года спустя Парселом с сотрудниками[44] для протонов в парафине и Блохом с сотрудниками[45] для протонов в воде. На рис. 5.36 представлена

[41] *Rabi I., Millman S., Kusch P., Zacharias J.*— Phys. Rev., **55**, 526 (1939).
[42] *Ramsay N. F.* Molecular Beams, Oxford, 1955.
[43] *Gorter C. J., Broer L. F.*— Physica, **9**, 591 (1942).
[44] *Purcell E. M., Torrey H. C., Pound R. V.*— Phys. Rev., **69**, 37 (1946).
[45] *Bloch F., Hansen W. W., Packard M. E.*— Phys. Rev., **69**, 127 (1946).

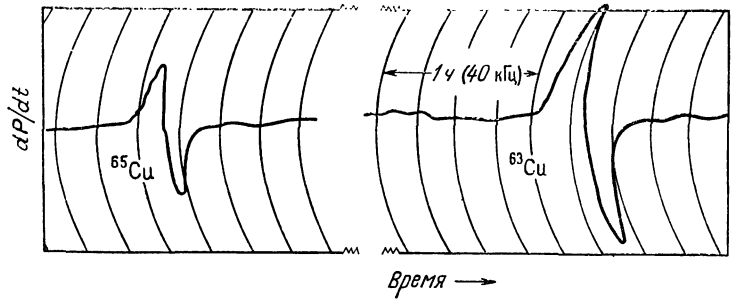


Рис. 5.36. Копия одной из ранних записей спектров ЯМР, полученных Паундом [*Pound R. V.*— Phys. Rev., 73, 523 (1948)] для ядер ^{63}Cu и ^{65}Cu в порошкообразном хлориде меди. Величина производной (dP/dt) контролировалась и записывалась с помощью самописца в процессе прохождения резонанса для двух изотопов меди. В этом эксперименте поле B поддерживалось постоянным (0,3 Тл), а частота линейно уменьшалась со временем. Отношение амплитуд двух линий находится в соответствии с отношением естественных распространенностей 7 : 3 для этих двух изотопов меди. Оба изотопа ^{63}Cu и ^{65}Cu имеют ядерный спин $I = {}^3/_2$, и результаты, представленные на данном рисунке, были первым точным измерением их ядерных магнитных моментов: $\mu = 2{,}22\mu_n$ для ^{63}Cu и $\mu = 2{,}38\mu_n$ для ^{65}Cu.

производная поглощенной мощности по времени для изотопов ^{63}Cu и ^{65}Cu в хлориде меди. Эти кривые были опубликованы Паундом в 1948 г. По оси абсцисс отложено время (цена деления — 15 мин). В процессе эксперимента частота ν изменялась линейно со временем (при постоянном поле B). В ранних исследованиях ЯМР запись кривой часто занимала несколько часов. Теперь существует техника, позволяющая быстрее получать экспериментальные данные, однако из-за больших времен релаксации часто при исследованиях ЯМР оказывается необходимым медленное прохождение измеряемого диапазона.

Помимо определения положения линии ЯМР (отношение ν к B) можно также определить ее форму и ширину, а в ряде случаев удается наблюдать тонкую структуру линии. Как будет показано в следующем кратком разделе, интерес к информации, получаемой из измерений ЯМР, проявляют специалисты в области химии, биофизики и физики твердого тела.

ЯМР-спектроскопия высокого разрешения

В экспериментах по ЯМР в конце 40-х годов нашего столетия основное внимание было сосредоточено на возможно более точном измерении положения линии. Это было очень важно, поскольку для многих ядер магнитные моменты были лишь предварительно оценены из оптической спектроскопии. Нако-

нец, с помощью ЯМР удалось найти точные значения g_N для этих ядер, получив информацию, которая представляла большой интерес для ядерной физики.

Эта стадия исследований давно завершилась. Величины g_N известны с достаточной точностью фактически для всех стабильных изотопов, а также для многих радиоактивных изотопов. Таким образом, специалисты в области физики твердого тела, которые применяют ЯМР для исследований электрического и магнитного полей в твердом теле, могут опираться на надежные значения g_N, имеющиеся в литературе.

Как впервые было отмечено Найтом[46], условие резонанса (ν/B) для данного ядра может зависеть от химического окружения. Основное внимание Найт уделил достаточно большому *сдвигу Найта* в резонансных условиях в металлах, что будет обсуждаться позднее, однако он также сообщил и о небольшом подобном эффекте в неметаллах. *Химический сдвиг* величины (ν/B) для ядра возникает из-за различия химических связей в разных неметаллических твердых телах и является результатом малого вклада электронов атомных и молекулярных орбиталей в локальное магнитное поле. Химический сдвиг обычно представляется как относительное смещение $(\Delta B/B)$ от предполагаемого «истинного» резонансного поля ядра при данной радиочастоте. Установлено, что для молекул, растворенных в жидких растворителях, величина $(\Delta B/B)$ иногда достигает 10^{-3}. Химический сдвиг меньшей величины иллюстрируется триплетом резонанса на рис. 5.37. Здесь приводится один из первых результатов исследований по химическим сдвигам — три пика ЯМР для протонов, соответствующие трем неэквивалентным по химической связи положениям атома водорода в этиловом спирте.

Химия и молекулярная биофизика извлекли немалую пользу из исследований с помощью ЯМР-спектроскопии высокого разрешения в таких направлениях, как структурная органическая химия[47] и положения водородных связей в биологических макромолекулах. Однако ЯМР-спектроскопия высокого разрешения возможна лишь тогда, когда резонансные линии оказываются узкими по сравнению с расстояниями между ними. Как это ни удивительно, но линии ЯМР оказываются узкими в газах и жидкостях, а в твердых телах они широкие. Причина состоит в том, что условия резонанса для ядер в жесткой решетке уширяются вследствие диполь-дипольного взаимодействия с сосе-

---

[46] *Knight W. D.*— Phys. Rev., **76**, 1259 (1949).

[47] Химическим аспектам ЯМР посвящена обширная литература, в том числе и монография: *Paudler W. W.* Nuclear Magnetic Resonance, Allyn and Bacon, 1971.

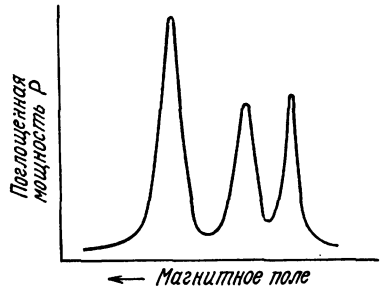


Рис. 5.37. Запись спектра ЯМР для протонов в этиловом спирте [*Arnold J. T., Dharmetti S. S., Packard M. E.*— J. Chem. Phys., **19**, 507 (1951). Эксперимент проводился при $B=0{,}76$ Тл на частоте $\nu=32{,}4$ МГц. Три линии принадлежат соответственно протонам CH_3-, CH_2- и OH-групп; расстояние между двумя внешними линиями спектра описывается отношением $\Delta B/B=5\cdot10^{-6}$.

дями. Напротив, быстрое движение и вращение молекул в жидкостях сглаживают локальные поля, создаваемые соседними диполями, что приводит к очень узким линиям!

Таким образом, вещества, которые представляют интерес для химии и биологии и существуют в жидком состоянии, являются хорошими объектами для ЯМР-спектроскопии высокого разрешения. К сожалению, в области физики твердого тела необходимо применять разные ухищрения, поскольку приходится иметь дело с широкими резонансными линиями.

Релаксация и ширина линии ЯМР в твердых телах

Постоянные времени τ_1 для спин-решеточной релаксации и τ_2 для спин-спиновой релаксации рассматривались выше в этом разделе в связи с ЭПР, и изложенные соображения применимы также и к ЯМР. Решения уравнений Блоха[31] для отклика системы спинов при линейном прохождении B/ν через условие резонанса зависят от τ_1 и τ_2.

Величина τ_1 обычно не влияет на ширину линии ЯМР, но важна для определения той скорости, с которой могут индуцироваться спиновые переходы в условиях, когда не достигается насыщение системы. Значение τ_1 может становиться очень большим, в частности при низких температурах. Обнаружив, что при комнатной температуре $\tau_1=300$ с для спиновой релаксации ядер ^7Li в монокристаллах LiF, Паунд и Парселл[48]

[48] *Purcell E. M., Pound R. V.*— Phys. Rev., **81**, 279 (1951).

смогли использовать этот очень медленный отклик для того, чтобы создать инверсную заселенность ядерных спинов, т. е. такую заселенность, при которой состояния с большей энергией оказываются более заполненными, чем состояния с меньшей энергией. Парселл и Паунд описали это явление на основе представления об *отрицательной температуре* для ядер, которая в данном случае оказалась равной —350 К. Это неравновесное состояние релаксирует через бесконечную спиновую температуру к нормальным условиям. Последующее развитие мазеров и лазеров привело к тому, что представления об инверсной заселенности и «отрицательных» температурах стали обычными.

Большая ширина линий ЯМР в твердых телах является результатом магнитных диполь-дипольных взаимодействий с ядерными магнитными моментами ближайших соседей. Соседний магнитный момент на расстоянии $\mathbf{r}$ вносит вклад в магнитное поле

$$\delta\mathbf{B} = (\mu_0/4\pi r^5)\,[3\,(\boldsymbol{\mu}\cdot\mathbf{r})\,\mathbf{r} - r^2\boldsymbol{\mu}], \qquad (5.134)$$

что можно проверить путем сравнения с выражением (5.6) для электрического диполя. Оценка с помощью уравнения (5.134) показывает, что момент величиной $g_N\mu_n$ на расстоянии a до ближайшего соседа создает поле $|\delta\mathbf{B}| \sim (\mu_0 g_N \mu_n/4\pi a^3)$, которое обычно порядка 10^{-3} Тл. Вклады от различных ближайших соседей складываются векторно, что приводит к изменению резонансных условий. В результате возникает ситуация, при которой эффективное поле увеличивается для одних ядер и уменьшается для других в данный момент времени. Таким образом, нетрудно оценить диполь-дипольное взаимодействие по наблюдаемым величинам ширины линии ЯМР, которые обычно лежат в пределах $\Delta B_{pp} = (5 \div 50) \cdot 10^{-4}$ Тл.

Динамическое сужение и диффузия в твердых телах

Существенно более узкие линии ЯМР в жидкостях обусловлены быстрым поступательным и вращательным движением молекул и атомов; движение приводит к тому, что усредненная сумма всех вкладов δB за один период колебания радиочастотного излучения оказывается равной нулю.

Динамическое сужение может также возникать и в твердых телах, если температура достаточно высока для миграции атомов из одного узла решетки (или междоузлия) в другой посредством диффузионного механизма. Явление *диффузионного сужения* впервые наблюдалось для твердого водорода[49]

[49] *Hatton J., Rollin B. V.*— Proc. Roy. Soc., **A199**, 222 (1949).

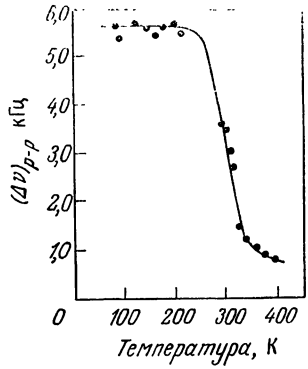


Рис. 5.38. Ширина линии ЯМР для ^{7}Li в стеклообразной композиции, включающей 20 % Li_2O, 80 % SiO_2. Данные по ширине линии получены преобразованием ΔB_{pp} в $\Delta\nu_{pp}$ и построены как функция температуры. Температурная зависимость для $T>260$ К согласуется с частотой перескоков для диффузии, которая изменяется как $\exp(-0{,}48/k_0T)$ (см. задачу 5.20) [*Bray P. J.* Magnetic Resonance, eds. C. K. Coogan et al., Plenum Press, 1970, p. 11].

в 1949 г. На рис. 5.38 приведены более поздние данные для стеклообразного твердого тела.

Коэффициент диффузии для атомов в твердом теле обычно записывается в виде

$$D=\xi a^2\nu_j=\xi a^2\nu_D\exp(-E/k_0T). \qquad (5.135)$$

Здесь $\nu_D=(k_0\theta_D/h)$ — максимальная частота колебаний в твердом теле, a — расстояние до ближайших соседей, а E — энергия активации для перемещения атома из одного положения в решетке в другое. Безразмерный множитель ξ зависит от типа решетки, но во всех случаях он порядка единицы. Дебаевская частота $\nu_D\sim 10^{12}\div 10^{13}$ Гц, но частота перескоков $\nu_j=\nu_D\cdot\exp(-E/k_0T)$ будет очень малой, если велика энергия активации.

Таким образом, уширенная диполь-дипольным взаимодействием линия ЯМР в твердом теле будет сужаться при тех температурах, при которых частота перескоков ν_j становится соизмеримой с шириной линии $\Delta\nu_{pp}$. В случае атомов лития в стеклообразном образце (рис. 5.38) это условие реализуется при температурах более низких, чем комнатная температура. При этом энергия активации для диффузии может быть определена по значению температуры, при которой $\Delta\nu_{pp}$ начинает уменьшаться. Отсюда следует, что ЯМР полезен для изучения диффузии в твердых телах; сведения о величине энергии активации диффузии позволяют выяснить, какой из механизмов ответствен за диффузионный процесс.

ЯМР в металлах; сдвиг Найта

Наблюдение ЯМР в металлах имеет свои особенности, которые связаны с высокой концентрацией свободных электронов. Прежде всего отметим, что вследствие скин-эффекта высоко-

частотное электромагнитное излучение проникает в металлы лишь на малую глубину. Поэтому исследования ЯМР в металлах легче всего проводить на образцах в виде тонких фольг или микроскопических шариков, помещенных в диэлектрическую среду.

Кроме того, наличие электронов проводимости существенно повышает скорость спин-решеточной релаксации, и поэтому значения величины τ_1 в металлах значительно ниже, чем в неметаллах. Для большинства металлов τ_1 изменяется как $1/T$ и достигает значений $\tau_1 > 1$ с лишь при температурах жидкого гелия.

Наиболее существенная особенность ЯМР в металлах по сравнению с неметаллами состоит в наличии *сдвига Найта*[46, 50] — изменения резонансного поля, которое обусловлено электронами с фермиевской энергией. Этот сдвиг всегда происходит в сторону более высоких частот при данном магнитном поле (иными словами, в сторону бóльших эффективных значений g_N); он прямо пропорционален напряженности приложенного магнитного поля, имеет бóльшую величину для тяжелых элементов и изменяется с изменением доли s-состояний в волновой функции электронов проводимости. Величина сдвига Найта обычно представляется безразмерным параметром $K = (\Delta\nu/\nu) = (-\Delta B/B)$, который изменяется в пределах от менее чем 0,001 % для ^{9}Be, до 1,5 % для ^{133}Cs и ^{205}Tl.

В теории сдвига Найта[51] рассматривается магнитное поле, создаваемое на атомном ядре посредством сверхтонкого взаимодействия между моментом ядра и моментами неспаренных s-электронов проводимости. Сдвиг Найта, как и сверхтонкое расщепление линии ЭПР, возникает в результате фермиевского контактного взаимодействия[39]. Вклад каждого электрона зависит от времени его нахождения на ядре и, следовательно, от $|\psi(0)|^2$.

При этом следует принимать во внимание лишь электроны на поверхности Ферми, поскольку все электроны с более низкой энергией занимают парные состояния. Таким образом, сдвиг Найта зависит от плотности состояний $g(\varepsilon_F)$ на поверхности Ферми[52] и от величины парамагнитной восприимчивости Паули $\chi_P = \mu_0 \mu_B^2 g(\varepsilon_F)$ [см. выражение (3.68)], обусловлен-

---

[50] В примечании 46 приведена первая статья Найта по этому вопросу; подробнее это явление рассматривается в его работе: *Knight W. D.*— In: Solid State Physics, v. 2, eds. F. Seitz, D. Turnbull, Academic Press, 1956, p. 93.

[51] *Townes C. H., Herring C., Knight W. D.*— Phys. Rev., 77, 852 (1950).

[52] Для газа *свободных электронов* $g(\varepsilon_F) = (3n/2\varepsilon_F)$ (см. выражение (3.69)), но периодический потенциал, введенный в зонной теории, может изменить величину $g(\varepsilon_F)$ в большую или меньшую сторону.

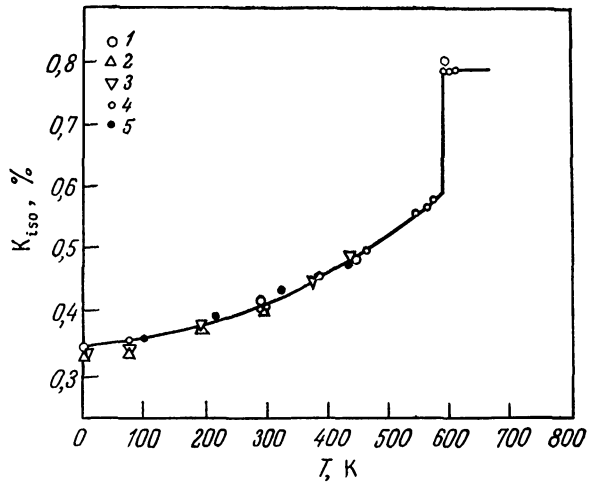


Рис. 5.39. Температурная зависимость изотропного сдвига Найта K_{iso} для ^{113}Cd (представлен в процентах от $\Delta\nu/\nu$). Рисунок взят из работы: *Kasowski R. V., Falicov L. M.*— Phys. Rev. Lett., 22, 1001 (1969). Приведены расчетные точки для кристаллического (при 0,298 и 462 К), жидкого (при 600 К) кадмия и экспериментальные результаты ряда авторов (*1* — теоретические расчеты, *2* — Гудрич и Кан, *3* — Шарма и Вильямс, *4* — Сеймур и Стайлз, *5* — Борса и Барнес). Касовски и Фаликов пришли к выводу, что для расплавленного кадмия $g(\varepsilon_F)$ и $|\psi_F(0)|^2$ такие же, как и для газа свободных электронов, а при абсолютном нуле значение $|\psi_F(0)|^2$ равно только 0,74 от соответствующей величины для свободных электронов; значение $g(\varepsilon_F)$ составляет еще меньшую долю.

ной электронами проводимости на поверхности Ферми. Следовательно, сдвиг Найта можно записать в виде

$$K = C\chi_P \langle |\psi_F(0)|^2 \rangle_{ср}. \qquad (5.136)$$

Здесь усреднение проводится по всем состояниям электронов на поверхности Ферми. Сдвиг Найта можно выразить и через константу сверхтонкого взаимодействия. Однако выражение (5.136) позволяет понять основные закономерности. Из этого выражения видно, почему величина K больше для более тяжелых элементов и каким образом она зависит от доли s-состояний в волновой функции электронов проводимости.

Для большинства металлов сдвиг Найта K слабо зависит от температуры; в интервале температур от абсолютного нуля до точки плавления он изменяется менее чем на 10 %. Интересным исключением являются экспериментальные данные для ^{113}Cd, которые приведены на рис. 5.39. Как было показано Ка-

совски и Фаликовым [53], такую температурную зависимость K можно объяснить температурной зависимостью $g(\varepsilon_F)$ и $|\psi_F(0)|^2$, поскольку в кадмии имеет место сложное перекрытие зон. Каждый атом кадмия отдает по два электрона. При этом первая зона Бриллюэна оказывается практически заполненной электронами, а вторая зона и крошечные карманы третьей и четвертой зон заполнены умеренно. В этом случае важную роль играют s—p-гибридизованные состояния второй зоны, но с повышением температуры они становятся в большей степени s-подобными. Кроме того, изменение периодического потенциала с нагреванием приводит к возрастанию заполнения s-состояний третьей и четвертой зон. Оба этих фактора обусловливают увеличение K. При плавлении распределение электронов становится подобным распределению в газе свободных электронов.

Таким образом, экспериментальные результаты по сдвигу Найта в сочетании с другой информацией о металле позволяют оценить $g(\varepsilon_F)$ и долю s-состояний для состояний на поверхности Ферми. В задаче 5.21 предлагается, используя данные рис. 5.39, выполнить упрощенную оценку изменения $g(\varepsilon_F)$ с температурой для кадмия.

Ядерный квадрупольный резонанс

Если энергетические уровни ядра расщеплены только магнитным полем, то $(2I+1)$ уровней соответствуют арифметической прогрессии, представленной уравнением (5.130). Правило отбора $\Delta m_I = \pm 1$ допускает $2I$ переходов между этими уровнями, и каждый из этих переходов требует одинаковой энергии $h\nu = g_N \mu_n B$. Однако если на энергетический спектр оказывает влияние взаимодействие квадрупольного момента Q с градиентом электрического поля, то для осуществления этих переходов уже необходимо $2I$ незначительно различающихся по величине квантов энергии. Изменение характера спектра ЯМР при наличии квадрупольных взаимодействий [54] возможно для ядер с $I \geqslant 1$:

Квадрупольный момент Q характеризует отклонение в распределении заряда ядра от сферического. Он является положительным, когда заряд вытянут вдоль спиновой оси и отрицательным, если заряд сжат вдоль этой оси. Строго Q представляется в виде бесследового симметричного тензора с пятью

---

[53] *Kasowski R. W., Falicov L. M.*— Phys. Rev. Lett., 22, 1001 (1969).

[54] Для кристаллических твердых тел оно впервые было описано в работе: *Pound R. V.*— Phys. Rev., 79, 685 (1950).

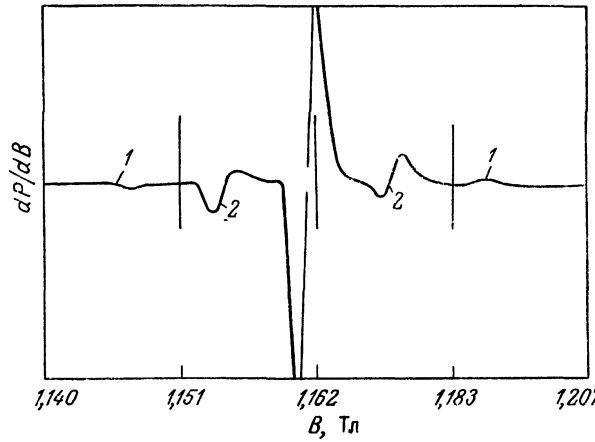


Рис. 5.40. Производная линии поглощения ЯМР для ^{11}B в порошкообразном CrB_2 на частоте 16 МГц. Сателлиты основной линии для этого некубического твердого тела возникают в результате взаимодействия ядерного квадрупольного момента с градиентом электрического поля (ГЭП). Некоторые кристаллиты ориентированы таким образом, что главная ось тензора ГЭП оказывается параллельной **B** (линии *1*); для вдвое большего числа кристаллитов эта ось перпендикулярна **B** (линии *2*). (Рисунок взят из книги: *Barnes R. G.*, Magnetic Resonance, eds. C. K. Coogan et al., Plenum Press, 1970, p. 63.)

независимыми компонентами. Однако для наших целей оказывается достаточным представить Q в виде скаляра

$$Q = \int (1/2e)(3z^2 - r^2)\rho(r)\, d^3r. \qquad (5.137)$$

Квадрупольный момент Q имеет размерность [длина]^2, и, как указано в последней колонке табл. 5.5, он обычно выражается в барнах (10^{-28} м^2) — единицах, используемых в ядерной физике.

Квадрупольный момент ядра взаимодействует с градиентом электрического поля (ГЭП) в месте расположения ядра. Если составляющая магнитного поля вдоль оси z равна B_z и соответствующий компонент тензора ГЭП равен d^2V/dz^2, то вместо выражения (5.130) энергию ядра следует записать в виде

$$u = u_0 - g_N\mu_n B_z m_I + eQ\left(\frac{d^2V}{dz^2}\right)\left[\frac{3m_I^2 - I(I+1)}{4I(2I-1)}\right]. \qquad (5.138)$$

Для стимулирования перехода $m_I \rightarrow (m_I - 1)$ теперь необходима энергия

$$h\nu = g_N\mu_n B_z - eQ\left(\frac{d^2V}{dz^2}\right)\left[\frac{3(2m_I - 1)}{4I(2I-1)}\right]. \qquad (5.139)$$

Она совпадает с энергией для обычного ЯМР лишь в случае переходов $1/2 \rightarrow -1/2$. Однако переходы между любой другой парой состояний будут реализовываться при иных частотах (для данного значения B_z) или иных B_z (для данной частоты), если ядро обладает отличным от нуля квадрупольным моментом и в месте его расположения в кристалле компонент тензора ГЭП d^2V/dz^2 также отличен от нуля*. Возникающие при этом «сателлитные» линии со стороны больших и меньших полей могут быть использованы для *оценки* градиента электрического поля в месте расположения атомного ядра, обладающего спином. Таким образом, ядерный квадрупольный резонанс играет важную роль в физике твердого тела.

На рис. 5.40 приведен типичный пример квадрупольного расщепления линии ЯМР. Ядро ^{11}B имеет спиновое квантовое число $I=3/2$. Следовательно, спектр ЯМР должен состоять из трех линий, если все ядра ^{11}B занимают эквивалентные узлы в кристаллической решетке и одинаковым образом ориентированы по отношению к направлению B_z. На рисунке можно видеть интенсивную центральную линию (соответствующую переходам из состояния с $m_I=1/2$ в состояние с $m_I=-1/2$) и *два* набора сателлитов, соответствующих кристаллитам различных ориентаций. Детальная информация о компонентах тензора градиента электрического поля может быть получена в подобных случаях из анализа характера изменения квадрупольного спектра при вращении образца.

Спектры ЭПР и ЯМР для парамагнитных ионов и атомных ядер в твердых телах часто оказываются довольно сложными и требуют непростого анализа. Читатель, который заинтересуется примерами, приведенными в настоящем разделе, может получить исчерпывающую информацию по всем этим вопросам в библиографии, представленной в конце данной главы.

Задачи

5.1. Сложите векторно компоненты электрического поля, связанные с зарядами $+q$ и $-q$ на рис. 5.1, и покажите, что электрическое поле, создаваемое диполем, действительно описывается уравнением 5.6.

5.2. Рассмотрим сферическую полость внутри однородного диэлектрика, у которого дипольный момент единицы объема равен **P**. Будем считать, что поляризация направлена вдоль оси z, а угол между осью z и вектором **r**, который направлен от центра полости к некоторой точке на стенке полости, обозначим через θ. Запишите поле $\mathbf{E}_{\text{пов}}$ в центре полости, создаваемое зарядами, распределенными по поверхности полости, через интеграл по θ и покажите, что интегрирование приводит к результату $\mathbf{P}/3\varepsilon_0$, который записан в соотношении (5.11). Как изменится выражение для $\mathbf{E}_{\text{пов}}$, если мы запи-

* Взаимодействие ядерного квадрупольного момента с градиентом электрического поля можно наблюдать по поглощению радиочастотной мощности и в отсутствие внешнего магнитного поля. — *Прим. перев.*

шем его для некоторой точки, смещенной от центра полости вдоль направления **P**?

5.3. Разложите в ряд до членов с w^5 экспоненты выражения (5.25) и покажите, что при условии слабого поля функция Ланжевена представляется сходящимся рядом вида:

$$L(w) \approx (w/3)\,[1 - w^2/15 + \ldots], \qquad w \ll 1.$$

Используя полный вид выражения (5.25), покажите, что, в случае когда поляризация составляет 50 % от величины насыщения, можно записать простое трансцендентное уравнение для $w_{0,5}$, которое просто решается с помощью логарифмической линейки и дает результат $w_{0,5}=1{,}80$. Нет никакой необходимости использовать таблицы для $L(w)$, чтобы показать это. Покажите, что ситуация, когда степень поляризации составляет 50 % от величины насыщения, требует для молекул воды при комнатной температуре неправдоподобно большого поля в 10^9 В/м.

5.4. Исследования температурных зависимостей действительной и мнимой частей диэлектрической проницаемости были выполнены для полярного твердого тела, у которого температура Дебая для колебаний решетки равна 153 К. В результате измерений диэлектрических свойств на частоте 110 кГц был выявлен максимум на кривой диэлектрических потерь, расположенный при температуре $T=270$ К. Покажите, что это согласуется с существованием двух возможных ориентаций диполей с барьером между ними в 0,4 эВ. Положив $A=5$ и $B=15$ в уравнении (5.33), постройте кривые зависимостей $\varkappa_1$ и $\varkappa_2$ для этого твердого тела на частоте 110 кГц. У вас должно получиться, что мнимая часть диэлектрической проницаемости в интервале температур от 250 до 290 К достигает 50 % от максимального значения.

5.5. Покажите, что решением уравнения движения (5.41) является соотношение (5.42) и это приводит к выражению (5.43) для поляризуемости. Подставляя последнее в соотношение Клаузиуса — Мосотти, покажите, что действительная и мнимая части диэлектрической проницаемости даются выражениями (5.46) и (5.47). Определите точки перегиба $\varkappa_1$ [выражение (5.46)] и используйте этот результат для нахождения величин γ, ω_1, ω_0, $e^2N_e/m\varepsilon_0$ и $\beta=\omega_0^2/m$ для некоторого твердого тела, у которого действительная часть диэлектрической проницаемости в видимой области спектра проходит через максимум и минимум. При этом максимуму $\varkappa_1$ соответствует частота $\omega=1{,}67\cdot10^{15}$ рад/с, а последующий за ним минимум $\varkappa_1=0{,}25$ имеет место при $\omega=3{,}6\cdot10^{15}$ рад/с. Вы должны получить, что постоянная возвращающей силы для электронов в этом твердом теле $\beta=10$ Н/м. Используя логарифмическую шкалу для частоты, постройте кривые зависимости от частоты для действительной и мнимой частей диэлектрической проницаемости. Вы должны получить, что в низкочастотной области спектра $\varkappa=\varkappa_1=2{,}25$.

5.6. Два атома с большой ионной поляризуемостью разведены на расстояние D друг от друга, находятся далеко от какого-либо другого вещества и экранированы от внешних полей. Какова должна быть поляризуемость α_i для того, чтобы спонтанная поляризация для этих атомов оказалась энергетически выгодной?

5.7. Рассмотрим электрон атома водорода, который движется по круговой орбите радиуса r с угловой частотой ω. Не принимая во внимание квантовые ограничения, напишите выражение для ω через e, m и r. Теперь рассмотрим ситуацию, когда магнитная индукция направлена перпендикулярно плоскости орбиты; действие силы Лоренца нарушает равновесие между центробежной и центростремительной силами. Покажите, что в первом порядке по B изменение угловой частоты можно записать в виде $\Delta\omega=(eB/2m)$. Этот результат известен как *теорема Лармора* и является основным в классиче-

ской теории диамагнетизма. Сколь велико должно быть магнитное поле для того, чтобы $\Delta\omega$ равнялось 1 % от частоты орбитального движения электрона в атоме водорода?

5.8. Рассмотрим систему парамагнитных атомов (N в единице объема), в каждом из которых все электроны, кроме одного, находятся в спаренных состояниях, а этот единственный электрон — в s-подобном состоянии. Какие значения m_l возможны? Нарисуйте схему энергетических состояний с учетом этих возможных значений m_l в магнитном поле, укажите энергию каждого состояния в магнитном поле. Покажите, что если заселенности состояний определяются распределением Больцмана, то *разность* заселенностей приводит к намагниченности единицы объема в виде

$$M = N\mu_B \operatorname{th}(B\mu_B/k_0T).$$

Покажите, что это выражение может быть получено из выражения (5.80) для условий настоящей задачи.

5.9. В этой задаче мы хотим решить уравнение (5.93) для намагниченности ферромагнетика в простом случае при $J=1/2$. Покажите, что $B_{1/2}(y)$ оказывается простой гиперболической функцией безразмерной переменной $y=(\gamma\mu_0\mu_1 M_s/k_0T)$, и постройте эту функцию для положительных значений y. На этот же график нанесите набор прямых линий, выходящих из начала координат, которые пересекают эту кривую. Эти прямые должны представлять линейные зависимости (M_s/M_{s0}) от y, соответствующие различным температурам. Вы можете использовать асимптотическое приближение $B_{1/2}(y)$ для случая слабого поля при $T=T_c$ для того, чтобы выразить T_c через $\gamma\mu_0\mu_1 M_{s0}$ и, следовательно, откалибровать наклоны ваших прямых линий. Определите из точек пересечения несколько комбинаций (M_s/M_{s0}) и (T/T_c) и сравните их с кривой и экспериментальными данными рис. 5.15.

5.10. Как показано на рис. 5.18, энергия магнона изменяется приблизительно как квадрат волнового вектора. Исходя из этого, получите зависимость плотности состояний для магнонов от их энергии в виде степенной функции. Считая далее, что состояния магнонов заполняются в соответствии с распределением Бозе — Эйнштейна, покажите, что получаемое таким образом полное число магнонов, возбуждаемых при любой температуре T, приводит к зависимости спонтанной намагниченности от температуры, которая согласуется с «законом $T^{3/2}$ Блоха». Таким же образом покажите, что магнонный вклад в теплоемкость пропорционален $T^{3/2}$. Посмотрите, как изменятся эти результаты, если принять во внимание, что дальнодействующее влияние анизотропии приводят к отличной от нуля минимальной энергии магнона, соответствующего спиновой волне с бесконечной длиной волны.

5.11. Киттель в своей монографии «Введение в физику твердого тела» для описания молекулярного поля в антиферромагнетике вместо выражения (5.97) предлагает более точные соотношения

$$H_A = (-\gamma M_B - \eta M_A); \qquad H_B = (-\gamma M_A - \eta M_B).$$

Покажите, что в этом случае отношение отрицательной температуры, отсекаемой прямой $(1/\chi_m)=f(T)$ на оси температур, к температуре Нееля T_N описывается формулой

$$(\theta/T_N) = (\gamma+\eta)/(\gamma-\eta).$$

5.12. Примените математический формализм задачи 5.11 для случая *ферримагнитного* материала, в котором магнитные моменты двух подрешеток имеют различные величины. Покажите, что это приводит к выражению (5.102) для магнитной восприимчивости при высоких температурах.

5.13. Вычислите момент количества движения L_s и магнитный момент μ для сферы массы M (масса равномерно распределена по объему), имеющей заряд Q (заряд равномерно распределен по поверхности), предполагая, что ее радиус равен r, а угловая скорость ω. Покажите, что магнитный момент и момент количества движения связаны соотношением (5.103).

5.14. Продифференцируйте выражения (5.115) и (5.117) для того, чтобы определить положение максимума и минимума кривой (dP/dB) для линий ЭПР лоренцевой и гауссовой формы. Как связаны между собой ΔB_{pp} и 2Г для гауссовой кривой? Скопируйте рис. 5.27 и нанесите на эту копию кривые гауссовой формы: одну — для ΔB_{pp}, такой же, как и для лоренцевой кривой, а другую для $\Delta B_{1/2}=2\Gamma$, такой же, как и для лоренцевой кривой. Определите положение $P_{\text{макс}}$ для экспериментальной кривой и затем оцените эффективный g-фактор. Определите также полуширину Г и выразите ее через эквивалентный интервал частот. Теперь, используя принцип неопределенности Гейзенберга, вычислите соответствующее значение времени релаксации τ_r и сравните его с данными рис. 5.25.

5.15. Постройте кривую, описывающую, каким образом должно меняться с температурой *внешнее* магнитное поле, если вы хотите наблюдать антиферромагнитный резонанс в образце MnF_2 (см. рис. 5.30) на частоте 10 ГГц.

5.16. В любой современной книге по физике можно найти, что радиальная часть волновой функции для основного состояния атома водорода имеет вид $\psi(r)=(\pi a^3)^{-1/2}\exp(-r/a)$. Вычислите $|\psi(0)|^2$ в $м^{-3}$ и полученный результат используйте для определения коэффициента $\mathscr{A}$ фермиевского контактного взаимодействия в этом атоме. С помощью этой величины (в джоулях) вычислите сверхтонкое расщепление $(\mathscr{A}/g\mu_B)$ (в теслах), которое, как вы знаете, для атома водорода должно быть равно 0,0508 Тл.

5.17. Напишите решения уравнения (5.128) для $\theta=0$ и $\theta=\pi/2$. Имеются три решения для каждого значения θ. Проверьте, что решения для $\theta=\pi/2$ соответствуют тем кривым, которые приведены на рис. 5.34, и изобразите графически три решения для $\theta=0$. Отметьте на этом рисунке два перехода, которые могут быть стимулированы квантами с энергией $h\nu>D$.

5.18. Покажите, что схемы энергетических уровней задачи 5.17 допускают такое количество линий ЭПР, удовлетворяющих правилу отбора $\Delta m_s=\pm 1$, которое оказывается достаточным для определения $g_\parallel$, $g_\perp$ и D по экспериментальной кривой (являющейся комбинацией параллельного и перпендикулярного спектров) вне зависимости от того, используются кванты с энергией $h\nu>D$ или с энергией $h\nu<D$.

5.19. Выразите ширину линии ЯМР для ^{63}Cu в хлориде меди в величинах $\Delta\nu_{pp}$ и ΔB_{pp} (рис. 5.36). Во сколько раз эта ширина линии больше, чем ширина линии для протонов группы OH в жидком этиловом спирте (см. рис. 5.37)?

5.20. Оцените температуру, при которой начинает уменьшаться ширина линии на рис. 5.38. Покажите, что это наблюдаемое уменьшение ширины линии согласуется с энергией активации для диффузии 0,48 эВ, как указано в подписи к рисунку. Данные табл. 2.2 позволяют предположить, что для стеклообразного вещества, которое в основном состоит из кремнезема, $\theta_D\simeq 600$ К.

5.21. В этой задаче предлагается провести упрощенный анализ экспериментальных данных рис. 5.39 по сдвигу Найта для кадмия. Найдите плотность этого металла (считая ее одинаковой для кристаллического и жидкого состояний) и определите плотность электронов, энергию Ферми ε_F, фермиев-

скую плотность состояний $g(\varepsilon_F)$, восприимчивость Паули χ_P для газа свободных электронов, принимая во внимание, что на каждый атом кадмия приходится два электрона. Предполагается, что это является хорошим допущением для жидкого кадмия. Из величины сдвига Найта для жидкого металла определите параметр C выражения (5.136). Теперь обратите внимание, что, как следует из подписи к рисунку, $|\psi_F(0)|^2$ при абсолютном нуле меньше единицы, и для простоты считайте, что она имеет ту же величину для кристаллического кадмия при всех температурах. При этом (некорректном) предположении оцените, как изменяется с температурой $g(\varepsilon_F)$ для кристаллического кадмия, и представьте ваши результаты графически, откладывая по оси ординат $g(\varepsilon_F)$ как долю от соответствующей величины для свободных электронов.

5.22. Исследуя ЯМР на ядрах ^{23}Na, имеющих в кристалле анизотропное окружение, вы предполагаете наблюдать сателлитные линии, обусловленные взаимодействием ядерного квадрупольного момента с градиентом электрического поля. Эксперимент проводится на частоте 30 МГц, ширина линии ЯМР $\Delta B_{pp} = 10^{-3}$ Тл, и сателлитные линии плохо разрешаются. Каков верхний предел для (d^2V/dz^2) в месте расположения ядер ^{23}Na? Каким образом необходимо изменить условия эксперимента, чтобы улучшить разрешение сателлитных линий?

Литература

Диэлектрики

Brown W. F.— In: Handbuch der Physik, v. 17, Springer 1956.
Bungat I., Popescu M. Physics of Solid Dielectrics, North-Holland, 1984.
Fröhlich H. Theory of Dielectrics, Oxford Univ. Press, 2nd ed., 1958. [Имеется перевод: *Фрелих Г.* Теория диэлектриков.— М.: ИЛ, 1960.]
Scaife B. K. P. Complex Permittivity, Crane Russak, 1971.
Желудев И. С. Физика кристаллических диэлектриков.— М.: Наука, 1968.

Пьезо- и сегнетоэлектрики

Burfoot J. C. Ferroelectrics, Van Nostrand, 1967. [Имеется перевод: *Барфут Дж.* Введение в физику сегнетоэлектрических явлений.— М.: Мир, 1970.]
Cady W. G. Piezoelectricity, Dover, 1964. [Имеется перевод: *Кэди У.* Пьезоэлектричество.— М.: ИЛ, 1962.]
Forsbergh P. W.— In: Handbuch der Physik, v. 17, Springer, 1956.
Фридкин В. К. Сегнетоэлектрики-полупроводники.— М.: Наука, 1976.
Jona F., Shirane G. Ferroelectric Crystals, Pergamon Press, 1962. [Имеется перевод: *Иона Ф., Ширане Д.* Сегнетоэлектрические кристаллы.— М: Мир, 1965.]
Lang S. B. Source Book of Pyroelectricity, Gordon and Breach, 1974.
Lines M. F., Glass A. M. Principles and Applications of Ferroelectrics and Related Materials, Oxford Univ. Press, 1977.

Диамагнетизм и парамагнетизм

Cullity B. D. Introduction to Magnetic Materials, Addison-Wesley, 1972.
Goodenough J. B. Magnetism and Chemical Bond, Krieger, 1976.
Kalvius G. M., Tebble R. S. Experimental Magnetism, Wiley, 1979.
Martin D. H. Magnetism in Solids, MIT Press, 1967.
Van Vleck J. H. Theory of Electric and Magnetic Susceptibilities, Oxford Univ. Press, 1932.

Wagner D. Introduction to the Theory of Magnetism, Pergamon Press, 1972.
White R. M. Quantum Theory of Magnetism, McGraw-Hill, 1970. [Имеется перевод: *Уайт Р.* Квантовая теория магнетизма.— М.: Мир, 1972.]

Ферро-, антиферро- и ферримагнетизм

Boll R. Soft Magnetic Materials, Wiley, 1979.
Bozorth R. M. Ferromagnetism, Van Nostrand, 1951. [Имеется перевод: *Бозорт Р. М.* Ферромагнетизм.— М.: ИЛ, 1956.]
Chikazumi S. Physics of Magnetism, Krieger, 1978.
Magnetic Oxides, ed. D. J. Craik, Wiley, 1975.
Magnetism of Metals and Alloys, ed. M. Cyrot, North-Holland, 1982.
Keffer F.— In: Handbuch der Physik, v. 18, Springer, 1966. (Спиновые волны).
Morrish A. H. Physical Principles of Magnetism, Krieger, 1980.
O'Dell T. H. Magnetic Bubbles, Halsted, 1974.
Magnetism, eds. G. T. Rado, H. Suhl, Academic Press, 5 vols., 1963—1973.
Standley K. J. Oxide Magnetic Materials, Oxford Univ. Press, 2nd ed., 1972.
Handbook on Ferromagnetic Materials, ed. E. P. Wohlfarth, North-Holland, 1980, 3 vols.
Zeiger H. J. Magnetic Interactions in Solids, Oxford Univ. Press, 1973.
Winkler G. Magnetic Garnets, Vieweg, 1981.

Магнитный резонанс

Abragam A., Goldman M. Nuclear Magnetism: Order and Disorder, Oxford Univ. Press, 1982. [Имеется перевод: *Абрагам А., Гольдман М.* Ядерный магнетизм: порядок и беспорядок.— М.: Мир, 1984.]
Andrew E. Nuclear Magnetic Resonance, Cambridge Univ. Press, 1958. [Имеется перевод 1-го издания: *Эндрю Э.* Ядерный магнитный резонанс.— М.: ИЛ, 1957.]
Ault A., Dudek G. An Introduction to Proton NMR Spectroscopy, Holden-Day, 1976.
Magnetic Resonance, eds. C. K. Coogan et al., Plenum Press, 1970.
Habeberlen U. High Resolution NMR in Solids, Academic Press, 1976.
Theoretical Foundations of Electron Spin Resonance, ed. J. E. Hamman, Academic Press, 1978.
Low W. Paramagnetic Resonance in Solids, Academic Press, 1960. [Имеется перевод: *Лоу У.* Парамагнитный резонанс.— М.: ИЛ, 1962.]
Pake G. E., Estle T. L. Physical Principles of Paramagnetic Resonance, Benjamin, 2nd ed., 1972.
Poole C. P. Electron Spin Resonance, Wiley, 2nd ed., 1983.
Slichter C. P. Principles of Magnetic Resonance, Springer, 2nd ed., 1980. [Имеется перевод 1-го издания: *Сликтер Ч.* Основы теории магнитного резонанса.— М.: Мир, 1967.]
Sparks M. Ferromagnetic Relaxation Theory, McGraw-Hill, 1964.
Wertz J. E., Bolton J. R. Electron Spin Resonance, McGraw-Hill, 1972.
Winter J. Magnetic Resonance in Metals, Oxford Univ. Press, 1971.

ПРИЛОЖЕНИЕ

Значения некоторых физических констант *

Величина	Символ	Значения
Скорость света в вакууме	c	$2{,}9979 \cdot 10^{8}$ м/с
Диэлектрическая проницаемость вакуума (электрическая постоянная)	$\varepsilon_0 = 10^7/4\pi c^2$	$8{,}8542 \cdot 10^{-12}$ Ф/м
Магнитная проницаемость вакуума (магнитная постоянная)	$\mu_0 = 4\pi \cdot 10^{-7}$	$1{,}2566 \cdot 10^{-6}$ Гн/м
Постоянная Планка	h $\hbar = h/2\pi$	$6{,}6262 \cdot 10^{-34}$ Дж·с $1{,}0546 \cdot 10^{-34}$ Дж·с
Элементарный заряд	e	$1{,}6022 \cdot 10^{-19}$ Кл
Электронвольт	эВ	$1{,}6022 \cdot 10^{-19}$ Дж
Атомная единица массы	а. е. м.	$1{,}6605 \cdot 10^{-27}$ кг
Масса покоя электрона	m	$9{,}1096 \cdot 10^{-31}$ кг
Масса покоя протона	M_p	$1{,}6726 \cdot 10^{-27}$ кг
Масса покоя нейтрона	M_n	$1{,}6749 \cdot 10^{-27}$ кг
Комптоновская длина волны электрона	$\lambda\!\!\!^{-}_c = \hbar/mc$	$3{,}8616 \cdot 10^{-13}$ м
Ангстрем	Å	$1{,}0000 \cdot 10^{-10}$ м
Боровский радиус	$a_0 = 4\pi\varepsilon_0\hbar^2/me^2$	$5{,}2918 \cdot 10^{-11}$ м
Постоянная тонкой структуры	$\alpha = \lambda\!\!\!^{-}_c/a_0$	1/137,036
Постоянная Ридберга	$R_\infty = \hbar^2/2ma_0^2$	$2{,}1799 \cdot 10^{-18}$ Дж 13,605 эВ
Магнетон Бора	$\mu_B = e\hbar/2m$	$9{,}2742 \cdot 10^{-24}$ Дж/Тл
Ядерный магнетон	$\mu_n = e\hbar/2M_p$	$5{,}0511 \cdot 10^{-27}$ Дж/Тл
Постоянная Больцмана	k_0.	$1{,}3806 \cdot 10^{-23}$ Дж/К
Число Авогадро	N_A	$6{,}0222 \cdot 10^{23}$ моль$^{-1}$
Газовая постоянная	$R = k_0 N_A$	8,3142 Дж/(моль·К)
Температура для $k_0T = 1$ эВ		$1{,}1605 \cdot 10^{4}$ К
Частота для $h\nu = 1$ эВ		$2{,}4180 \cdot 10^{14}$ Гц
Длина волны для $hc/\lambda = 1$ эВ		$1{,}2398 \cdot 10^{-6}$ м

* Значения фундаментальных физических констант взяты из работы: *Taylor B. N., Parker W. H., Langenberg D. N.* — Rev. Mod. Phys., 41, 375 (1969).

Предметный указатель

Азбеля — Канера модель 310, 311
— — резонанс 309, 310, 311
Азимутальное квантовое число 534
Акустическая ветвь 140
Аллотропный переход 18, 23
Алмаз, колебательный спектр 145, 146
— молярная теплоемкость 158
— структура 68, 69
— энергетические зоны 279
Альфвена волны 247
Алюминий, ферми-поверхность 286
— электропроводность 188
Амбиполярная проводимость 405
Аморфные твердые тела 13
— — — валентная зона 433
— — — дифракция 91—95
— — — проводимость 432
Аморфный германий, фотоэмиссия 434
— кремний, дифракция электронов 94
— — проводимость 435, 436
Амфотерные центры 386
Ангармонические эффекты 117, 148, 170
Аномальный скин-эффект 296
Антимонид индия, электропроводность 420
Антисегнетоэлектричество 526
Антиструктурное разупорядочение 105
Антиферромагнетизм 533, 551—555
Антиферромагнитный резонанс 569—571
Арсенид галлия, зоны проводимости 443, 449, 450
— — теплопроводность 177
Атомные координаты 64
— орбитали 285
— энергетические уровни 262, 263, 273
Атомный радиус 32, 33
— фактор рассеяния 83, 84, 92

Базис 43, 44
Баллистический перенос, фононы 174
Блоха закон $T^{3/2}$ 547
— стенки (границы) 549
— функции 250—254, 263
Блоха — Флокэ теорема 250—253, 280
Бозе — Эйнштейна статистика 157, 213, 359
Больцмана кинетическое уравнение 177, 200, 410
— распределение, постоянные диполи 507, 535
— — спины электронов 557
— — — ядер 582
— — энергия решетки 155
— — — электронов 195, 410
Бора магнетон 221, 527, 556
— радиус
Борна — Кармана граничные условия 133, 211
Браве (пространственная) решетка 43
— — двумерная 52
— — трехмерная 56
Брейта — Раби уровни 573
Бриллюэна зона (зоны) 100
— — гексагональная решетка двумерная 270
— — многомерные твердые тела 265
— — объем 135, 265
— — одномерная решетка 125
— — представление зон повторяющихся 295
— — — — приведенных 253, 254, 260, 268
— — — — расширенных 253, 254, 269

— — простая кубическая решетка 130, 271
— рассеяние 154, 183
— функция 535, 537
Броуновское движение 178, 184
Брэгга условие дифракции 80, 98, 99

Вакансия 103
Ванадий, фононный спектр 136
Ван Хова сингулярности 137, 276, 299, 300, 433, 504
Вейсенберга метод 89
Вейсса коэффициент 540
Векторы обратной решетки 97, 126, 130, 253
Взаимодействие вандерваальсово 18, 19, 26, 429
— спинов электрон — ядро 571—574
— фермиевское контактное 572, 576
— электрон — дырка 456—458
Вигнера — Зейтца метод ячеек 278
Видемана — Франца закон 188, 200
Водородная модель 299, 381
— связь 36, 37
Возбужденные состояния примесей 396
Волновая функция периодическая 211
Волновой вектор 96
— — сохранение 99
— — фонон 120
— — электрон 211
Вольфрам, электропроводность 188
Восприимчивость диамагнитная 224
— магнитная 225
— парамагнитная 224
Вырождение атомных состояний 273, 448
— в полупроводнике 407
— электронного газа 193, 218, 364

Галлий, теплоемкость 330
Галогениды щелочных металлов, наиболее выгодные структуры 26
Ганна эффект 445
Гексагональная плотная упаковка (г. п. у.) 39, 56, 65, 67
— решетка 53, 57, 64
— — *Бриллюэна* зона 270
Гелий, теплопроводность 179
Геликонные волны 246, 247
Германий аморфный 433
— дефекты *Френкеля* 106
— зоны 451
— край оптического поглощения 458
— поверхности постоянной энергии 452
— подвижность дырок 407
— скорость электронов 440, 442
— собственная щель 458
— теплопроводность 176
— фотоэмиссия 434
Гиромагнитное отношение 556
Гомеополярная связь 21
Горячие электроны 405, 439
Грюнайзена выражение 230
Гука закон 117, 124, 138, 145
g-фактор 527, 528

Дебай, единица измерения 35, 500
Дебая модель 158, 159, 183
— температура 160, 165—167
— уравнения 512
Дебая — Шеррера метод 87—90
Де Бройля длина волны 77
Де Гааза — ван Альфена эффект 320—323, 354, 462
Детальное равновесие 359, 392, 465
Дефекты линейные 91, 174
— точечные 91, 174
— *Френкеля* 104
Джозефсона переход 349
Диамагнетизм, сверхпроводники 337
— электроны орбитальные 529—531
— — свободные 313—325
Дигидрофосфат калия (KDP) 519—521
Динамическое сужение 586, 587
Дипольная ориентация 503, 507—512
Дипольное взаимодействие 572, 576
Дипольный момент 19, 35, 498
— — эффективный 536
Дислокации 107
— винтовые 108
— краевые 108
— плотность 110
Дисперсионное уравнение 120
— — магноны 584
— — фононы 124, 139, 181
Диссипации постоянная 514
Дифракция, методы 91—95
— нейтронов 79, 80
— рентгеновских лучей 75—82
— электронов 78, 95
Диффузии коэффициент, атомы 589
Диффузионная длина электронов 486
Диффузионный ток 485
Диэлектрическая восприимчивость 497, 509

— проницаемость (постоянная) 147, 497
— — плазма 245
— — сероводород 509, 510
Диэлектрические потери 505, 514
Длина когерентности 338
Доменные стенки 524, 525, 549
Домены сегнетоэлектрические 524—526
— магнитные 549—551
Донорные примеси (центры) 363, 380—382
Дрейфовая скорость 198
Дрейфовый ток 484
Друде модель 196
Дырки, статистика 368, 369
Дюлонга и *Пти* закон 122, 158

Железо, электропроводность 188
Жесткости коэффициент. *См.* Модуль объемной упругости 118

Закон действующих масс 107, 376, 391, 392, 479
Зеемана расщепление 574
Зинера туннелирование 303, 508
Золото, амфотерный центр 404
— электропроводность 188
Зоммерфельда модель проводимости 225—228, 299, 364
Зона валентная 276, 361
— запрещенная (щель) 255, 260
— идеально простая структура 447
— проводимости 277, 361
— разрешенная 255
— твердые тела 447—452
Зонная схема, аморфные тела 436
— — металл 276
— — полуметалл 462, 463
— — полупроводник 276, 362
— — сверхпроводник 344
Зоны, алмаз 279
— алюминий 286
— арсенид галлия 443, 449
— германий 451, 452
— кремний 276, 292, 451
— магний 276
— мышьяк 463
— натрий 275
— никель 282, 540
— сульфид кадмия 448

Идеальный проводник 329
Изотопический эффект, сверхпроводимость 340
— — теплопроводность 174—176
Индий сверхпроводящий 328
— температура *Дебая* 165
Ионная поляризуемость 146—151, 504, 515, 516
— связь 25
Ионность 35
Ионный радиус 33
Ионы щелочных металлов, радиус 33
Иттрий, теплоемкость 162

Кадмий, электропроводность 188
Кальций, работа выхода 234
Квадрупольный момент 590
Квантование магнитного потока 336, 341
Кварц, монокристалл 15
— пьезоэлектричество 518, 519
Клаузиуса — Мосотти соотношение 506, 512, 515, 522
Кнудсеновское течение фононов 174, 178
Ковалентная связь 21, 22
Колебательные моды, цепочка двухвалентная 138—141
— — — одноатомная 122—127
Коллективные электронные эффекты 187, 325—345
Комбинационные полосы фононов 151
Комбинированная плотность состояний 300
— — — эффективная масса 300
Кона аномалии 131
Край подвижности 437
Крамерса — Кронига соотношения 504
Кремний аморфный 94, 434—437
— донорные состояния 379, 384
— зоны 276, 451
— решеточное поглощение 152
— собственная щель (запрещенная зона) 276, 373
— теплопроводность 176
Криотрон 348
Кристаллические системы 54—60
Кристаллографические направления 60, 62
Критическая температура (переход в сверхпроводящее состояние) 325, 328
Критическое поле 331—336
Кронига — Пенни модель 258—261, 285
Кулоновское притяжение, ионы 26
— — электрон — дырка 455
Куперовские пары 209, 336, 340—342
Кюри — Вейсса закон 533, 541

Кюри закон 222, 533, 536
— постоянная 534, 536
— температура 520, 533, 539—545

Лазер, полупроводник 480, 481
Ландау диамагнетизм 224, 313
— уровень 314, 460, 461
Ланде фактор расщепления 527, 528
Ланжевена функция 205, 504, 508, 534
Ларморова частота прецессии 530
Лауэ метод 87, 88
— уравнения 100
Лиддана — Сакса — Теллера соотношение 150
Локализованные состояния, электроны 257, 379
Локальные фононные моды 152, 153
Лондона глубина проникновения 335
Лоренца метод 501, 503
— модель 200—208
— сила 205, 307, 314
— число 188, 189, 200
Лоренцева форма линии 148, 563—565
Люминесценция 473, 478

Магнетит (Fe_3O_4) 90, 538
Магнетосопротивление 208, 228, 418
Магнетофононный резонанс 325, 461
Магнитная восприимчивость 497, 526, 530, 533
— индукция 310, 315, 497
— проницаемость 497
Магнитные домены 549—551
— подзоны 313—325
— подрешетки 551
Магнитный момент 527
Магнитопоглощение 464
Магноны 79, 546
Маделунга постоянная 30, 31, 111
Максвелла — Больцмана распределение 194
Максвелла уравнения 498
Малоугловая граница 109, 110
Маттисена правило 189, 219, 230
Медь, *Азбеля — Канера* резонанс 312
— колебательный спектр 128, 136, 161
— теплоемкость 122, 168
— ферми-поверхность 295
— электропроводность 189
Междоузельный атом 104
Мейснера эффект 335, 338
Металлическая связь 39—42
Металлы, свойства 187—190
— Ферми энергия 214, 215
Метод вращения кристалла 87—89
— качания 89
— квантового дефекта 386
— ЛКАО 261, 281
— ОПВ 285
— порошка 87
— ППВ 387
— прецессии 89
— элементарных возбуждений 249, 250
— ячеек 278
Миллера индексы 60—62
Многовалентные центры 386
Модель почти свободных электронов (ПСЭ) 261, 282
Модуль объемной упругости 118, 125
Молекулярное поле 540, 553
Мышьяк, поверхность *Ферми* 463

Найта сдвиг 584, 587—589
Намагниченность 222—224, 497, 526, 534, 541
Натрий, зонная структура 275
— электропроводность 207
Незатухающие токи 336
Непрерывности уравнения 202, 470
Непрямая щель 453
Никель, заполнение зон 282, 540
— плотность состояний 282
— ферромагнитное состояние 539
— электропроводность 207
Нормальные (N) процессы 171
— — влияние на теплопроводность 177
Носители, время жизни 468, 470
— диффузия и дрейф 484—487
— неосновные 377, 469
— основные 377, 469
— подвижность 405—410
— эксклюзия (обеднение) 469

Область истощения 394
Обменная энергия 544
Обменное сужение 563
Обменный интеграл 544, 567
Оболочечная модель 145
Обратное пространство 95
Операция винтового поворота 48, 59
— скольжения 48, 55, 59, 112
Оптическая ветвь 141—151
Оптические переходы 452—456
— фононы 412, 415, 437
Оптический переход непрямой 457

— — прямой 457, 458
Основное состояние, расщепление кристаллическим полем 398
Остаточные лучи 146
— — частота 149
Отрицательная дифференциальная проводимость. *См.* Ретроградная скорость

Парамагнетизм, магнитные атомы 533—540
— *Паули* 221—225, 301, 588
— электронный газ 221—225
Параэлектрическая восприимчивость 503, 588
Паули принцип 209, 212, 218, 222
Перенос с участием нескольких зон 419—422
Перовскит 74, 77
Плазмон 79, 246
Плоские группы 51, 54
Плотность состояний 221, 222, 298—301
Плотноупакованные структуры 64—67
Поворот доменов 550
Подвижность дрейфовая 199
— дырочная 405
Полинга шкала электроотрицательностей 34, 35
Положительная дырка 293
Положительные атомные остатки (ионы) 190—193
Полуметаллы 462
Полупроводники аморфные 432—439
— идеально простая зонная структура 447
— органические 429
— примесные 278, 361, 376
— собственные 277, 361
— *n*-типа 362
— *p*-типа 362
Поляризационное облако 415
Поляроны 415, 428
Поперечные волны 118, 126
Потенциал деформации, константа 410—412
Представление зон повторяющихся 295, 296
— — приведенных 253, 255, 268
— — расширенных 253, 255, 269
Примеси акцепторные 363, 381, 382
— донорные 363, 378—386
Пробой, диэлектрик 508
— *Зинера* 303, 508
— лавинный 303, 441, 508
— тепловой 438
Проводимость амбиполярная 405
— дырочная 293, 405
— модель *Зоммерфельда* 225—228
— — *Лоренца* 204
— полупроводник 404—410
— — собственный 419
— примесная 422—428
— сверхпроводник 329
— тепловая 167, 180, 189, 199
— типичные значения 188, 207
— электрическая 188—190, 199
Продольные волны 120, 126, 135
— — дисперсионное уравнение 121
Промежуточное состояние 337
Пространственные группы 59
Процессы переброса (*U*-процессы) 172, 173, 177—180
— — температура 177
— протекания (перколяционные) 427, 437
Прыжковая проводимость 424, 437
Псевдопотенциал 287
Пуазейлевское течение 179, 180
Пьезоэлектрические материалы, таблица 519
Пьезоэлектричество 517—519
p—n-переход 303, 487

Работа выхода 192, 193, 234, 235
Радиационное разрушение 105
Разрешенные переходы 454
Рассеяние комбинационное (рамановское) 153, 154
— междолинное 415
— на границе, фононы 174
— — дефектах, фононы 174, 175
— — — электроны 229—231, 414—416
— — ионизованных примесях 412—414
— — колебаниях решетки 410—412
— нейтронов когерентное 79, 128—130, 144—146
— — магнонами 548
— — некогерентное 79, 136
— — неупругое 81, 128—130, 548
— — фононами 79, 128—130, 144—146
— поляризационная мода 415
— упругое 195, 229, 247
— фононов в сплавах 176
— — *Казимира* условие 174, 179
— — фононами 153
Расщепление в кристаллическом поле 398, 399, 578—580

Рекомбинация излучательная 473, 478—481
— — вынужденная 478
— примесная 471—477
— *Оже́* 482
— фононный каскад 474
— *Холла — Шокли — Рида* (ХШР) 471, 482—484
Релаксация дипольная 510
— — постоянная времени 510
— — ширина линии ЯМР 585
— парамагнитная 561
— спин-решеточная 560—565
— — рамановский процесс 560
Рентгеновская дифракция 75—80
Ретроградная скорость 442
Решетка г. ц. к. 39, 46, 57
— — *Бриллюэна* зона 272, 286, 446
— косоугольная 53—55
— — *Бриллюэна* зона 101
— — обратная 97
— к. п. у. 39, 46, 66
— о. ц. к. 39, 46, 68
— — *Бриллюэна* зона 114
— примитивная кубическая 43, 68, 69
Ричардсона константа, таблица 234
— кривая 233, 234
Ромбические структуры 57—59
Ртуть, сверхпроводимость 328

Самокомпенсация 435
Самосогласованное поле, движение электрона 247, 249
Сверхпроводимость, применения 344—349
— теория БКШ 327
— — микроскопическая 339
Сверхпроводники жесткие (II рода) 333, 337, 346
— мягкие (I рода) 333, 337, 347
Сверхпроводящее состояние 304, 325
Сверхпроводящие полупроводники 327
— твердые тела, таблица 328
Сверхпроводящий магнит 345
Сверхтонкое расщепление 574—580
Свинец, сверхпроводимость 329, 332
— фононный спектр 128
— электропроводность 188, 329, 332
Свободная энергия 105
Свободные электроны, модель 192, 216
— — хаотическое и упорядоченное движение 193
Сегнетова соль 519
Сегнетоэлектрические домены 524, 525
— твердые тела 519—524
— — — применения 523, 524
— — — таблица 520
Сегнетоэлектричество 519—524
Серебро, электропроводность 207
Сильная связь, модель 261, 280
Симметрия зеркальная 48
— инверсия 48
— отражение 48
— поворот 48—51
— решетки 47—51
— трансляция 42, 43, 48
Скин-эффект 309
— аномальный 295, 296, 309
— нормальный 309
Сложные операции 48, 50, 59
Смешанные ионно-ковалентные связи 33
— ковалентные и вандерваальсовы связи 24, 25
Собственная щель 370—373, 452, 458
Собственные дефекты 103
Спиновые волны 546
Спиновый гамильтониан 574—578
Спин-орбитальное взаимодействие 448, 449
Спиральные структуры 38, 59, 109, 110
Средняя длина свободного пробега, упругое рассеяние 195, 196
— — — — фононы 171
— — — — электроны 199
Стехиометрия 107
Структура, алмаз 69
— вюрцит 71
— германий 434
— графит 25
— ДНК 38, 39
— изотактический полистирол 59
— кремнезем стеклообразный 92
— куприт (Cu_2O) 73
— лед 37
— перовскит 74
— теллур 24
— хлорид натрия (каменная соль) 72
— — цезия 72
— цинковая обманка (сфалерит) 69
— K_2PtCl_4 75
— $MoAl_{12}$ 73
— Si 270, 434
— SiF_4 73
— TCNQ 430, 431
Структурный фактор 85—87
Сульфид кадмия, зоны 448

Суперпозиции принцип 117
Сфалерит (цинковая обманка) 69, 70

Тайна закон 331
Тантал, работа выхода 234
Теплоемкость 330, 331
Теплопроводность решеточная 167—170, 176, 177
— электронная 180, 181, 188
Тетрагональная структура 57—59, 75
Титанат бария — стронция 519
— свинца — циркония (PZT) 519
Томаса — Ферми радиус экранирования 246
Томсоновское рассеяние 78, 83
Тонкая структура 565, 578
Точечные группы 45, 52—57, 517
Тригональная (ромбоэдрическая) структура 57—59, 62
Триклинная структура 57—59
Туннелирование *Зинера* 303, 508
— *Кронига — Пенни* модель 258
— между сверхпроводниками 304, 342, 343, 348
Туннельный диод 304

Углерод, атомный фактор рассеяния 84
Угловой момент, электрон 527, 528, 534, 556
— — ядро 556
Урбаха правило 456

Фазовая скорость, колебания решетки 125, 140
— — электроны 287
Фактор спектроскопического расщепления 527, 557
— спинового вырождения 388
Фаулера — Нордгейма формула 238
Ферми — Дирака интегралы 216
— статистика 210
— функция распределения 213
Фермиевское контактное взаимодействие 572, 575
Ферми поверхность 228, 295, 296
— — алюминий 285
— — германий 452
— — кремний 292
— — медь 295
— — мышьяк 463
— — сечение 322, 461
— сфера 227, 228
— энергия (уровень) 214, 316, 588
— — температурная зависимость 371, 395
Ферримагнетизм 533, 551—555
Ферримагнитный резонанс 567—569
Ферриты 554
Ферромагнетизм 533, 538—546
Ферромагнитные материалы, таблица 539
Ферромагнитный резонанс 567—569
Флокэ теорема 250
Фонон 79, 120
— длина волны 166, 167
— комбинационные полосы 151
— продольный акустический (*LA*) 410
— числа заполнения 156
Фононный каскад 474
Формфактор. *См.* Атомный фактор рассеяния
Фотопроводимость 476
Фторид лития, теплопроводность 174, 175

Халькогенидные стекла 91, 92, 438
Хартри — Фока методы 84, 249
Хилла уравнение 250
Химический (электрохимический) потенциал 216
Хлорид натрия, решетка 72
— цезия 72
Холла подвижность 206, 414, 417, 418
— постоянная (коэффициент) 206, 228, 293
— угол 418
— эффект 206, 417
Холл-фактор 417
Хоустона модель 304, 305
Хунда правило 528

Центрированная прямоугольная решетка 52
Центры окраски 104
Циклические граничные условия. *См. Борна—Кармана* граничные условия
Циклотронный резонанс 307, 308, 459
Цинк, электропроводность 188

Частота перескоков 587

Шоттки разупорядочение 104
— эмиссия 237

Шредингера уравнение 211, 249, 299, 313, 381
Штарка эффект 305
Шубникова — де Гааза эффект 323, 419, 460
Щель (зона) запрещенная 255, 260
— собственная 370

Эвальда метод 30, 111
— построение 100
Эвьена метод 30, 111
Эйнштейна коэффициенты 479
— модель 156—158
— соотношение 485
— функция 157
Экситоны 455, 481
Экстремальная орбита 308, 312
Электрон, магнитный момент 221
— подвижность 199
— скорость групповая 288, 289
— — фазовая 288
— спин 557
— среднее время свободного пробега 196, 229
— средняя длина свободного пробега 196, 199
Электронная поляризуемость 513
— теплопроводность 199
Электронно-дырочные пары 369—375
Электронные состояния, плотность 212
Электронный газ 186, 191, 192
— — вырожденный 218—221, 222, 228, 364
— — классический 195—209, 365
— — невырожденный 222, 246, 365
— — плазменные колебания 244, 245
Электронный парамагнитный резонанс (ЭПР) 557—560
Электронный спиновый резонанс. *См.* ЭПР
Электропроводность 198
Электрострикция 517—519
Элементарная ячейка 45, 46
Эллипсоидальные поверхности энергии 292, 451, 452
Эмиссия автоэлектронная 236, 237
— во внешнем поле 236
— вторичная 241—244
— многофононная 474
— термоэлектронная 231—235
— *Шоттки* 237
— фотоэлектрическая (фотоэмиссия) 239
Энергетические зоны, методы расчета 278—287
Энергия нулевых колебаний 156
— связи 17, 28, 111
Эффективная масса 225, 290—293
— — плотности состояний 300, 453
— — тензор 290, 291, 294
Эффективное поле 485

Ядерные магнитные моменты, таблица 572
Ядерный магнетон 528, 556, 581
— резонанс квадрупольный 590—592
— — магнитный (ЯМР) 581—583
— — — в металлах 587
Ямки травления 109, 110
Ячейка *Вигнера — Зейтца* 278
— примитивная 46, 270
— элементарная 45
— k-пространство

Оглавление

Предисловие редакторов перевода 5
Предисловие 8
Общая литература 11

Глава 1. Кристаллическая структура и форма твердых тел 13

1.1. Типы межатомных связей 17
1.2. Операции симметрии 42
1.3. Структуры реальных кристаллов 60
1.4. Дифракция в кристаллах 75
1.5. Обратное пространство 95
1.6. Дефекты кристаллической решетки 102
Задачи 111
Литература 114

Глава 2. Динамика решетки 117

2.1. Упругие волны, смещения атомов и фононы 118
2.2. Колебательные моды одноатомной решетки 122
2.3. Колебательный спектр решетки с базисом 137
2.4. Статистика фононов и теплоемкость решетки 154
2.5. Теплопроводность 167
Задачи 181
Литература 184

Глава 3. Электроны в металлах 186

3.1. Некоторые особенности металлического состояния 187
3.2. Классическая теория свободных электронов 194
3.3. Квантовая теория свободных электронов 210
3.4 Зонная теория твердых тел 247
3.5. Динамика движения электронов 287
3.6. Сверхпроводимость 325
Задачи 349
Литература 355

Глава 4. Полупроводники 357

4.1. Равновесная статистика электронов 359
4.2. Явления переноса в полупроводнике 404
4.3. Зонная структура реальных полупроводников 445
4.4. Явления, обусловленные избыточными носителями 464

Задачи 487
Литература 493

Глава 5. Диэлектрические и магнитные свойства твердых тел 496

5.1. Диэлектрические свойства 497
5.2. Магнитные свойства твердых тел 526
5.3. Магнитный резонанс 555
Задачи 592
Литература 596

Приложение. Значения некоторых физических констант . . 598

Предметный указатель 599

Учебное издание

Джон Блейкмор

ФИЗИКА ТВЕРДОГО ТЕЛА

Заведующий редакцией проф. А. Н. Матвеев. Зам. зав. редакцией С. М. Жебровский.
Научн. редакторы А. Н. Куксенко, И. Г. Нахимсон, М. Ф. Путов, В. И. Самсонова.
Мл. научн. редакторы Г. Г. Сорокина, Р. Х. Зацепина.
Художник В. Б. Прищепа. Художественный редактор К. В. Радченко.
Технический редактор Е. С. Потапенкова. Корректор Н. Л. Гиря.

ИБ № 6471

Сдано в набор 1.04.88. Подписано к печати 21.10.88. Формат 60×90¹/₁₆. Бумага типографская № 1. Печать высокая. Гарнитура литературная. Объем 19,00 бум. л. Усл. печ. л. 38,0. Усл. кр.-отт. 38,00. Уч.-изд. л. 37,74. Изд. № 2/5399. Тираж 15 200 экз. Зак. № 916. Цена 3 руб.

Издательство «МИР» В/О «Совэкспорткнига» Государственного комитета СССР по делам издательств, полиграфии и книжной торговли. 129820, ГСП, Москва, И-110, 1-й Рижский пер., 2.

Ленинградская типография № 4 ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» им. Евгении Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 191126, Ленинград, Социалистическая ул., 14.