

# ОГЛАВЛЕНИЕ

Предисловие . . . . .	6
<b>Глава 1. Симметрия кристаллов . . . . .</b>	<b>8</b>
1.1. Трансляционная симметрия кристалла . . . . .	8
1.2. Точечная симметрия кристаллов . . . . .	11
1.2.1. Элементы и преобразования точечной симметрии (11).	
1.2.2. Группы: определение (13). 1.2.3. Примеры групп симметрии (14). 1.2.4. Группа симметрии куба и группа симметрии тетраэдра (15).	
1.3. Пространственная симметрия кристалла . . . . .	16
1.4. Макроскопическая симметрия кристаллов. Кристаллический класс	18
1.5. Классификация кристаллов по симметрии . . . . .	18
1.6. Преобразование координат при точечных преобразованиях симметрии . . . . .	20
1.7. Классификация физических величин. Физические тензоры в кристаллах . . . . .	22
1.8. Примеры физических законов . . . . .	23
1.8.1. Закон Ома (23). 1.8.2. Пьезоэлектричество (25). 1.8.3. Эффект Сасаки (27). 1.8.4. Закон Гука. Тензор упругих модулей (28).	
Задачи. . . . .	29
<b>Глава 2. Колебания, волны и разложения в ряды Фурье . . . . .</b>	<b>31</b>
2.1. Введение . . . . .	31
2.2. Граничные условия Борна–Кармана. . . . .	32
2.3. Гармонические плоские волны . . . . .	35
2.4. Плотность состояний . . . . .	36
2.5. Фазовая и групповая скорость волны . . . . .	38
2.6. Обратная решетка . . . . .	39
2.7. Рассеяние рентгеновских лучей кристаллом . . . . .	43
2.8. Зона Бриллюэна . . . . .	46
Задачи. . . . .	50
<b>Глава 3. Колебания кристаллической решетки . . . . .</b>	<b>52</b>
3.1. Введение . . . . .	52
3.2. Одномерная цепочка с одним атомом в примитивной ячейке . . . . .	53
3.3. Одномерная цепочка с двумя атомами в примитивной ячейке . . . . .	57
3.4. Трехмерный кристалл. . . . .	63
3.5. Фононы . . . . .	67
3.6. Энергия колебаний и теплоемкость кристаллической решетки . . . . .	69
3.6.1. Модель Эйнштейна (70). 3.6.2. Модель Дебая (71).	
3.7. Экспериментальные методы исследования закона дисперсии фононов . . . . .	75
Приложение. Операторы рождения и уничтожения фононов . . . . .	81

Задачи. . . . .	87
<b>Глава 4. Электроны в идеальном кристалле . . . . .</b>	<b>90</b>
4.1. Введение . . . . .	90
4.2. Теорема Блоха. . . . .	91
4.3. Теорема о скорости и ускорении. . . . .	96
4.4. Зонная структура полупроводников. . . . .	102
4.4.1. Приближение сильной связи (103). 4.4.2. Приближение слабой связи (110).	
4.5. Металлы, изоляторы и полупроводники. . . . .	117
Задачи. . . . .	123
<b>Глава 5. Статистика электронов и дырок в полупроводниках . . . . .</b>	<b>125</b>
5.1. Собственный полупроводник . . . . .	125
5.2. Примесный полупроводник . . . . .	132
5.3. Статистика электронов и дырок в примесном полупроводнике. . . . .	136
5.4. Сильно легированные полупроводники . . . . .	145
Задачи. . . . .	150
<b>Глава 6. Кинетические явления в полупроводниках . . . . .</b>	<b>155</b>
6.1. Проводимость полупроводников . . . . .	155
6.2. Время свободного пробега (время релаксации) электронов в кристалле . . . . .	164
6.3. Рассеяние электронов на примесях . . . . .	166
6.4. Рассеяние на тепловых колебаниях решетки. . . . .	175
6.5. Рассеяние на акустических фононах . . . . .	177
6.6. Рассеяние на оптических фононах . . . . .	186
6.7. Температурная зависимость подвижности. . . . .	189
6.8. Горячие электроны . . . . .	192
Задачи. . . . .	200
<b>Глава 7. Поведение избыточных (неравновесных) носителей заряда в полупроводниках . . . . .</b>	<b>203</b>
7.1. Введение . . . . .	203
7.2. Нейтрализация пространственного заряда . . . . .	203
7.3. Дрейф неравновесных носителей заряда . . . . .	206
7.4. Расплывание пакета . . . . .	208
7.5. Эффект Дембера . . . . .	210
7.6. Рекомбинация избыточных носителей в полупроводниках. . . . .	211
7.7. Нелинейная теория рекомбинации . . . . .	212
7.7.1. Прямая рекомбинация электрона и дырки (бимолекулярная рекомбинация) (212). 7.7.2. Мономолекулярная рекомбинация. Статистика рекомбинации Шокли–Рида (214).	
7.8. Дрейфовая длина и диффузионная длина. . . . .	219

7.9. Неустойчивость тока в полупроводниках . . . . .	222
7.9.1. <i>N</i> -образная характеристика. Эффект Ганна (222).	
7.9.2. <i>S</i> -образная характеристика. Тепловой пробой. Шнурование тока (226).	
Задачи. . . . .	228
<b>Глава 8. Оптические свойства полупроводников . . . . .</b>	<b>229</b>
8.1. Введение. Прямые и непрямые оптические переходы . . . . .	229
8.2. Квантовомеханический подход к задаче о поглощении света . . . . .	232
8.3. Поглощение света свободными носителями заряда. . . . .	247
8.4. Непрямые межзонные переходы . . . . .	250
8.5. Индуцированное (вынужденное) излучение. Формулы Планка и Эйнштейна . . . . .	254
8.5.1. Усиление света в полупроводниках. Лазер на гомопереходе (257). 8.5.2. Лазер на гетеропереходе (261). 8.5.3. Лазер на квантовых ямах (263).	
8.6. Влияние магнитного поля на межзонное поглощение света . . . . .	264
8.7. Влияние электрического поля на край фундаментального поглощения света. Эффект Келдыша–Франца . . . . .	273
Задачи. . . . .	280
<b>Глава 9. Кинетические явления в полупроводниках в магнитном поле . . . . .</b>	<b>281</b>
9.1. Циклотронный (диаманитный) резонанс . . . . .	281
9.1.1. Классическое рассмотрение (281). 9.1.2. Квантовая теория циклотронного резонанса. Мазер на циклотронном резонансе (291).	
9.2. Эффект Холла и магнетосопротивление . . . . .	293
9.2.1. Классический эффект Холла (293). 9.2.2. Квантовый эффект Холла (300). 9.2.3. Магнетосопротивление. Отрицательное магнетосопротивление (313).	
9.3. Кинетическое уравнение . . . . .	317
Задачи. . . . .	324

## Предисловие

Эта книга основана на базе курса лекций, читаемых в Санкт-Петербургском электротехническом университете (ЛЭТИ) им. Ульянова-Ленина на базовой кафедре оптоэлектроники при Физико-техническом институте им. А. Ф. Иоффе РАН. Один из авторов книги академик В. И. Перель читал эти лекции на протяжении 30 лет. В настоящее время Г. Г. Зегря продолжает чтение этих лекций по рекомендации Владимира Иделевича. В. И. Перель был одним из создателей базовой кафедры оптоэлектроники в ЛЭТИ, вырастил целую плеяду талантливых учеников. Без преувеличения можно сказать, что Владимир Иделевич является одним из создателей современной физики полупроводников. Его пионерские работы в области квантовой кинетики, спиновых эффектов и рекомбинационных явлений в полупроводниках получили широкую известность и пользуются признанием во всем мире.

Книга рассчитана на студентов университетов и технических вузов, обучающихся по физическим специальностям, а также на аспирантов, научных работников и преподавателей. Для понимания книги требуется знакомство с основами квантовой механики и статистической физики. Никаких специальных сведений о полупроводниках, выходящих за рамки курса общей физики, не требуется. Особенностью книги является то, что в ней все формулы выводятся достаточно подробно для того, чтобы сделать ее доступной широкому кругу читателей.

В книге рассматриваются следующие темы: симметрия кристаллов и колебания кристаллической решетки; зонная структура полупроводников; кинетические явления в полупроводниках; оптические свойства полупроводников. Изложена современная концепция квантового эффекта Холла, одного из удивительных явлений физики полупроводников XX века. Большое внимание уделено сопоставлению физических моделей с экспериментальными результатами.

Поскольку книга является учебным пособием, в ней почти отсутствуют ссылки на оригинальные работы. В соответствующих частях книги даны ссылки на рекомендуемые монографии и учебные пособия, где читатель может найти более детальное изложение конкретного вопроса.

Мы выражаем свою искреннюю признательность М. М. Глазову, тщательно прочитавшему всю рукопись, за замечания, способствовавшие улучшению книги. Мы благодарны Р. А. Сурису, А. В. Горбатьюку, А. П. Дмитриеву, В. Ю. Кочаровскому за обсуждение ряда вопросов, В. Б. Корсакову за помощь в написание первых трех глав, А. А. Грешнову за помощь в написании § 9.2. Особую благодарность выражаем Е. А. Ефремовой за подготовку рукописи к изданию.

Авторы также выражают благодарность А. Н. Пихтину и Д. Р. Хохлову за ценные замечания и советы, высказанные при рецензировании рукописи.

Выражаем искреннюю благодарность заведующему кафедрой Оптоэлектроники В. И. Кучинскому за поддержку при написании книги.

## Глава 1

## СИММЕТРИЯ КРИСТАЛЛОВ

## 1.1. Трансляционная симметрия кристалла

Наиболее важной датой в истории физики твердого тела считается 1912 г., когда Макс фон Лауэ<sup>1)</sup> доказал, исследуя кристаллы посредством дифракции рентгеновских лучей, что они состоят из периодических рядов атомов.

Именно упорядоченное расположение атомов соответствует термодинамически равновесному состоянию твердого тела, то есть кристалл — это равновесное состояние системы атомов при достаточно низких температурах. Не все твердые тела кристаллические: например, при быстром охлаждении расплава можно получить «застывшую жидкость» — аморфное твердое тело, в котором атомы расположены беспорядочно. Примером такого твердого тела может служить обычное стекло.

Однако аморфное состояние не является термодинамически равновесным, поэтому рано или поздно аморфное тело, предоставленное самому себе, перейдет в равновесное состояние — кристаллизуется. Другое дело, что время кристаллизации может быть сколь угодно большим. Например, при нормальных условиях стекло превращается в кристалл за  $\sim 100$  лет.

Под упорядоченным расположением атомов в кристалле подразумевается его пространственная периодичность. Иными словами, для каждого кристалла можно выбрать три вектора  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$ , не лежащих в одной плоскости, таких, что при смещении кристалла как целого на любой из этих векторов, он совмещается сам с собой.

Операция перемещения кристалла на вектор  $\mathbf{A}_{n_1 n_2 n_3} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$ , где  $n_1$ ,  $n_2$ ,  $n_3$  — целые числа, называется *трансляцией*. Говорят, что кристалл обладает *трансляционной симметрией*, так как при трансляции на вектор  $\mathbf{A}_{n_1 n_2 n_3}$  он совмещается сам с собой.

Ясно, что векторы  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$  можно выбирать различными способами (см. рис. 1.1). Параллелепипед, образованный этими векторами, называется *элементарной ячейкой*. Элементарная ячейка минималь-

ного объема называется *примитивной* ячейкой, а вектора  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$ , на которых она построена, *примитивными* или *основными* векторами трансляций. В дальнейшем  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$  обозначают именно основные вектора трансляций. Векторы  $\mathbf{A}_{n_1 n_2 n_3}$  мы будем называть векторами прямой решетки.

Таким образом, примитивная ячейка является частным случаем элементарной ячейки. Основные вектора также можно выбрать различными способами. На рис. 1.1  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  и  $\mathbf{a}'_1$ ,  $\mathbf{a}'_2$  — две возможных пары основных векторов, а  $\mathbf{a}'_1$ ,  $\mathbf{a}'_2$  — неосновные вектора трансляций.

Посредством соответствующих операций трансляций все пространство кристаллической структуры заполняется примитивными ячейками. Примитивная ячейка — это периодически повторяющаяся в пространстве часть кристаллической решетки, имеющая форму параллелепипеда, с каждой точкой которой связана совокупность атомов. Совокупность атомов примитивной ячейки называется *базисом*, базис повторяется в пространстве и образует кристаллическую структуру.

Пусть кристалл имеет примитивную ячейку, образованную основными векторами трансляций  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$ . Возьмем точку в пространстве и с помощью трансляций этой точки на вектора  $\mathbf{A}_{n_1 n_2 n_3}$ , где  $n_1$ ,  $n_2$ ,  $n_3$  пробегает все целые числа, получим состоящую из точек пространственную решетку. Легко заметить, что все точки (узлы) такой решетки эквивалентны, то есть имеют одинаковое окружение; иными словами, из каждого узла видна одна и та же картина решетки расположения узлов. Такие решетки называются *решетками Браве*. Решетка Браве полностью определяет трансляционную симметрию кристалла.

Основные вектора трансляций должны удовлетворять следующему условию: построенные на них вектора прямой решетки, начинающиеся в некотором узле решетки Браве, должны кончаться на *всех* остальных узлах решетки Браве, то есть среди векторов прямой решетки  $\mathbf{A}_{n_1 n_2 n_3}$  найдется такой, который соединяет два любых заданных узла решетки Браве.

Если базис кристаллической решетки состоит из одного атома, иными словами, на примитивную ячейку приходится один атом, то кристаллическая решетка называется *простой*. В этом случае все атомы кристалла располагаются по узлам одной решетки Браве.

Если невозможно выбрать примитивную ячейку таким образом, чтобы ей соответствовал лишь один атом, то есть базис состоит из нескольких атомов, то решетка называется сложной. В этом случае каждому атому базиса соответствует своя подрешетка однотипных атомов, идентичная решетке Браве кристалла. Пример двумерной сложной решетки изображен на рис. 1.1. Белые и черные атомы могут быть химически идентичны, но по положению в кристаллической решетке они разные. Атомы кристалла *однотипны*, если они химически идентичны и с каждого из них видна одна и та же картина кристаллической решетки.

<sup>1)</sup> Макс фон Лауэ (1879–1960).

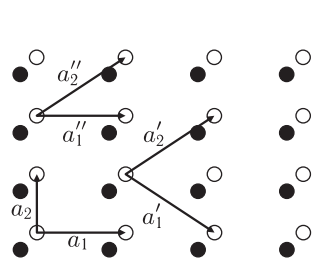


Рис. 1.1. Двумерный кристалл. Кружки разных цветов означают различные типы атомов

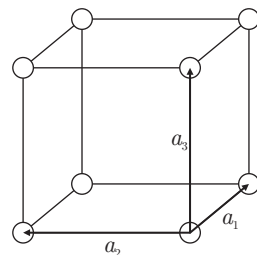


Рис. 1.2. Простая кубическая решетка  $|a_1| = |a_2| = |a_3|$

Таким образом, увидеть решетку Браве можно, если смотреть только на однотипные атомы. Кристалл можно представить себе двумя способами: взять базис и транслировать его многократно с помощью примитивных векторов трансляций или взять несколько в точности одинаковых решеток Браве и вставить их друг в друга. Видно, что двумерный кристалл на рис. 1.1 состоит из двух вставленных друг в друга решеток Браве.

Для кубических кристаллов существуют 3 типа решеток Браве:

1. **Простая кубическая решетка** (в природе очень редко встречается, рис. 1.2).
2. **Кубическая гранецентрированная решетка, ГЦК**. В центре каждой грани находится еще по одному атому (рис. 1.3).
3. **Объемноцентрированная кубическая решетка, ОЦК**. Дополнительный атом помещен в центр куба (рис. 1.4).

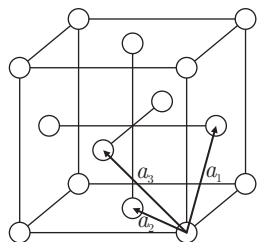


Рис. 1.3. Кубическая гранецентрированная решетка

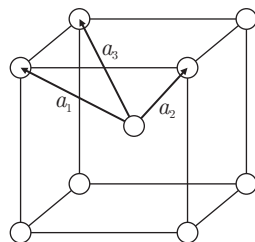


Рис. 1.4. Кубическая объемноцентрированная решетка

Итак, *решетка Браве отражает трансляционную структуру (симметрию) кристалла*, а не его *кристаллическое* строение. Кроме трансляционной, кристаллы обладают точечной симметрией: симметрией по отношению к поворотам и отражениям. Для определения точечной симметрии необходимо наряду с типом решетки Браве учитывать строение соединения (базис примитивной ячейки).

Сделаем замечание по поводу терминологии. Иногда под элементарной ячейкой понимают примитивную ячейку с базисом. Часть решетки Браве, содержащую несколько примитивных ячеек и отражающую точечную симметрию кристалла называют *условной* или *кристаллографической элементарной ячейкой*. На рис. 1.2–1.4 изображены именно такие ячейки.

## 1.2. Точечная симметрия кристаллов

Точечное преобразование в теории симметрии — это преобразование, которое оставляет в покое, то есть неподвижной, хотя бы одну точку фигуры. Если при некотором точечном преобразовании фигура переходит сама в себя, то говорят, что она симметрична относительно этого преобразования (обладает соответствующим элементом симметрии: осью вращения, плоскостью отражения и т. д.).

### 1.2.1. Элементы и преобразования точечной симметрии

**1. Ось симметрии.** Преобразование: поворот вокруг оси на определенный угол  $\alpha = 2\pi/n$ , соответствующая ось симметрии обозначается как  $C_n$ ,  $n$  называется порядком оси симметрии.

На рис. 1.5 изображены фигуры, обладающие осями симметрии второго  $C_2$  и третьего  $C_3$  порядков. Ясно, что если фигура симметрична относительно поворота на угол  $2\pi/n$ , то она будет переходить в себя и при преобразованиях  $C_n^k$  — поворотах на угол  $k \cdot 2\pi/n$ , где  $1 \leq k \leq n$  — целое число. Очевидно, что  $C_n^n$  — это тождественное преобразование.

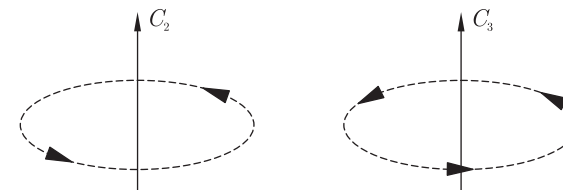


Рис. 1.5. Фигуры, обладающие осями симметрии второго  $C_2$  и третьего  $C_3$  порядков

Можно показать, что кристаллическая решетка (фигура, обладающая трансляционной симметрией) может иметь только оси симметрии порядков 2, 3, 4, 6<sup>1)</sup>.

**2. Плоскость симметрии.** Преобразование: отражение относительно плоскости, этот элемент симметрии обозначается через  $\sigma$ , см. рис. 1.6.

Если у фигуры есть и плоскость симметрии и ось симметрии, то:

<sup>1)</sup> Ансельм А. И. Введение в теорию полупроводников. — М.: Наука, 1978.

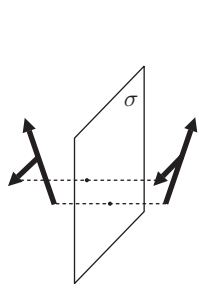


Рис. 1.6. Плоскость симметрии

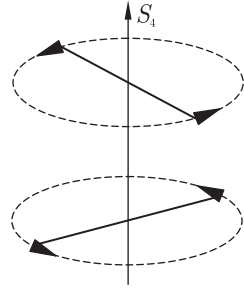


Рис. 1.7. Зеркально-поворотная ось четвертого порядка

- если плоскость симметрии проходит через ось симметрии, то эта плоскость симметрии обозначается через  $\sigma_v$ ;
- если плоскость симметрии перпендикулярна оси симметрии, этот элемент будем обозначать через  $\sigma_h$ .

Легко видеть, что других возможностей для взаимного расположения оси вращения и плоскости отражения нет.

**3. Инверсия относительно некоторой точки.** Обозначается буквой  $i$ , соответствующая точка называется центром инверсии.

Преобразование инверсии заключается в одновременном изменении знака всех координат (в координатной системе, центром которой является данная точка), то есть изменении направления всех осей на обратное:  $\{x, y, z\} \rightarrow \{-x, -y, -z\}$ . При этом правовинтовая система координат переходит в левовинтовую, и наоборот. О теле, симметричном относительно этого преобразования, говорят, что оно обладает центром симметрии (инверсии).

**4. Зеркально-поворотная ось порядка  $n$ .** Преобразование: последовательный поворот вокруг оси на угол  $2\pi/n$  и отражение в плоскости, перпендикулярной этой оси. Можно показать, что при нечетном  $n$  зеркально-поворотная ось сводится к оси симметрии  $C_n$  и перпендикулярной к ней плоскости симметрии  $\sigma_h$ .

Фигура, обладающая зеркально-поворотной осью 4-го порядка изображена на рис. 1.7.

Преобразование  $S_n$  можно представить в виде:

$$S_n = \sigma_h \cdot C_n. \quad (1.1)$$

Знак произведения в этом выражении означает последовательность действий (справа налево): сначала второй множитель ( $C_n$ ), потом первый ( $\sigma_h$ ), то есть сначала выполняется поворот, а после этого — отражение. Очевидно, что  $S_n^{2k} = C_n^{2k}$ ; кроме того,  $\sigma_h C_2 = C_2 \sigma_h = S_2 = i$ . Отсюда следует, что  $i \sigma_h = C_2$  и  $i C_2 = \sigma_h$ . Таким образом, три элемента симметрии  $C_2$ ,  $\sigma_h$  и  $i$  взаимно связаны, так что наличие двух из них автоматически приводит к наличию третьего.

Чем больше у фигуры возможных преобразований симметрии, тем она симметричнее. Отметим, что симметрия кристалла не может быть выше, чем симметрия решетки Браве, но может быть ниже.

Пусть у нас имеются две фигуры. Для каждой из этих фигур можно найти набор (группу) всех преобразований симметрии. Если эти наборы преобразований одинаковы, то говорят, что у фигур одинаковая симметрия.

**1.2.2. Группы: определение.** Пусть имеется множество элементов  $\{A, B, C, D \dots\}$  и это множество обладает следующими свойствами:

1. Для элементов множества определено понятие произведения (композиции), то есть любым двум элементам  $A$  и  $B$  ставится в соответствие третий элемент из этого же множества:  $A \cdot B = C$ . Этот элемент  $C$  называют произведением  $A$  и  $B$ . Результат «умножения», вообще говоря, зависит от порядка, то есть не обладает коммутативностью.

2. Среди элементов группы есть один и только один элемент, который называется тождественным или единичным элементом  $E$  и обладает следующим свойством:  $A \cdot E = E \cdot A = A$  для любого  $A$  (свойство единицы). Другими словами, постулируется существование единицы (единичного элемента) в этом множестве.

3. Для любого элемента  $A$  найдется в этом же множестве обратный элемент, который мы обозначим через  $A^{-1}$  и для которого выполняется равенство:  $A \cdot A^{-1} = A^{-1} \cdot A = E$ . Это условие существования обратного элемента.

4. «Умножение» должно быть ассоциативным, то есть

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Множество целых чисел не образует группы относительно операции умножения чисел — нет обратного элемента  $A^{-1}$ . Множество действительных чисел также не образует группу относительно арифметического умножения, так как нет обратного элемента для нуля.

Действительные и целые числа образуют группы по отношению к операции сложения, в этом случае  $E = 0$ ,  $A^{-1} = -A$ .

**Теорема.** Множество всех преобразований симметрии определенной фигуры, то есть множество преобразований, при которых фигура переходит сама в себя, образует группу. Композицией («произведением») двух элементов этой группы является последовательное применение соответствующих преобразований симметрии: произведение  $A \cdot B$  означает, что сначала было выполнено преобразование  $B$ , а затем преобразование  $A$ .

**Доказательство.** Свойство полноты следует из того, что множество содержит *все* преобразования симметрии фигуры. Единичный элемент — отсутствие преобразования (тождественное преобразование, обозначается через  $E$ ); обратный элемент — *обратное преобразова-*

ние. Например, для  $C_n$  это  $C_n^{n-1}$ ; для  $\sigma$  это  $\sigma^{-1}$ . Можно доказать и ассоциативность умножения.

**1.2.3. Примеры групп симметрии.** Группа симметрии равностороннего треугольника. Эта группа называется  $C_{3v}$ :

Равносторонний треугольник — симметричная фигура. Теория групп позволяет охарактеризовать симметрию этой фигуры — определить все преобразования симметрии, то есть найти группу преобразований симметрии фигуры. Если не выводить треугольник из плоскости, то он будет обладать следующими элементами симметрии (рис. 1.8, слева):

1. Ось симметрии третьего порядка (перпендикулярно плоскости, через центр треугольника): треугольник вращается в своей плоскости по часовой стрелке на углы  $2\pi/3$  (преобразование  $C_3$ ) и  $4\pi/3$  (преобразование  $C_3^2$ ).
2. Три плоскости симметрии, эти плоскости проходят через ось симметрии третьего порядка, поэтому их обозначают  $\sigma_v$ <sup>1)</sup>.
3. Тожественное преобразование.

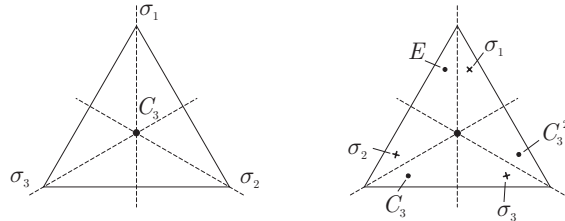


Рис. 1.8. Группа симметрии равностороннего треугольника

Итак, имеем:  $C_3$ ,  $C_3^2$ ,  $\sigma_1$ ,  $\sigma_2$ ,  $\sigma_3$ ,  $E$ . Составим таблицу умножения этой группы.

Чтобы составить таблицу умножения для группы преобразований симметрии равностороннего треугольника, возьмем точку, не лежащую на каком-либо элементе симметрии, и применим к ней последовательно все преобразования и произведения преобразований. На рис. 1.8 (на треугольнике справа) исходная точка обозначена как  $E$  — тождественное преобразование, остальные точки — результат поворотов исходной вокруг оси  $C_3$ , крестики — результат отражений от плоскости симметрии.

Используя этот рисунок, легко составить таблицу умножения. Например, нужно найти произведение  $\sigma_3\sigma_2$ . Точка  $\sigma_2$  уже есть на рисунке. Ее нужно подвергнуть отражению  $\sigma_3$ . В результате попадаем в точку  $C_3^2$ . Таким образом,  $\sigma_3\sigma_2 = C_3^2$ .

<sup>1)</sup> Отметим, что операцию отражения в плоскости нельзя представить в виде произведения поворотов.

Таблица 1.1. Таблица умножения группы симметрии равностороннего треугольника

$C_{3v}$	$E$	$C_3$	$C_3^2$	$\sigma_1$	$\sigma_2$	$\sigma_3$
$E$	$E$	$C_3$	$C_3^2$	$\sigma_1$	$\sigma_2$	$\sigma_3$
$C_3$	$C_3$	$C_3^2$	$E$	$\sigma_2$	$\sigma_3$	$\sigma_1$
$C_3^2$	$C_3^2$	$E$	$C_3$	$\sigma_3$	$\sigma_1$	$\sigma_2$
$\sigma_1$	$\sigma_1$	$\sigma_3$	$\sigma_2$	$E$	$C_3^2$	$C_3$
$\sigma_2$	$\sigma_2$	$\sigma_1$	$\sigma_3$	$C_3$	$E$	$C_3^2$
$\sigma_3$	$\sigma_3$	$\sigma_2$	$\sigma_1$	$C_3^2$	$C_3$	$E$

Действуя подобным образом, получим таблицу умножения для группы симметрии равностороннего треугольника  $C_{3v}$  (табл. 1.1; так как умножение некоммукативно, то необходимо определить порядок умножения: сначала производится преобразование из левого столбца, потом — из верхней строки). В верхнем левом углу табл. 1.1 стоит название группы.

Как уже говорилось, произведение двух элементов в группе преобразований некоммукативно, то есть в общем случае  $A \cdot B \neq B \cdot A$ . Это свойство произведения можно продемонстрировать на элементах группы  $C_{3v}$ , например:

$$C_3 \cdot \sigma_1 = \sigma_2, \quad (1.2)$$

$$\sigma_1 \cdot C_3 = \sigma_3. \quad (1.3)$$

Видно, что эти два элемента группы не коммутируют.

**Подгруппа** — это подмножество элементов группы, которое само образует группу. У группы  $C_{3v}$  есть следующие подгруппы:

$$E, C_3, C_3^2 - (3); \quad E, \sigma_1 - (2); \quad E, \sigma_2 - (2); \quad E, \sigma_3 - (2); \quad E - (1).$$

**Порядок** подгруппы — это число ее элементов (см. цифры в скобках). Можно показать, что порядок подгруппы является делителем числа элементов в группе. В группе  $C_{3v}$  содержатся 6 элементов, соответственно ее подгруппы имеют порядки 1, 2 и 3.

С математической точки зрения группа полностью определена таблицей умножения.

Две группы  $G$  и  $G'$  называются **изоморфными**, если между их элементами существует взаимно однозначное соответствие, причем для любых  $A, B \in G$ ,  $A', B' \in G'$ , таких, что  $A \leftrightarrow A'$ ,  $B \leftrightarrow B'$  выполняется  $A \cdot B \leftrightarrow A' \cdot B'$  (символ  $\leftrightarrow$  обозначает соответствие). Иными словами, группы изоморфны, если можно так переименовать элементы одной из этих групп, чтобы ее таблица умножения совпала с таблицей умножения второй группы.

**1.2.4. Группа симметрии куба и группа симметрии тетраэдра.** Это группы, с которыми мы в дальнейшем будем работать чаще всего.



**Группа симметрии куба.** Группа осей куба  $O$  содержит 24 элемента и состоит из преобразований поворота вокруг всех осей симметрии:  $E$ ;  $3C_4$ ;  $3C_4^2$ ;  $3C_2^3$ ;  $4C_3$ ;  $4C_3^2$ ;  $6C_2$  (рис. 1.9).

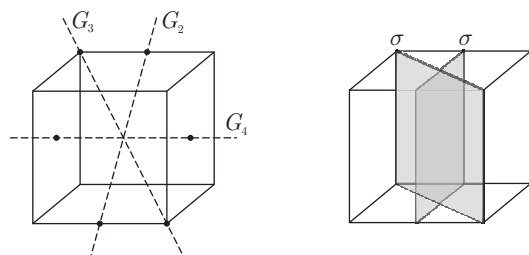


Рис. 1.9. Оси симметрии куба и плоскости симметрии куба. Показаны примеры различных элементов симметрии

Кроме того, у куба есть плоскости симметрии и центр инверсии  $i$ . Если операцию инверсии  $i$  проводить совместно с 24 поворотами, то получится еще 24 преобразования. Итого получается 48 элементов, которые исчерпывают всю группу куба (в них войдут все плоскости симметрии). Эта группа из 48 элементов называется группой  $O_h$ . Симметрия  $O_h$  определяет кристаллический класс объемных материалов кремния и германия (см. ниже).

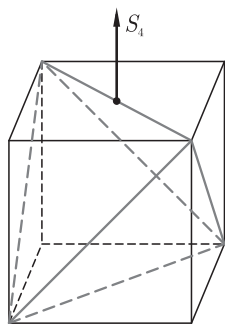


Рис. 1.10. Тетраэдр, вписанный в куб

**Группа симметрии тетраэдра.** Группа тетраэдра является подгруппой группы куба, то есть все преобразования симметрии тетраэдра являются подмножеством преобразований симметрии куба. Симметрию тетраэдра имеют кристаллы GaAs, InSb. На рис. 1.10 изображен тетраэдр, вписанный в куб. Видно, что у тетраэдра нет центра инверсии.

В группу тетраэдра входят преобразования  $E$ ,  $3C_4^2$ ,  $4C_3$ ,  $4C_3^2$  (12 обычных поворотов), 6 зеркальных поворотов:  $3S_4$ ,  $3S_4^3$  и 6 отражений в плоскости:  $6\sigma$  (плоскости симметрии проходят через ребра тетраэдра). Группа тетраэдра обозначается через  $T_d$ , содержит 24 преобразования.

### 1.3. Пространственная симметрия кристалла

До сих пор мы рассматривали два типа преобразований симметрии: трансляции и точечные преобразования.

Симметрия относительно трансляций — это главное свойство кристалла. Ее полностью отражает решетка Браве.

Симметрия решетки Браве относительно точечных преобразований называется **сингонией** или **кристаллической системой**. Всего имеется 7 групп симметрии, описывающих различные сингонии, то есть 7 разных групп, каждая из которых включает в себя все точечные преобразования симметрии решеток Браве определенного типа.

Кристалл характеризуется не только решеткой Браве, которая полностью определяет его трансляционную симметрию, но и базисом: типом и взаимным расположением атомов, составляющих примитивную ячейку (строением примитивной ячейки). Симметрия простого кристалла, у которого каждому узлу решетки Браве соответствует один атом, совпадает с симметрией решетки Браве, то есть с ее сингонией. Что касается сложного кристалла, то очевидно, что его точечная симметрия не может быть выше чем симметрия его решетки Браве: однако в общем случае не все точечные преобразования симметрии, переводящие в себя решетку Браве, переводят в себя кристаллическую структуру (пример будет приведен ниже). Таким образом, в общем случае группа точечной симметрии кристалла является подгруппой группы симметрии его решетки Браве (сингонии). Таких подгрупп, то есть групп точечной симметрии, совместимых с трансляционной симметрией, насчитывается 32.

Например, группа симметрии куба  $O_h$  (кубическая сингония) имеет пять подгрупп, совместимых с трансляцией, в том числе три из рассмотренных нами групп:  $O_h$ ,  $O$ ,  $T_d$ .

Не нужно думать, что точечные преобразования и трансляции исчерпывают всю совокупность преобразований симметрии кристаллической решетки (хотя они исчерпывают преобразования симметрии решетки Браве).

Кристаллическая решетка может обладать дополнительными элементами симметрии:

1. Винтовые оси  $n$ -го порядка. Соответствующие преобразования заключаются в поворотах на углы  $2\pi n$  вокруг винтовой оси, сопровождаемых переносами на расстояния  $ra/n$  вдоль той же оси. Здесь  $p = 1, 2, \dots, n - 1$ ;  $a$  — постоянная решетки вдоль той же оси.

2. Плоскости зеркального скольжения. Соответствующее преобразование заключается в отражении от этой плоскости, сопровождаемом трансляцией вдоль оси, лежащей в той же плоскости, на расстояние  $a/2$ .

Группа пространственной симметрии кристалла содержит все преобразования симметрии кристаллической решетки: трансляции, точечные преобразования, преобразования, связанные с винтовыми осями и плоскостями зеркального скольжения (можно показать, что других преобразований симметрии кристаллической решетки нет). Все пространственные группы симметрии были найдены Е. С. Федоровым в конце XIX века. Их оказалось 230.



## 1.4. Макроскопическая симметрия кристаллов.

### Кристаллический класс

С макроскопической точки зрения кристалл характеризуется симметрией направлений. Именно от направления электрического поля в общем случае зависит усредненная плотность тока в кристалле, от направления силы, действующей на кристалл зависит деформация и т.д. Для характеристики макроскопической симметрии кристаллов наряду с группами пространственной симметрии (в которые входят преобразования, переводящие все узлы кристаллической решетки в узлы того же типа) можно ввести группы *симметрии направлений*, которые переводят только направления в кристалле в эквивалентные направления. Поскольку параллельные переносы (трансляции) вообще не меняют направлений, в группы симметрии направлений не входят трансляции, а винтовые оси и плоскости зеркального скольжения заменяются простыми поворотными осями и плоскостями отражения.

Таким образом, группы симметрии направлений — это группы точечной симметрии, совместимые с трансляционной симметрией кристалла. Как уже говорилось, таких групп насчитывается 32 (все подгруппы всех сингоний).

Каждая группа симметрии направлений характеризует *кристаллический класс*, то есть существуют 32 кристаллических класса.

## 1.5. Классификация кристаллов по симметрии

Наиболее полной характеристикой симметрии кристалла является группа его пространственной симметрии (одна из 230 групп). Однако, с макроскопической точки зрения кристалл может быть охарактеризован более грубо.

**Сингония.** Сингония или *кристаллическая система* — это точечная группа симметрии решетки Браве. Как уже говорилось, существует всего 7 сингоний, то есть 7 различных групп, каждая из которых включает в себя все точечные преобразования симметрии решеток Браве определенного типа.

Наиболее высоко симметричной является кубическая сингония — группа  $O_h$ , к этой сингонии относятся все виды кубических решеток: простая (рис. 1.2), ГЦК (рис. 1.3) и ОЦК (рис. 1.4). Наименее симметрична триклинная сингония, которая имеет всего два элемента симметрии: единичный элемент  $E$  и инверсию  $i$ . Центром инверсии обладают все решетки Браве, его существование непосредственно следует из трансляционной симметрии решетки.

**Тип решетки Браве.** Каждой сингонии соответствует один или несколько типов решеток Браве. Например, триклинной сингонии соответствует только один тип решетки, а кубической — три типа: простая

кубическая, ГЦК и ОЦК. В общей сложности выделяют 14 типов решеток Браве.

**Кристаллический класс.** Кристаллический класс — группа симметрии направлений кристалла. Эта характеристика симметрии определяет вид тензорных свойств кристалла (см. ниже). Всего кристаллических классов 32.

Обычно, если симметрию кристалла характеризуют одним словом, то это слово означает сингонию. Например, кубическими называют кристаллы с сингонией  $O_h$  независимо от типа решетки Браве и кристаллического класса.

### Примеры кристаллических структур

**Кристалл NaCl.** Структура кристалла NaCl изображена на рис. 1.11, ионы натрия и хлора представлены шарами черного и белого цветов. Нетрудно понять, что этот кристалл имеет гранецентрированную кубическую решетку Браве, его примитивная ячейка (базис) состоит из двух атомов: атома хлора и атома натрия. Кристалл принадлежит кубической сингонии  $O_h$  и кристаллическому классу  $O_h$ , то есть обладает полной симметрией куба.

**Структура алмаза.** Такой структурой обладают кристаллы алмаза, кремния и германия (рис. 1.12).

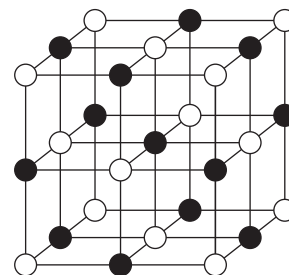


Рис. 1.11. Кристалл NaCl

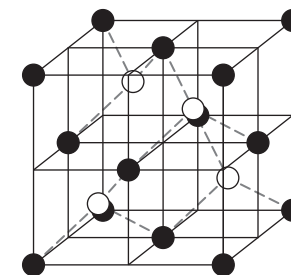


Рис. 1.12. Решетка алмаза. Круги разных цветов показывают атомы двух типов

Примитивная ячейка этих кристаллов состоит из двух химически одинаковых атомов, то есть кристалл содержит атомы двух типов (на рисунке изображены разными цветами). Каждый атом находится в центре правильного тетраэдра, в вершинах которого расположены атомы другого типа, эти четыре атома являются его ближайшими соседями. На рис. 1.12 ближайшие соседи соединены пунктирными линиями. Тетраэдры, окружающие атомы разных типов, ориентированы по-разному.

Структуре алмаза соответствует гранецентрированная кубическая решетка Браве, то есть атомы каждого типа образуют ГЦК-подре-

сетку. Подрешетки атомов разных типов сдвинуты друг относительно друга на четверть пространственной диагонали гранецентрированного куба.

Таким образом, так же как и структура NaCl, структура алмаза принадлежит сингонии  $O_h$  и кристаллическому классу  $O_h$ .

**Структура цинковой обманки.** В структуре алмаза атомы разных типов, то есть атомы, принадлежащие двум разным решеткам Браве, химически идентичны. Если атомы двух разных типов химически различны, то мы получаем структуру цинковой обманки ZnS. Такая структура отражает симметрию полупроводниковых соединений  $A_3B_5$ <sup>1)</sup>, например GaAs.

Структура цинковой обманки относится к кубической сингонии  $O_h$  и кристаллическому классу  $T_d$  — группе симметрии тетраэдра. Тип решетки Браве — ГЦК.

### 1.6. Преобразование координат при точечных преобразованиях симметрии

До сих пор мы изображали элементы и преобразования симметрии геометрически. Между тем, любое преобразование симметрии однозначно описывается алгебраически: законом изменения координат при этом преобразовании.

Пусть точка в пространстве задается радиус-вектором  $\mathbf{r}$ . В выбранной системе координат положение этой точки определяется тремя координатами  $(x, y, z)$ , которые являются проекциями радиус-вектора  $\mathbf{r}$  на координатные оси.

Рассмотрим точечные преобразования. Как известно, они оставляют неподвижной хотя бы одну точку объекта, пусть это будет начало координат.

В новой системе координат точка (радиус-вектор  $\mathbf{r}$ ) будет задаваться координатами  $(x', y', z')$ . Координаты  $(x', y', z')$  связаны с координатами  $(x, y, z)$  линейным преобразованием:

$$x' = c_{xx}x + c_{xy}y + c_{xz}z, \quad (1.4)$$

$$y' = c_{yx}x + c_{yy}y + c_{yz}z, \quad (1.5)$$

$$z' = c_{zx}x + c_{zy}y + c_{zz}z, \quad (1.6)$$

или, если использовать матричные обозначения:

$$\mathbf{r}' = \hat{\mathbf{C}} \cdot \mathbf{r},$$

где

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad \mathbf{r}' = \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \quad \hat{\mathbf{C}} = \begin{pmatrix} c_{xx} & c_{xy} & c_{xz} \\ c_{yx} & c_{yy} & c_{yz} \\ c_{zx} & c_{zy} & c_{zz} \end{pmatrix},$$

<sup>1)</sup> Здесь цифры «3» и «5» обозначают соединения элементов третьей и пятой групп периодической системы Менделеева.

$\hat{\mathbf{C}}$  — матрица перехода от одной системы координат к другой, ее столбцы представляют собой координаты единичных векторов базиса старой системы координат в новой системе координат.

Приведем примеры матриц преобразования координат для некоторых точечных преобразований симметрии:

Инверсия ( $i$ ):

$$\hat{\mathbf{C}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Поворот  $C_4$  вокруг оси  $z$ :

$$\hat{\mathbf{C}} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Отражение от плоскости, проходящей через начало координат перпендикулярно оси  $x$  ( $\sigma_x$ ):

$$\hat{\mathbf{C}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Поворот  $C_3$  вокруг оси  $z$ :

$$\hat{\mathbf{C}} = \begin{pmatrix} -1/2 & -\sqrt{3}/2 & 0 \\ \sqrt{3}/2 & -1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Пусть мы теперь последовательно совершаем два точечных преобразования, первому из них соответствует матрица преобразования координат  $\hat{\mathbf{C}}_1$ , второму —  $\hat{\mathbf{C}}_2$ . Если вначале точка имела координаты  $\mathbf{r}$ , то после первого преобразования она будет иметь координаты  $\mathbf{r}' = \hat{\mathbf{C}}_1 \cdot \mathbf{r}$ , а после второго — координаты  $\mathbf{r}''$ :

$$\mathbf{r}'' = \hat{\mathbf{C}}_2 \cdot \mathbf{r}' = \hat{\mathbf{C}}_2 \cdot (\hat{\mathbf{C}}_1 \cdot \mathbf{r}) = (\hat{\mathbf{C}}_2 \cdot \hat{\mathbf{C}}_1) \cdot \mathbf{r}.$$

Таким образом, произведению преобразований симметрии соответствует произведение соответствующих матриц преобразования координат, то есть преобразования симметрии и соответствующие им матрицы перемножаются по одной таблице умножения. Поэтому группа преобразований точечной симметрии изоморфна группе соответствующих матриц преобразования координат (см. определение изоморфизма групп). Говорят, что группа матриц преобразования координат является *представлением* группы преобразований симметрии.

Рассмотрим преобразования групп куба ( $O_h$ ) и тетраэдра ( $T_d$ ). Если координатные оси направлены вдоль ребер куба, то можно показать, что точечным преобразованиями группы  $O_h$  соответствуют все перестановки координат с любыми изменениями знаков, например:

$$(x, y, z) \rightarrow (y, z, x) \text{ — поворот } C_3,$$

$$(x, y, z) \rightarrow (-x, -y, -z) \text{ — инверсия},$$

$$(x, y, z) \rightarrow (y, -x, z) \text{ — поворот } C_4 \text{ вокруг оси } z, \text{ и т. п.}$$

Всего таких преобразований насчитывается  $48 = 3! \cdot 2^3$ , где  $3!$  — число возможных перестановок,  $2^3$  — число возможных расстановок знаков; то есть столько, сколько элементов симметрии в группе куба.

Преобразования группы тетраэдра  $T_d$ , которая является подгруппой группы  $O_h$ , соответствуют лишь те из перестановок координат группы куба, при которых меняется четное количество знаков: 0 или 2. Число таких перестановок конечно же равно 24 — количеству элементов группы тетраэдра.

### 1.7. Классификация физических величин. Физические тензоры в кристаллах

Для количественного описания физических явлений используются величины различных типов: они ведут себя различным образом при преобразовании координат.

Простейшим случаем является скалярная величина — величина, которая не меняется при преобразованиях системы координат. Примерами скалярных величин могут служить масса, заряд, плотность заряда, электростатический потенциал. Значения этих величин не зависят от того, какой системой координат мы пользуемся.

Многие физические величины характеризуются не только одним численным значением, но и направлением в пространстве. Примерами являются сила, скорость, напряженность электрического поля. Эти величины являются векторными. Чтобы однозначно определить векторную величину, например силу  $\mathbf{F}$ , мы должны задать ее компоненты — проекции силы на координатные оси:  $F_x, F_y, F_z$ . Очевидно, что значения компонент векторной величины зависят от выбора системы координат и меняются определенным образом при преобразованиях системы координат.

**Вектор** (тензор первого ранга)  $\mathbf{A}$  — величина, которая в каждой системе координат определяется тремя числами, называемыми компонентами вектора,  $(A_x, A_y, A_z)$ , причем при переходе от одной системы координат к другой компоненты вектора преобразуются как координаты точки (радиус-вектор точки  $\mathbf{r}$  — частный случай вектора с компонентами  $(x, y, z)$ ). Это означает, что если  $\hat{C}$  — матрица преобразования координат, то

$$\mathbf{A}' = \hat{C}\mathbf{A},$$

или

$$A'_\alpha = \sum_{\alpha''} c_{\alpha\alpha''} A_{\alpha''},$$

где индексы  $\alpha, \alpha''$  принимают значения  $x, y, z$ ;  $c_{\alpha\alpha''}$  — компоненты матрицы преобразования,  $\mathbf{A}, \mathbf{A}'$  — столбцы координат вектора в старой и новой системах координат.

Во многих случаях для описания физических величин удобно использовать более сложные величины — тензоры второго и более высокого рангов.

Тензором  $n$ -го порядка называется физическая величина, характеризующаяся элементами  $A_{\alpha_1, \alpha_2, \dots, \alpha_n}$  (индексы  $\alpha_1, \alpha_2, \dots, \alpha_n$  принимают значения  $x, y, z$ ), которые при преобразованиях координат преобразуются по закону:

$$A'_{\alpha'_1, \alpha'_2, \dots, \alpha'_n} = \sum_{\alpha'_1, \alpha'_2, \dots, \alpha'_n} c_{\alpha'_1 \alpha_1} c_{\alpha'_2 \alpha_2} \dots c_{\alpha'_n \alpha_n} A_{\alpha_1, \alpha_2, \dots, \alpha_n},$$

то есть как произведения координат. Примеры тензоров высших рангов даны ниже.

**Тензорная форма физических законов.** Целый ряд физических законов может быть записан в виде зависимости некоторой величины  $B$  — отклика — от другой величины  $A$  — воздействия. Если воздействие мало, то эта зависимость линейная (первый член ряда Тейлора) и может быть записана в символическом виде

$$B = DA.$$

Величина  $D$  является свойством кристалла и внешних параметров, но не зависит от воздействия.

Вообще говоря, воздействие и отклик описываются тензорами и общая линейная связь между ними может быть представлена в координатном виде:

$$B_{\beta_1, \beta_2, \dots, \beta_m} = \sum_{\alpha_1, \alpha_2, \dots, \alpha_n} D_{\beta_1, \beta_2, \dots, \beta_m, \alpha_1, \alpha_2, \dots, \alpha_n} A_{\alpha_1, \alpha_2, \dots, \alpha_n}. \quad (1.7)$$

При изменении системы координат, зависимость (1.7) записывается в виде:

$$B'_{\beta'_1, \beta'_2, \dots, \beta'_m} = \sum_{\alpha_1, \alpha_2, \dots, \alpha_n} D'_{\beta'_1, \beta'_2, \dots, \beta'_m, \alpha_1, \alpha_2, \dots, \alpha_n} A'_{\alpha_1, \alpha_2, \dots, \alpha_n}. \quad (1.8)$$

Можно показать, что для того, чтобы две формулы (1.7) и (1.8) описывали одну и ту же зависимость отклика от воздействия необходимо и достаточно, чтобы свойство  $D_{\beta_1, \beta_2, \dots, \beta_m, \alpha_1, \alpha_2, \dots, \alpha_n}$  преобразовывалось бы при преобразованиях координат как компоненты тензора ранга  $n + m$ , то есть как произведение координат. Симметрия направлений кристалла (кристаллический класс) определяет структуру тензоров, то есть линейные связи между компонентами, описывающих их свойства.

### 1.8. Примеры физических законов

**1.8.1. Закон Ома.** В качестве примера рассмотрим зависимость усредненной по элементарной ячейке плотности тока  $\mathbf{j}$  в кристалле от приложенного электрического поля  $\mathbf{E}$ . Если функция  $\mathbf{j}(\mathbf{E})$  является

достаточно гладкой, мы можем разложить ее в ряд Тейлора в окрестности точки  $\mathbf{E} = 0$  и для слабых полей ограничиться линейным членом разложения. Тогда в определенной системе координат эта зависимость будет иметь вид:

$$\begin{aligned} j_x &= \sigma_{xx}E_x + \sigma_{xy}E_y + \sigma_{xz}E_z, \\ j_y &= \sigma_{yx}E_x + \sigma_{yy}E_y + \sigma_{yz}E_z, \\ j_z &= \sigma_{zx}E_x + \sigma_{zy}E_y + \sigma_{zz}E_z, \end{aligned}$$

или, в матричной форме:

$$\mathbf{j} = \hat{\sigma} \cdot \mathbf{E},$$

где

$$\mathbf{j} = \begin{pmatrix} j_x \\ j_y \\ j_z \end{pmatrix}, \quad \mathbf{E} = \begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}, \quad \hat{\sigma} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix}.$$

При этом  $\mathbf{E}$  — воздействие,  $\mathbf{j}$  — отклик,  $\hat{\sigma}$  — свойство кристалла.

В новой системе координат закон Ома будет иметь вид:

$$\mathbf{j}' = \hat{\sigma}' \cdot \mathbf{E}'.$$

В соответствии с общими правилами  $\hat{\sigma}$  является тензором второго ранга, то есть в преобразовании  $\hat{\sigma} \rightarrow \hat{\sigma}'$  компоненты  $\hat{\sigma}'$  должны быть связаны с компонентами  $\hat{\sigma}$  так:

$$\sigma'_{\alpha\beta} = \sum_{\alpha'\beta'} c_{\alpha\alpha'} c_{\beta\beta'} \sigma_{\alpha'\beta'},$$

где  $c_{\alpha\beta}$  — компоненты матрицы преобразования координат. Иными словами, компоненты тензора  $\sigma_{\alpha\beta}$ , как и компоненты любого тензора, преобразуются так же, как произведение его индексов.

Это правило позволяет легко преобразовывать компоненты тензора при простых точечных преобразованиях. Например, при отражении в плоскости, перпендикулярной оси  $x$ , координаты преобразуются следующим образом:  $(x, y, z) \rightarrow (-x, y, z)$ . Преобразуя компоненты тензора  $\sigma_{\alpha\beta}$  как произведения координатполучим:

$$\begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \rightarrow \begin{pmatrix} \sigma_{xx} & -\sigma_{xy} & -\sigma_{xz} \\ -\sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ -\sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix}.$$

Симметрия кристалла накладывает определенные ограничения на вид любого тензора, описывающего свойства кристалла (чего, конечно, нельзя сказать о тензорах, описывающих внешнее воздействие и результат этого воздействия, например о тензоре деформации: деформировать кристалл можно как угодно).

Именно, при преобразованиях симметрии точечной группы направления в кристалле переходят в эквивалентные, тогда и тензор, характеризующий свойства этого кристалла, не должен менять своего вида при

этих преобразованиях симметрии (то есть его компоненты не должны меняться).

Докажем, что для кристалла класса  $O_h$  любой тензор второго ранга, описывающий некоторое свойство этого кристалла, сводится к одной (скалярной) величине.

Как уже говорилось, группа симметрии куба состоит из всех возможных перестановок координат со всеми возможными расстановками знаков. Посмотрим, как преобразуются компоненты тензора второго ранга при некоторых из этих преобразований.

Возьмем преобразование, при котором координаты  $x$  и  $y$  меняются местами:

$$(x, y, z) \rightarrow (y, x, z): \quad xx \rightarrow yy, \quad \sigma_{xx} \rightarrow \sigma_{yy}.$$

Так как тензор кристалла класса  $O_h$  не должен меняться при всех преобразованиях симметрии группы куба, то получаем, что  $\sigma_{xx} = \sigma_{yy}$ . Аналогично доказывается, что  $\sigma_{xx} = \sigma_{zz}$  и  $\sigma_{yy} = \sigma_{zz}$ .

Рассмотрим теперь другое преобразование группы куба:

$$(x, y, z) \rightarrow (-x, y, z): \quad xy \rightarrow -xy, \quad \sigma_{xy} \rightarrow -\sigma_{xy}.$$

Отсюда следует, что  $\sigma_{xy} = 0$ , таким же образом доказывается, что равны нулю все недиагональные компоненты тензора.

Итак, недиагональные компоненты любого тензора второго ранга, характеризующего свойства кристалла класса  $O_h$ , равны нулю, а диагональные равны между собой, то есть тензор превращается в скаляр. В частности, для закона Ома получаем:

$$j_x = \sigma E_x, \quad j_y = \sigma E_y, \quad j_z = \sigma E_z,$$

или в векторной форме:

$$\mathbf{j} = \sigma \mathbf{E}.$$

Таким образом, кристалл класса  $O_h$  ведет себя по отношению к проводимости как изотропная среда. Это утверждение справедливо для всех свойств кристалла класса  $O_h$ , описываемых тензорами второго ранга, например, диэлектрической проницаемости. Напомним, что к классу  $O_h$  принадлежат кристаллы Si и Ge.

Группа симметрии тетраэдра  $T_d$  не включает в себя преобразование  $(x, y, z) \rightarrow (-x, y, z)$ , которым мы пользовались при доказательстве равенства нулю недиагональных компонент тензора второго ранга. Однако вместо этого преобразования можно использовать преобразование  $(x, y, z) \rightarrow (-x, y, -z)$ , входящее в группу  $T_d$ , которое также переводит  $\sigma_{xy}$  в  $-\sigma_{xy}$ . Поэтому и для кристаллов класса  $T_d$  тензор второго ранга всегда сводится к скалярной величине, в частности, таким свойством обладает структура цинковой обманки (например, кристалл GaAs).

Таким образом, любой тензор второго ранга, описывающий свойства кристалла класса  $T_d$  или  $O_h$ , сводится к скаляру.

**1.8.2. Пьезоэлектричество.** Прямой пьезоэлектрический эффект называется электрическая поляризация кристалла (возникновение дипольного момента) при его деформации. Обратный пьезоэлектрический эффект — деформация кристалла при помещении его в электрическое поле.

Любое физическое явление включает воздействие и отклик. В случае прямого пьезоэлектрического эффекта в роли воздействия выступает деформация, в роли отклика — дипольный момент.

При деформации кристалла каждая его точка смещается на вектор  $u = \mathbf{u}(\mathbf{r})$ . Ясно, что деформация характеризуется не величиной смещения каждой точки, а разницей в смещениях соседних точек, то есть производными компонент  $u$  по координатам.

Деформацию удобно описывать тензором деформации  $\varepsilon_{\alpha\beta}$ , который определяется следующим образом:

$$\varepsilon_{\alpha\alpha} = \frac{\partial u_\alpha}{\partial r_\alpha}, \quad \varepsilon_{\alpha\beta} = \frac{1}{2} \left( \frac{\partial u_\alpha}{\partial r_\beta} + \frac{\partial u_\beta}{\partial r_\alpha} \right).$$

Этот тензор по определению является симметричным:  $\varepsilon_{\alpha\beta} = \varepsilon_{\beta\alpha}$ . При таком определении недиагональных компонент в тензор деформации не попадают смещения точек кристалла, при которых  $\frac{\partial u_x}{\partial y} = -\frac{\partial u_y}{\partial x}$  и т.п., такие перемещения соответствуют вращению кристалла как целого, а не деформации.

Поляризацию кристалла (дипольный момент единицы объема) будем обозначать через  $\mathbf{P}$ .

При слабых деформациях мы можем разложить функцию  $\mathbf{P}(\varepsilon)$  в ряд Тейлора и ограничиться первым членом разложения, линейным по деформации:

$$P_\alpha = \sum_{\delta\nu} Q_{\alpha\delta\nu} \varepsilon_{\delta\nu}.$$

Тензор третьего ранга  $Q_{\alpha\delta\nu}$  характеризует пьезоэлектрические свойства кристалла.

Для любого кристалла, обладающего центром инверсии, в частности, для кубического кристалла, все компоненты тензора  $Q$  равны нулю. Действительно, при операции инверсии  $(x, y, z) \rightarrow (-x, -y, -z)$  любая компонента тензора третьего ранга меняет знак (так как меняет знак произведение любых трех координат):  $Q_{\alpha\delta\nu} \rightarrow -Q_{\alpha\delta\nu}$ . Но если кристалл имеет центр инверсии, то он не должен менять своих свойств при операции инверсии откуда следует  $Q_{\alpha\delta\nu} = 0$ .

Рассмотрим теперь кристаллический класс  $T_d$ . Как уже говорилось, если координатные оси направлены вдоль ребер куба, то преобразования симметрии группы  $T_d$  соответствуют все возможные перестановки координат со всеми возможными изменениями четного количества знаков. Отсюда следует, что при преобразованиях группы  $T_d$  знаки могут менять все компоненты тензора  $Q$ , кроме компонент, у которых

все три индекса разные:  $Q_{xyz}, Q_{xzy}, Q_{yxz}, Q_{yzx}, Q_{zxy}, Q_{zyx}$ . Таким образом, только эти 6 компонент отличны от нуля и, так как в группу  $T_d$  входят все перестановки координат, равны между собой. Пьезоэлектрические свойства кристалла класса  $T_d$  характеризуются одной константой  $Q = Q_{xyz} = Q_{xzy}$ .

Например, Ge и Si не являются пьезоэлектриками, а GaAs является, и его пьезоэлектрические свойства характеризуются одной константой.

**1.8.3. Эффект Сасаки.** Кубический кристалл в линейном приближении проводит электрический ток как изотропная среда. Однако в достаточно сильных полях закон Ома нарушается: зависимость тока от поля оказывается нелинейной. Чтобы описать слабую нелинейность, мы можем оставить в разложении плотности тока по электрическому полю, кроме линейного вклада, следующий ненулевой член разложения. Так как любой тензор третьего ранга, описывающий свойства кристалла класса  $O_h$ , равен нулю, то следующим ненулевым членом разложения будет член третьего порядка:

$$j_\alpha = \sigma E_\alpha + \sum_{\beta\gamma\delta} \gamma_{\alpha\beta\gamma\delta} E_\beta E_\gamma E_\delta.$$

С учетом этого члена ток в общем случае уже не будет направлен вдоль поля даже для кристаллов класса  $O_h$ . Этот эффект называется эффектом Сасаки и проявляется в достаточно сильных полях.

Исследуем теперь свойства тензора четвертого ранга  $\gamma_{\alpha\beta\gamma\delta}$ , вытекающие из кубической симметрии кристалла.

Преобразования координат группы  $O_h$  соответствуют все возможные перестановки координат со всеми возможными изменениями знаков. Поэтому все компоненты тензора, содержащие нечетное количество одинаковых индексов (например,  $\gamma_{xyzx}, \gamma_{xxzz}$ ), равны нулю. Это утверждение верно для любого тензора четвертого ранга, описывающего физические свойства кристалла класса  $O_h$ .

Нетрудно также показать, что оно верно и для кристалла класса  $T_d$ . Действительно, если какая-либо координата встречается среди индексов компоненты тензора четвертого ранга нечетное число раз, то по крайней мере одна из двух других координат встретится четное число раз или не встретится вообще. А это значит, что найдется преобразование группы тетраэдра (изменение знаков двух координат), при котором такая компонента меняет знак, откуда следует, что она должна быть равна нулю.

Итак, отличны от нуля четыре набора компонент тензора, содержащие четные количества одинаковых индексов:

$$\gamma_{\alpha\alpha\alpha\alpha}, \quad \gamma_{\alpha\alpha\beta\beta}, \quad \gamma_{\alpha\beta\alpha\beta}, \quad \gamma_{\alpha\beta\beta\alpha}.$$

Так как группы  $O_h$  и  $T_d$  включают преобразования, которые соответствуют любым перестановкам координат, элементы каждого набора равны между собой.

В первый набор входят три равных друг другу компонента:  $\gamma_{xxxx} = \gamma_{yyyy} = \gamma_{zzzz} = \Gamma_{11}$ .

Остальные наборы содержат 18 компонент, по 6 в каждом. Однако в сумме, определяющей плотность тока, компоненты  $\gamma_{\alpha\beta\gamma\delta}$ , отличающиеся перестановкой последних трех индексов, умножаются на произведения одинаковых компонент **E**, например:  $\gamma_{xyyx}E_yE_yE_x$  и  $\gamma_{xyxy}E_yE_xE_y$ . Поэтому тензор  $\gamma_{\alpha\beta\gamma\delta}$  должен быть симметричен по последним трем индексам. Отсюда следует, что все 18 компонент типа  $\gamma_{\alpha\alpha\beta\beta}$ ,  $\gamma_{\alpha\beta\alpha\beta}$ ,  $\gamma_{\alpha\beta\beta\alpha}$  равны. Обозначим эту величину через  $\Gamma_{12}/3$ .

В итоге мы получаем следующие выражения для компонент плотности тока кристалла класса  $O_h$ :

$$\begin{aligned} j_x &= \sigma E_x + \Gamma_{11}E_x^3 + \Gamma_{12}E_x(E_y^2 + E_z^2), \\ j_y &= \sigma E_y + \Gamma_{11}E_y^3 + \Gamma_{12}E_y(E_x^2 + E_z^2), \\ j_z &= \sigma E_z + \Gamma_{11}E_z^3 + \Gamma_{12}E_z(E_x^2 + E_y^2). \end{aligned}$$

Анизотропия проводимости характеризуется коэффициентами  $\Gamma_{11}$  и  $\Gamma_{12}$ .

Если поле приложено вдоль ребра куба (например, вдоль оси  $x$ ), то ток течет в направлении поля:

$$j_x = \sigma E + \Gamma_{11}E^3, \quad j_y = 0, \quad j_z = 0.$$

Ток также параллелен полю, если поле направлено вдоль диагонали куба:

$$j = \sigma E + \frac{\Gamma_{11} + 2\Gamma_{12}}{3}E^3.$$

В случае приложения поля в других направлениях направление тока отличается от направления поля. Это и есть эффект Сасаки.

**1.8.4. Закон Гука. Тензор упругих модулей.** При деформации кристалла в нем возникают напряжения. Эти напряжения описываются тензором деформации  $\sigma_{\alpha\beta}$ , который определяется следующим образом. В каждой точке кристалла можно провести плоскость, внешняя нормаль которой направлена по оси  $\beta$ ,  $\beta = (x, y, z)$ . В деформированном (напряженном) кристалле к каждой точке этой плоскости будет приложена сила, направление которой, вообще говоря, не совпадает с направлением нормали, то есть оси  $\beta$ . Компонента тензора напряжений  $\sigma_{\alpha\beta}$  равна проекции этой силы, приходящейся на единицу площади, на ось  $\alpha$ .

Например, на единичный элемент плоскости, перпендикулярной оси  $x$ , в общем случае действует нормальная сила  $\sigma_{xx}$  (давление или растяжение) и сдвиговые (тангенциальные) силы  $\sigma_{yx}$  и  $\sigma_{zx}$ , направленные вдоль осей  $y$  и  $z$  соответственно.

Можно показать, что тензор напряжений симметричен<sup>1)</sup>.

Ясно, что при деформации кристалла в нем возникают определенные напряжения, и наоборот, при создании напряжений кристалл определенным образом деформируется. Связь между напряжением и деформацией описывается тензором четвертого ранга  $C_{\alpha\beta\gamma\delta}$  — тензором упругих модулей кристалла:

$$\varepsilon_{\alpha\beta} = \sum_{\gamma\delta} C_{\alpha\beta\gamma\delta} \sigma_{\gamma\delta}.$$

Так как тензоры напряжения и деформации симметричны, то тензор упругих модулей симметричен по двум первым и двум последним индексам:  $C_{\alpha\beta\gamma\delta} = C_{\alpha\beta\delta\gamma}$ ;  $C_{\alpha\beta\gamma\delta} = C_{\beta\alpha\gamma\delta}$ .

Исследуем тензор упругих модулей кубического кристалла. При рассмотрении эффекта Сасаки мы пришли к выводу, что тензор четвертого ранга, описывающий свойства кристалла класса  $T_d$  или  $O_h$ , имеет четыре независимых набора компонент. Компоненты, принадлежащие одному набору, равны между собой:

$$C_{\alpha\alpha\alpha\alpha}, \quad C_{\alpha\alpha\beta\beta}, \quad C_{\alpha\beta\alpha\beta}, \quad C_{\alpha\beta\beta\alpha}.$$

Из-за симметрии тензора упругих модулей по двум первым и двум последним индексам компоненты последних двух наборов равны:  $C_{\alpha\beta\alpha\beta} = C_{\alpha\beta\beta\alpha}$ . Таким образом, упругие свойства кубических кристаллов (классы  $T_d$  и  $O_h$ ) характеризуются тремя независимыми постоянными. Обычно используют следующие обозначения:

$$\begin{aligned} C_{\alpha\alpha\alpha\alpha} &= C_{11}, \\ C_{\alpha\alpha\beta\beta} &= C_{12}, \\ C_{\alpha\beta\alpha\beta} &= C_{\alpha\beta\beta\alpha} = C_{44}. \end{aligned}$$

## Задачи

**Задача 1.1.** В каких средах возможен гальваномагнитный эффект (возникновение электрического тока в однородном магнитном поле), описываемый соотношением  $\mathbf{j} \propto \mathbf{B}$ .

*Решение.* Электрический ток является полярным вектором, а магнитное поле — аксиальным (или псевдо-) вектором. При пространственных поворотах полярный и аксиальный векторы преобразуются одинаково, а при пространственной инверсии — по-разному (полярный вектор меняет знак). Поэтому гальваномагнитный эффект допустим лишь в средах, группа симметрии которых не включает в себя про-

<sup>1)</sup> Ландау Л. Д., Лифшиц Е. М. Теория упругости. — М.: ФИЗМАТЛИТ, 2001.

странственную инверсию. Это условие является необходимым, но не достаточным.

**Задача 1.2.** Может ли неоднородное распределение температуры вызывать (а) электрический ток и (б) магнитное поле в кубических полупроводниках.

*Указание.* Неоднородное распределение температуры удобно характеризовать ее градиентом,  $\nabla T$ , который является полярным вектором.

**Задача 1.3.** Тензором какого ранга описываются магнитоупругий эффект (возникновение намагниченности среды при ее деформации).

*Указание.* Деформации характеризуются тензором второго ранга, а магнитное поле — псевдотензор первого ранга (аксиальный вектор).

**Задача 1.4.** Объясните кажущееся противоречие: электрический ток и электрическое поле связаны законом Ома ( $\mathbf{j} = \hat{\sigma} \mathbf{E}$ ), при этом электрическое поле инвариантно к инверсии времени, а электрический ток — меняет при инверсии времени знак.

*Ответ.* При инверсии времени меняют знаки также величины, описывающие диссипацию, в данном случае тензор проводимости  $\hat{\sigma}$ .

## Глава 2

# КОЛЕБАНИЯ, ВОЛНЫ И РАЗЛОЖЕНИЯ В РЯДЫ ФУРЬЕ

## 2.1. Введение

При описании физических явлений в твердых телах нам часто придется иметь дело с плоскими волнами, как вещественными, так и комплексными.

Вещественная одномерная плоская волна, распространяющаяся вдоль оси  $x$ , описывается функцией  $|A| \cos(kx + \varphi)$ , где  $|A|$  — амплитуда волны,  $k$  — волновое число,  $\varphi$  — фаза. Волновое число  $k$  определяет длину волны  $\lambda = 2\pi/k$ , а фаза  $\varphi$  — начальный сдвиг по оси  $x$ :  $kx + \varphi = k(x + \varphi/k)$ , то есть сдвиг равен  $+\varphi/k$ .

Комплексная одномерная плоская волна (например, волновая функция свободного электрона) описывается комплексной экспонентой:  $A \exp(ikx)$ , где  $A$  — комплексная амплитуда:  $A = |A| \exp(i\varphi)$ .

Функция  $|A| \cos(kx + \varphi)$  является вещественной частью комплексной экспоненты, поэтому и вещественные волны удобно описывать комплексной плоской волной, помня при этом, что физической величине соответствует вещественная часть функции.

Все сказанное легко обобщается на трехмерный случай (да и вообще на случай любой размерности). Вещественная плоская волна в пространстве имеет вид  $|A| \cos(\mathbf{k}\mathbf{r} + \varphi)$ , комплексная:  $A \exp(i\mathbf{k}\mathbf{r})$ . Волновой вектор  $\mathbf{k}$  задает направление распространения и длину волны  $\lambda = 2\pi/k$ ,  $k = |\mathbf{k}|$ .

Любая функция  $f(\mathbf{r})$ , удовлетворяющая весьма общим условиям, может быть представлена в виде наложения всех плоских волн с разными амплитудами, то есть разложена в интеграл Фурье:

$$f(\mathbf{r}) = \iiint dk_x dk_y dk_z F(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} = \int d^3k F(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}}, \quad (2.1)$$

где  $F(\mathbf{k})$  — фурье-образ функции  $f(\mathbf{r})$ .

Любая одномерная периодическая функция  $f(x)$  с периодом  $L$  может быть разложена в ряд Фурье:

$$f(x) = \sum_{k=2\pi n/L} F(k) e^{ikx}, \quad (2.2)$$



где  $n$  пробегает все целые числа. Это и понятно: в разложении периодической функции присутствуют только такие волны, для которых  $L$  также является периодом:  $e^{ikx} = e^{ik(x+L)}$ , что выполняется только при «разрешенных»  $k = 2\pi n/L$ . Иными словами, на длине  $L$  должно укладываться целое число длин волн:  $\lambda = L/n$ .

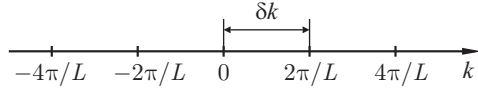


Рис. 2.1. Одномерная обратная решетка

Разрешенные значения  $k$  образуют равномерную решетку на оси  $k$  с интервалом  $\delta k = 2\pi/L$  между соседними значениями (рис. 2.1). В некотором, достаточно большом интервале волновых чисел  $\Delta k$  содержится  $\frac{\Delta k}{2\pi/L} = \frac{L\Delta k}{2\pi}$  разрешенных значений  $k$ .

## 2.2. Граничные условия Борна–Кармана

Иногда удобнее иметь дело с суммами, чем с интегралами. Однако функции, описывающие физические явления в бесконечном кристалле, часто не являются периодическими, то есть раскладываются в интеграл, а не в ряд Фурье.

Можно считать кристалл конечным (каким он на самом деле и является) и в соответствии с этим ставить для функций естественные граничные условия: например, требовать, чтобы функция обращалась в нуль на границах кристалла. В этом случае функцию можно разложить в ряд на отрезке или в ограниченном объеме (в кубе, например).

Но, во-первых, ограниченность достаточно большого кристалла не сказывается на физических явлениях в объеме, а корректно учесть краевые эффекты и поставить правильные граничные условия обычно непросто.

При этом кристалл конечного размера уже не обладает трансляционной и точечной симметрией бесконечного кристалла, а использование свойств симметрии часто сильно упрощает решение задачи.

Чтобы рассматривать бесконечный кристалл, но при этом сохранить возможность раскладывать функции в ряды и работать с дискретными величинами, используют искусственный прием: требуют, чтобы функция была периодической по каждой из координат с некоторым большим периодом  $L$ . Такие искусственные периодические граничные условия называются условиями Борна–Кармана.

Другими словами, мы делим весь кристалл на большие кубические области с ребром  $L$  и требуем, чтобы функция в каждом кубе вела себя одинаково. Такой куб называют нормировочным.

Если  $L$  достаточно велико, то такое искусственное ограничение не скажется на физических явлениях, происходящих внутри кристалла, так же как не сказывается на них влияние поверхности, если размеры кристалла велики. Поэтому окончательное решение задачи не должно зависеть от  $L$  при  $L \rightarrow \infty$ .

Итак, в соответствии с условиями Борна–Кармана, мы требуем, чтобы функция была периодической по каждой из координат. В трехмерном случае:

$$\begin{aligned} f(x+L, y, z) &= f(x, y, z), \\ f(x, y+L, z) &= f(x, y, z), \\ f(x, y, z+L) &= f(x, y, z). \end{aligned} \quad (2.3)$$

Плоская волна  $e^{i\mathbf{k}\mathbf{r}}$  удовлетворяет этим условиям только при «разрешенных» волновых векторах  $\mathbf{k}$ :

$$k_x = \frac{2\pi}{L}n_1, \quad k_y = \frac{2\pi}{L}n_2, \quad k_z = \frac{2\pi}{L}n_3, \quad (2.4)$$

где  $n_1, n_2, n_3$  — целые числа. В разложении периодической по каждой координате функции будут присутствовать только такие волны.

Разрешенные значения волнового вектора образуют в  $k$ -пространстве простую кубическую решетку, двумерный аналог которой изображен на рис. 2.2.

На каждое разрешенное значение волнового вектора в  $k$ -пространстве приходится элементарный объем  $\delta^3 k$ :

$$\delta^3 k = \left(\frac{2\pi}{L}\right)^3 = \frac{(2\pi)^3}{V}, \quad (2.5)$$

где  $V = L^3$  — объем нормировочного куба, то есть куба с ребром  $L$ . При  $V \rightarrow \infty$  элементарный объем  $\delta^3 k \rightarrow 0$ .

Нетрудно посчитать количество разрешенных значений волновых векторов, попадающих в достаточно большой объем  $k$ -пространства  $\Delta^3 k$ :

$$\frac{\Delta^3 k}{\delta^3 k} = \frac{V}{(2\pi)^3} \Delta^3 k. \quad (2.6)$$

Пространство волновых векторов называют *обратным пространством*.

При квантовом описании плоская волна описывает состояние частицы, разные волновые векторы соответствуют разным состояниям. Поэтому число разрешенных значений волнового вектора часто называют числом состояний.

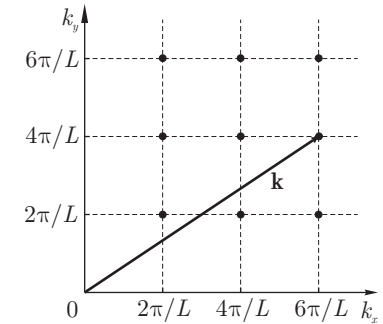


Рис. 2.2. Двумерная обратная решетка

Имеет место следующая

**Теорема:** Если куб достаточно велик, то число состояний, приходящихся на объем  $\Delta^3 k$ , не зависит от того, какие граничные условия поставлены на границах куба, а определяется лишь его объемом.

Правда, нужно сделать оговорку. Если мы, к примеру, зададим нулевые граничные условия на границах нормировочного куба, то разрешенными состояниями (волнами, по которым раскладывается в ряд функция) будут функции  $\sin k_x x \cdot \sin k_y y \cdot \sin k_z z$  с разрешенными компонентами волнового вектора  $k_x = \pi n_x/L$ ,  $k_y = \pi n_y/L$ ,  $k_z = \pi n_z/L$ . Как видим, интервал между разрешенными значениями  $k$  в два раза меньше, чем при периодических граничных условиях, то есть плотность разрешенных значений волновых векторов в  $k$ -пространстве в восемь раз выше. Однако, в этом случае можно считать, что все разрешенные значения компонент  $k_x$ ,  $k_y$ ,  $k_z$  положительны: волновой вектор  $(-k_x, k_y, k_z)$  описывает то же самое состояние, что и  $(k_x, k_y, k_z)$ . Поэтому здесь мы можем говорить только о количестве разрешенных состояний в объеме  $\Delta^3 k$ , расположенном в одном из восьми октантов  $k$ -пространства (с положительными значениями  $k_x$ ,  $k_y$ ,  $k_z$ ). В случае периодических граничных условий этому объему соответствуют восемь объемов во всех октантах  $k$ -пространства, таким образом, количество разрешенных состояний будет тем же.

Часто граничные условия и, как следствие, дискретные разрешенные состояния возникают непосредственно из условий задачи и имеют явный физический смысл. Например, рассмотренные выше нулевые граничные условия реализуются в кубическом резонаторе, вследствие чего в нем могут существовать волны лишь с разрешенными волновыми числами  $k_x = \pi n_x/L$ ,  $k_y = \pi n_y/L$ ,  $k_z = \pi n_z/L$ , где  $L$  — длина стороны резонатора (между стенками резонатора должно укладываться целое число полуволен).

Те же граничные условия для квантовой частицы в прямоугольной потенциальной яме приводят к тому, что частица может занимать лишь те состояния, для которых между стенками укладывается целое число полуволен Де Бройля, то есть ее импульс  $p = \hbar k$  может иметь только разрешенные дискретные значения.

В этих примерах нормировочный объем  $V$  уже не является абстрактным объемом, введенным для удобства, а соответствует объему резонатора или потенциальной ямы.

Покажем, как при  $V \rightarrow \infty$  осуществляется переход от суммирования по дискретным состояниям к интегрированию по  $\mathbf{k}$ -пространству.

Пусть есть сумма по разрешенным значениям  $k$  функции  $\varphi(\mathbf{k})$ . Домножим и разделим эту сумму на  $\delta^3 k = (2\pi)^3/V$  и будем считать, что функция  $\varphi(\mathbf{k})$  почти не меняется на интервале  $\delta^3 k$ .

$$\sum_{\mathbf{k}} \varphi(\mathbf{k}) = \frac{V}{(2\pi)^3} \sum_{\mathbf{k}} \varphi(\mathbf{k}) \delta^3 k. \quad (2.7)$$

Предел при  $\delta^3 k \rightarrow 0$  ( $V \rightarrow \infty$ ) суммы, стоящей в правой части равенства, и есть интеграл:

$$\lim_{\delta^3 k \rightarrow 0} \sum_{\mathbf{k}} \varphi(\mathbf{k}) \delta^3 k = \int \varphi(\mathbf{k}) d^3 k. \quad (2.8)$$

Отсюда следует правило преобразования суммы в интеграл:

$$\sum_{\mathbf{k}} \varphi(\mathbf{k}) \rightarrow V \int \frac{\varphi(\mathbf{k})}{(2\pi)^3} d^3 k, \quad \text{где } V \rightarrow \infty. \quad (2.9)$$

Итак, на каждое разрешенное состояние в обратном пространстве приходится объем  $\delta^3 k = (2\pi)^3/V$ .

В элементе объема обратного пространства  $\Delta^3 k$  содержится  $\frac{V}{(2\pi)^3} \Delta^3 k$  разрешенных состояний. Соответственно,  $\frac{\Delta^3 k}{(2\pi)^3}$  — число состояний в элементе  $\Delta^3 k$ , приходящихся на единицу объема прямого пространства <sup>1)</sup>.

## 2.3. Гармонические плоские волны

Гармонической плоской волной (комплексной) называется функция  $\Psi(\mathbf{r}, t) = A e^{i(\mathbf{k}\mathbf{r} - \omega t)}$ ,  $A = |A| \exp(i\varphi)$  по-прежнему обозначает комплексную амплитуду.

Вещественная гармоническая плоская волна  $|A| \cos(kx - \omega t + \varphi)$  является вещественной частью комплексной волны.

Частота  $\omega$  задает период колебаний:  $T = 2\pi/\omega$ . Частота гармонических плоских волн, описывающих различные физические процессы (будь то упругие колебания, электромагнитные волны или движение свободной квантовомеханической частицы), как правило, однозначно определяется волновым вектором:  $\omega = \omega(\mathbf{k})$ . Эта зависимость называется законом дисперсии.

**Волны и частицы (напоминание).** Из квантовой механики известно, что каждой частице соответствует волна, а каждой элементарной волне — частица:

Характеристики волны	Характеристики частицы	Связь
волновой вектор $\mathbf{k}$	импульс $\mathbf{p}$	$\mathbf{p} = \hbar \mathbf{k}$
частота $\omega$	энергия $\mathcal{E}$	$\mathcal{E} = \hbar \omega$

**Универсальные соотношения, не зависящие от типа волны-частицы:**

<sup>1)</sup> Здесь рассматриваются состояния без учета спина. Спин — это собственный момент элементарной частицы, имеющий квантовую природу и не связанный с ее движением в пространстве.

$$\begin{aligned} \text{Длина волны: } \lambda &= \frac{2\pi}{k} = \frac{2\pi\hbar}{p} \\ \text{Период: } T &= \frac{2\pi}{\omega} = \frac{2\pi\hbar}{\mathcal{E}} \end{aligned}$$

**Соотношение, зависящее от типа волны-частицы, — закон дисперсии.** Так как энергия от частоты, также как импульс от волнового вектора, отличается лишь постоянным множителем  $\hbar$ , то законом дисперсии часто называют не только зависимость  $\omega(\mathbf{k})$ , но и зависимости  $\mathcal{E}(\mathbf{p})$  или  $\mathcal{E}(\mathbf{k})$ .

Примеры законов дисперсии:  $\omega = ck$  — для фотона,  $\mathcal{E} = \frac{p^2}{2m}$  — для свободного электрона.

## 2.4. Плотность состояний

**Квазичастица** — это объект, который описывается гармонической плоской волной  $A \exp(i\mathbf{k}\mathbf{r} - \omega t)$ .

Каждая квазичастица обладает определенной энергией  $\mathcal{E} = \mathcal{E}(\mathbf{k})$ , а волна — соответствующей частотой,  $\omega = \mathcal{E}/\hbar$ . Зависимость  $\mathcal{E}(\mathbf{k})$  (или  $\omega(\mathbf{k})$ ), как уже говорилось, называется законом дисперсии, это самая важная характеристика квазичастицы. Иногда вместо зависимости  $\mathcal{E}(\mathbf{k})$  используют зависимость  $\mathcal{E}(\mathbf{p})$ , где  $\mathbf{p} = \hbar\mathbf{k}$  — импульс частицы. В простейших случаях энергия зависит только от модуля волнового вектора (свободные электроны, фотоны в вакууме), но в общем случае зависит и от направления (большинство квазичастиц в кристаллах).

Обозначим через  $\Delta N$  число разрешенных состояний квазичастицы для единичного нормировочного объема ( $V = 1$ ), энергии которых находятся в интервале от  $\mathcal{E}$  до  $\mathcal{E} + \Delta\mathcal{E}$ , где  $\Delta\mathcal{E} \ll \mathcal{E}$ .

**Плотностью состояний** квазичастицы  $\rho(\mathcal{E})$  называется предел отношения  $\frac{\Delta N}{\Delta\mathcal{E}}$  при  $\Delta\mathcal{E} \rightarrow 0$ , то есть число состояний на единицу объема, приходящихся на единичный интервал энергии.

Вычислим  $\Delta N$ . Проведем в  $k$ -пространстве две поверхности постоянной энергии, соответствующие энергиям  $\mathcal{E}$  и  $\mathcal{E} + \Delta\mathcal{E}$ , где  $\Delta\mathcal{E} \ll \mathcal{E}$ . Эти поверхности для двумерного случая изображены на рис. 2.3, где также нарисована двумерная решетка, в узлах которой располагаются разрешенные значения волновых векторов.

Число состояний  $\Delta N$  с энергиями от  $\mathcal{E}$  до  $\mathcal{E} + \Delta\mathcal{E}$  равно отношению объема  $\Delta^3 k$ , заключенного между поверхностями, к объему  $\delta^3 k$ , приходящемуся на одно разрешенное состояние.

Если энергия квазичастицы  $\mathcal{E}(\mathbf{k})$  зависит только от модуля волнового вектора, то поверхности постоянной энергии в  $k$ -пространстве представляют собой сферы (в двумерном пространстве на картинке — круги). Объем тонкого шарового слоя толщиной  $\Delta k$  равен  $4\pi k^2 \Delta k$ . Учитывая, что  $\delta^3 k = (2\pi)^3/V$ , а мы считаем число состояний на еди-

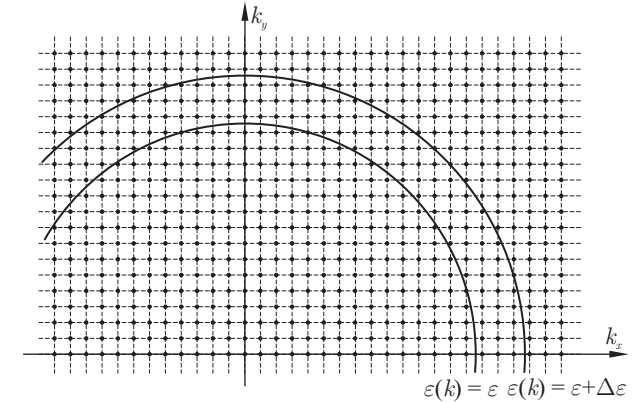


Рис. 2.3. Поверхности постоянной энергии для квадратной плоской (обратной) решетки

ницу объема ( $V = 1$ ), для плотности состояний получаем:

$$\rho(\mathcal{E}) = \lim_{\Delta\mathcal{E} \rightarrow 0} \frac{4\pi k^2 \Delta k}{(2\pi)^3 \Delta\mathcal{E}} = \frac{k^2}{2\pi^2 (d\mathcal{E}/dk)}, \quad (2.10)$$

где  $k^2$  и  $d\mathcal{E}/dk$  выражены через энергию  $\mathcal{E}$ .

### Примеры

Рассмотрим два самых важных случая: линейный и квадратичный законы дисперсии.

Линейным законом дисперсии,  $\omega = ck$ ,  $\mathcal{E} = c\hbar k$ , характеризуются такие квазичастицы как фотоны (кванты света) и длинноволновые (акустические) фононы (кванты звука). Постоянная  $c$  в законе дисперсии имеет смысл соответственно скорости света или звука.

Квадратичный закон дисперсии,  $\mathcal{E} = \hbar^2 k^2/(2m) = p^2/(2m)$ , характерен для волн, соответствующих свободным частицам с определенной массой  $m$ , например, электронам.

**а) Фотоны, акустические фононы:  $\mathcal{E} = c\hbar k$  <sup>1)</sup>**

$$\frac{d\mathcal{E}}{dk} = c\hbar, \quad k^2 = \frac{\mathcal{E}^2}{c^2\hbar^2}, \quad (2.11)$$

$$\rho(\mathcal{E}) = \frac{k^2}{2\pi^2 (d\mathcal{E}/dk)} = \frac{\mathcal{E}^2}{2\pi^2 (c\hbar)^3}. \quad (2.12)$$

<sup>1)</sup> Для фотонов  $c$  — скорость света, для акустических фононов — это скорость звука.

Число состояний  $N(\mathcal{E})$  в объеме  $V$  с энергиями меньше  $\mathcal{E}$  равно:

$$N(\mathcal{E}) = \int \int dV d\mathcal{E} \rho(\mathcal{E}) = \frac{V}{6\pi^2(c\hbar)^3} \mathcal{E}^3. \quad (2.13)$$

**б) Свободные электроны:**  $\mathcal{E} = \hbar^2 k^2 / (2m) = p^2 / (2m)$

$$\frac{d\mathcal{E}}{dk} = \frac{\hbar^2 k}{m}, \quad k^2 = \frac{2m\mathcal{E}}{\hbar^2}, \quad (2.14)$$

$$\rho(\mathcal{E}) = \frac{k^2}{2\pi^2 (d\mathcal{E}/dk)} = \frac{mk}{2\pi^2 \hbar^2} = \frac{m^{3/2}}{\sqrt{2}\pi^2 \hbar^3} \mathcal{E}^{1/2}. \quad (2.15)$$

## 2.5. Фазовая и групповая скорость волны

Есть две важные характеристики волны с определенным законом дисперсии: фазовая скорость и групповая скорость.

Фазовая скорость — скорость движения поверхности постоянной фазы:  $\mathbf{k}\mathbf{r} - \omega t = \text{const}$ . Направим ось  $x$  вдоль вектора  $\mathbf{k}$ . Тогда уравнение, определяющее поверхность постоянной фазы примет вид:  $kx - \omega t = \text{const}$ , откуда  $x = \text{const} + \frac{\omega}{k}t$ , то есть поверхность постоянной фазы движется со скоростью  $\omega/k$ :

$$v_{ph} = \frac{\omega}{k}. \quad (2.16)$$

Направление фазовой скорости совпадает с направлением волнового вектора. Уравнение (2.16) можно переписать как

$$\mathbf{v}_{ph} = \frac{\omega}{k^2} \mathbf{k}.$$

Для фотона в вакууме и акустического фонона с не слишком большой энергией (см. гл. 3)  $v_{ph} = c$ : фазовая скорость не зависит от волнового вектора.

Для электрона:

$$\mathbf{v}_{ph} = \frac{\hbar \mathbf{k}}{2m} = \frac{\mathbf{p}}{2m}. \quad (2.17)$$

Групповая скорость — это скорость движения волнового пакета, составленного из гармонических волн с волновыми векторами, лежащими вблизи заданного вектора  $\mathbf{k}_0$ , («группы» волн):

$$f(\mathbf{r}, t) = \int_{\Delta^3 k} d^3 k A(\mathbf{k}) e^{i[\mathbf{k}\mathbf{r} - \omega(\mathbf{k})t]}. \quad (2.18)$$

Здесь интегрирование производится по небольшому объему обратного пространства  $\Delta^3 k$ : волновые векторы и частоты волн, составляющих волновой пакет, не слишком сильно отличаются от величин  $\mathbf{k}_0$  и  $\omega_0 = \omega(\mathbf{k}_0)$  — «центральных» волнового вектора и частоты пакета.

Так как отклонения  $\mathbf{k}$  от  $\mathbf{k}_0$  невелики, мы можем разложить функцию  $\omega(\mathbf{k})$  в ряд по величине  $\Delta \mathbf{k} = \mathbf{k} - \mathbf{k}_0$  с точностью до первого (линейного) члена:

$$\omega(\mathbf{k}) \approx \omega_0 + \frac{\partial \omega}{\partial \mathbf{k}} \Delta \mathbf{k}. \quad (2.19)$$

Здесь  $\partial \omega / \partial \mathbf{k}$  — градиент частоты в точке  $\mathbf{k}_0$ .

Подставляя это разложение в интеграл, определяющий волновой пакет, получим:

$$\begin{aligned} f(\mathbf{r}, t) &\approx \int_{\Delta^3 k} d^3 k A(\mathbf{k}) \exp \left[ i \left( (\mathbf{k}_0 + \Delta \mathbf{k}) \mathbf{r} - (\omega_0 + \frac{\partial \omega}{\partial \mathbf{k}} \Delta \mathbf{k}) t \right) \right] = \\ &= \exp [i(\mathbf{k}_0 \mathbf{r} - \omega_0 t)] \int A(\mathbf{k}) \exp \left[ i \Delta \mathbf{k} \left( \mathbf{r} - \frac{\partial \omega}{\partial \mathbf{k}} t \right) \right] d^3 k, \quad (2.20) \\ \Delta \mathbf{k} &= \mathbf{k} - \mathbf{k}_0. \end{aligned}$$

Видно, что зависимость амплитуды волнового пакета  $|f|$  от координат и времени сводится к зависимости только от величины  $\mathbf{r} - (\partial \omega / \partial \mathbf{k}) t$ . Иными словами, пакет движется как целое со скоростью  $\partial \omega / \partial \mathbf{k}$ . Это и есть групповая скорость  $\mathbf{v}_g$ :

$$\mathbf{v}_g = \frac{\partial \omega}{\partial \mathbf{k}}. \quad (2.21)$$

Заметим, что мы учли лишь линейный член в разложении  $\omega(\mathbf{k})$  в окрестности  $\mathbf{k}_0$ , в этом приближении огибающая (форма) волнового пакета сохраняет свою форму. На самом же деле, если дисперсия не линейная, пакет медленно расплывается.

Если частота (энергия квазичастицы) зависит только от модуля волнового вектора, то групповая скорость направлена вдоль  $\mathbf{k}$  и по абсолютной величине равна  $\partial \omega / \partial k$ .

Для свободного электрона получаем:  $\mathbf{v}_g = \hbar \mathbf{k} / m = \mathbf{p} / m$ . Как и следовало ожидать, групповая скорость совпадает с классической скоростью электрона.

Отметим, что  $v_{ph}$  в два раза меньше скорости электрона с импульсом  $p$ . У волн с линейной дисперсией (фотонов и акустических фононов) групповая скорость равна фазовой. Поэтому одномерный волновой пакет из таких волн не расплывается, то есть сохраняет свою форму.

## 2.6. Обратная решетка

Рассмотрим функцию координат  $n(\mathbf{r})$ , которая обладает трансляционной симметрией кристалла, например, локальную концентрацию электронов в равновесных условиях. Трансляционная симметрия функции  $n(\mathbf{r})$  означает, что эта функция не изменяется при сдвиге на произвольный вектор трансляции  $\mathbf{A}_{n_1 n_2 n_3} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$ , то есть

$n(\mathbf{r} + \mathbf{A}_{n_1 n_2 n_3}) = n(\mathbf{r})$  (здесь  $n_1, n_2, n_3$  — целые числа,  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  — основные вектора трансляций). Иными словами,  $n(\mathbf{r})$  обладает свойством трехмерной периодичности кристалла. Такую функцию можно разложить в ряд Фурье:

$$n(\mathbf{r}) = \sum_{\mathbf{k}} n_{\mathbf{k}} e^{i\mathbf{k}\mathbf{r}}, \quad (2.22)$$

где суммирование ведется по «разрешенным» значениям  $\mathbf{k}$ . Выясним, что это за разрешенные значения, то есть какие  $\mathbf{k}$  могут присутствовать в разложении. Из свойства периодичности  $n(\mathbf{r} + \mathbf{A}) = n(\mathbf{r})$  следует:

$$\sum_{\mathbf{k}} n_{\mathbf{k}} e^{i\mathbf{k}\mathbf{A}} e^{i\mathbf{k}\mathbf{r}} = \sum_{\mathbf{k}} n_{\mathbf{k}} e^{i\mathbf{k}\mathbf{r}} \quad (2.23)$$

(индексы у вектора трансляции  $\mathbf{A}$  для краткости опускаем).

Два ряда Фурье равны в том и только том случае, если равны коэффициенты при  $\exp(i\mathbf{k}\mathbf{r})$  с одинаковыми волновыми векторами  $\mathbf{k}$ . Поэтому

$$n_{\mathbf{k}} e^{i\mathbf{k}\mathbf{A}} = n_{\mathbf{k}} \quad \text{или} \quad e^{i\mathbf{k}\mathbf{A}} = 1. \quad (2.24)$$

Отсюда следует, что разрешенные  $\mathbf{k}$  — это любые волновые вектора, удовлетворяющие условию

$$\mathbf{k}\mathbf{A} = 2\pi N, \quad (2.25)$$

где  $N$  — целое число, а  $\mathbf{A}$  — произвольный вектор трансляции решетки:

$$\mathbf{A}_{n_1 n_2 n_3} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3. \quad (2.26)$$

Можно проверить, что условию (2.25) удовлетворяют все вектора вида

$$\mathbf{k} = \mathbf{B}_{m_1 m_2 m_3} = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 + m_3 \mathbf{b}_3, \quad (2.27)$$

где  $m_1, m_2, m_3$  — целые числа, а  $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$  — основные вектора обратной решетки, связанные с основными векторами прямой решетки соотношениями:

$$\mathbf{b}_1 = 2\pi \frac{[\mathbf{a}_2 \mathbf{a}_3]}{[\mathbf{a}_2 \mathbf{a}_3] \mathbf{a}_1}, \quad \mathbf{b}_2 = 2\pi \frac{[\mathbf{a}_3 \mathbf{a}_1]}{[\mathbf{a}_3 \mathbf{a}_1] \mathbf{a}_2}, \quad \mathbf{b}_3 = 2\pi \frac{[\mathbf{a}_1 \mathbf{a}_2]}{[\mathbf{a}_1 \mathbf{a}_2] \mathbf{a}_3}. \quad (2.28)$$

Произведение  $([\mathbf{a}_1 \mathbf{a}_2] \mathbf{a}_3)$  представляет собой объем параллелепипеда, построенного на векторах прямой решетки  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ ; следовательно, это объем элементарной ячейки. Произведения  $[\mathbf{a}_1 \mathbf{a}_2]$  и т.д. изображают площади трех граней элементарной ячейки. Таким образом, векторы  $\mathbf{b}_i$  имеют размерность обратной длины, а по величине равны умноженным на  $2\pi$  обратным высотам параллелепипеда, построенного на векторах  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ . Из (2.28) следует, что между  $\mathbf{b}_i$  и  $\mathbf{a}_i$  имеют место соотношения:

$$\mathbf{a}_i \mathbf{b}_k = \begin{cases} 0, & \text{если } i \neq k, \\ 2\pi, & \text{если } i = k. \end{cases}$$

Это означает, что вектор  $\mathbf{b}_1$  перпендикулярен к векторам  $\mathbf{a}_2, \mathbf{a}_3$  и аналогично для  $\mathbf{b}_2$  и  $\mathbf{b}_3$ .

Подведем итоги. Любая удовлетворяющая условиям Борна–Кармана функция  $f(\mathbf{r})$  может быть разложена в ряд Фурье:

$$f(\mathbf{r}) = \sum_{\mathbf{k}} F(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}}. \quad (2.29)$$

Входящие в эту сумму значения вектора  $\mathbf{k}$  образуют в пространстве векторов  $\mathbf{k}$  (в обратном пространстве) кубическую решетку с ребром элементарного куба  $2\pi/L$ , где  $L$  — условная длина (в окончательных результатах  $L \rightarrow \infty$ ). Эта решетка физического смысла не имеет и введена только для удобства.

Для функции, периодической с периодом кристалла, то есть функции, удовлетворяющей условиям

$$n(\mathbf{r} + \mathbf{A}) = n(\mathbf{r}), \quad (2.30)$$

где  $\mathbf{A} = \mathbf{A}_{n_1 n_2 n_3}$  — любой вектор трансляции (вектор прямой решетки), фурье-разложение имеет вид:

$$n(\mathbf{r}) = \sum_{\mathbf{B}} n_{\mathbf{B}} e^{i\mathbf{B}\mathbf{r}}, \quad (2.31)$$

где  $\mathbf{B} = \mathbf{B}_{m_1 m_2 m_3}$  — векторы обратной решетки, удовлетворяющие условиям  $e^{i\mathbf{B}\mathbf{A}} = 1$ , то есть

$$\mathbf{B}_{m_1 m_2 m_3} \mathbf{A}_{n_1 n_2 n_3} = 2\pi N, \quad (2.32)$$

где  $N$  — целое число. Нетрудно увидеть, пользуясь формулами (2.26) и (2.27), что  $N = m_1 n_1 + m_2 n_2 + m_3 n_3$ . Решетка в обратном пространстве, основанная на векторах  $\mathbf{B}$ , называется обратной решеткой кристалла. Обратная решетка имеет очень важный физический смысл (см. ниже). В частности, с помощью понятия обратной решетки удобно описывать закономерности рассеяния рентгеновских лучей.

**Отступление. Индексы Миллера.** Известно (см. ниже (2.40)), что уравнение вида  $\mathbf{b} \cdot \mathbf{r} = \text{const}$ , где  $\mathbf{b}$  — постоянный вектор, описывает плоскость, перпендикулярную к вектору  $\mathbf{b}$  и находящуюся на расстоянии  $\text{const}/b$  от начала координат. Выберем начало координат в каком-нибудь из узлов решетки Браве. Пусть вектор  $\mathbf{b} = \mathbf{B}$ , где  $\mathbf{B}$  — вектор обратной решетки (см. (2.27)). Тогда уравнение (2.32) представляет собой плоскость, заполненную бесконечным множеством узлов решетки Браве. О таких плоскостях говорят как о *кристаллических*. При заданных  $m_1, m_2, m_3$  и пробегаяющей различные значения постоянной  $N = m_1 n_1 + m_2 n_2 + m_3 n_3$  уравнение (2.32) определяет бесчисленное множество кристаллических плоскостей, которые все параллельны друг другу. Числа  $m_1, m_2, m_3$  в (2.32) можно представлять себе всегда взаимно простыми, не имеющими общего делителя. Они называются индексами Миллера данного семейства кристаллических плоскостей и

обозначаются как  $(m_1 m_2 m_3)$ . Так, плоскости граней кубического кристалла имеют индексы Миллера (100), (010), (001),  $(\bar{1}00)$ ,  $(0\bar{1}0)$ ,  $(00\bar{1})$  (рис. 2.4). Плоскости, параллельные диагональной плоскости основного параллелепипеда решетки, имеют индексы (111) и т. д. (рис. 2.4).

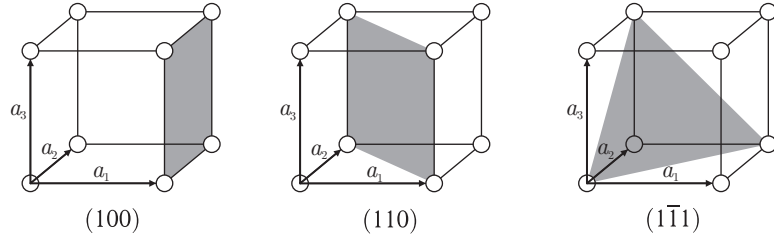


Рис. 2.4. Три атомные плоскости и их индексы Миллера для простой кубической решетки Браве

Аналогичные обозначения используются для направлений в прямой решетке, но при этом, чтобы избежать путаницы с индексами Миллера (для направлений в обратной решетке), применяют квадратные, а не круглые скобки. Так, пространственная диагональ простой кубической решетки имеет направление  $[111]$ , а в общем случае радиус-вектор  $n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$  имеет направление  $[n_1 n_2 n_3]$  по отношению к началу отсчета.

В кубическом кристалле, в силу симметрии кристалла, плоскости (100), (010) и (001) эквивалентны. В совокупности их обозначают как плоскости  $\{100\}$ . Аналогичное правило применяют и в отношении направлений: направления  $[100]$ ,  $[010]$ ,  $[001]$ ,  $[\bar{1}00]$ ,  $[0\bar{1}0]$ ,  $[00\bar{1}]$  в кубическом кристалле называют направлениями  $\langle 100 \rangle$ .

Вид обратной решетки определяется видом прямой решетки, например, гранецентрированной решетке Браве соответствует объемноцентрированная обратная решетка. Можно доказать, что прямая и обратная решетки обладают одинаковой точечной симметрией<sup>1)</sup>.

Итак, объем примитивной ячейки кристалла  $v_0$  — это объем параллелепипеда, построенного на векторах  $\mathbf{a}_1$ ,  $\mathbf{a}_2$ ,  $\mathbf{a}_3$ :

$$v_0 = (\mathbf{a}_1 [\mathbf{a}_2 \mathbf{a}_3]). \quad (2.33)$$

Используя определение векторов  $\mathbf{b}_1$ ,  $\mathbf{b}_2$ ,  $\mathbf{b}_3$  (2.28), можно найти объем примитивной ячейки обратной решетки:

$$(\mathbf{b}_1 [\mathbf{b}_2 \mathbf{b}_3]) = \frac{(2\pi)^3}{v_0}, \quad (2.34)$$

Выше, при рассмотрении граничных условий Борна–Кармана, мы пришли к выводу, что в объеме обратного пространства  $\Delta^3 k$  содер-

жится  $\frac{V}{(2\pi)^3} \Delta^3 k$  разрешенных состояний волнового вектора ( $V$  — объем кристалла). Отсюда следует, что примитивная ячейка обратной решетки содержит

$$\frac{V}{(2\pi)^3} \frac{(2\pi)^3}{v_0} = \frac{V}{v_0} \quad (2.35)$$

разрешенных состояний волнового вектора, то есть столько, сколько примитивных ячеек прямой решетки содержит объем кристалла.

## 2.7. Рассеяние рентгеновских лучей кристаллом

Рассмотрим электромагнитную волну, в поле которой находится свободный электрон. Движение нерелятивистского электрона определяется главным образом электрическим полем волны: если скорость электрона мала по сравнению со скоростью света, влиянием магнитного поля можно пренебречь. Переменное электрическое поле раскачивает электрон, который из-за этого излучает электромагнитные волны. Эти волны называются рассеянными: часть энергии падающей волны переизлучается электроном. Свободный электрон рассеивает волны во всех направлениях, кроме направления своих колебаний: зависимость мощности излучения от угла пропорциональна  $\sin^2 \theta$  ( $\theta$  отсчитывается от направления колебаний электрона). Диаграмма направленности излучения, рассеянного электроном, изображена на рис. 2.5.

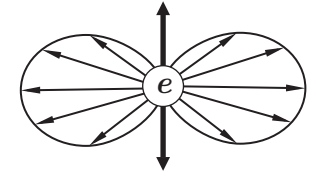


Рис. 2.5. Диаграмма направленности рассеянного излучения

Иная картина наблюдается при рассеянии электромагнитных волн кристаллом в том случае, когда длина волны излучения сравнима с межатомными расстояниями в кристалле, то есть с постоянной решетки<sup>1)</sup>. Такими длинами волн обладает излучение рентгеновского диапазона. Кристалл, в зависимости от длины волны падающего излучения, либо практически не рассеивает это излучение, либо рассеивает его только в определенных направлениях.

Объясняется это тем, что волны, рассеянные разными участками кристалла, складываются, то есть происходит интерференция волн. Вследствие периодичности кристаллической структуры для большинства направлений рассеяния эта интерференция носит деструктивный характер: рассеянные волны гасят друг друга. И лишь в некоторых выделенных направлениях волны складываются в фазе, то есть усиливают друг друга.

<sup>1)</sup> Это соответствует частотам в диапазоне  $\nu = 3 \cdot 10^{18} \text{ с}^{-1}$  и энергиям кванта  $\epsilon = 12 \text{ кэВ}$ .

<sup>1)</sup> Ансельм А. И. Введение в теорию полупроводников. — М.: Наука, 1978.



Посмотрим, какой ток возникает в кристалле при облучении его электромагнитной волной рентгеновского диапазона. Частота рентгеновской волны намного выше всех характерных электронных частот в кристалле, поэтому электроны в поле волны ведут себя как свободные. Если падающая волна имеет частоту  $\omega_0$ , то и ток ведет себя как гармоническая функция времени с частотой  $\omega_0$ .

Амплитуда плотности тока в точке  $\mathbf{r}$  пропорциональна произведению амплитуды электрического поля в этой точке на плотность электронов  $n(\mathbf{r})$ . Это справедливо для нерелятивистского движения электронов.

Будем считать, что влиянием рассеянных волн на электроны можно пренебречь и их движение определяется только первичной падающей волной  $\mathbf{E}(\mathbf{r})$ . Тогда для плотности тока имеем:

$$\mathbf{j}(\mathbf{r}) \sim \mathbf{E}(\mathbf{r})n(\mathbf{r}) \sim e^{i\mathbf{k}_0\mathbf{r}} n(\mathbf{r}). \quad (2.36)$$

Здесь  $\mathbf{k}_0$  — волновой вектор падающей волны,  $\mathbf{E}(\mathbf{r}) \sim e^{i\mathbf{k}_0\mathbf{r}}$ .

Концентрация электронов  $n(\mathbf{r})$ , как периодическая функция с трансляционной симметрией кристалла, раскладывается в ряд Фурье по векторам  $\mathbf{B}$  обратной решетки кристалла:

$$n(\mathbf{r}) = \sum_{\mathbf{B}} n_{\mathbf{B}} e^{i\mathbf{B}\mathbf{r}}. \quad (2.37)$$

Для плотности тока получаем:

$$\mathbf{j}(\mathbf{r}) \sim \sum_{\mathbf{B}} n_{\mathbf{B}} e^{i\mathbf{k}_0\mathbf{r}} e^{i\mathbf{B}\mathbf{r}} = \sum_{\mathbf{B}} n_{\mathbf{B}} e^{i(\mathbf{B}+\mathbf{k}_0)\mathbf{r}}. \quad (2.38)$$

Итак, в разложении плотности тока, а, следовательно, и в разложении Фурье создаваемого им электромагнитного поля  $\mathbf{E}_{\text{расс}} \sim \mathbf{j}$ , присутствуют только компоненты Фурье с волновыми векторами  $\mathbf{k} = \mathbf{k}_0 + \mathbf{B}$ .

Таким образом, волновой вектор рассеянной волны может отличаться от волнового вектора падающей волны только на произвольный вектор обратной решетки.

Дадим квантовую интерпретацию этого явления. Плоской электромагнитной волне соответствует квазичастица — фотон. Импульс фотона, как и импульс любой квазичастицы, равен  $\hbar\mathbf{k}$ , а энергия —  $\hbar\omega$ .

Импульс фотона при рассеянии на кристалле, как видно из (2.38), может изменяться на величину  $\hbar\mathbf{B}$ . Дискретный набор векторов  $\hbar\mathbf{B}$  — это дискретный набор импульсов, которые идеальная кристаллическая решетка может передать фотону при рассеянии. И это выполнено не только для фотона, но и для любой квазичастицы, рассеивающейся на идеальном кристалле.

Найдем теперь множество волновых векторов фотонов, для которых разрешено рассеяние на кристалле.

В этом разделе мы изучаем упругое рассеяние. Это означает, что энергия рассеянного фотона должна быть равна энергии падающего.

Закон дисперсии электромагнитной волны гласит:  $\omega = ck$ . Следовательно, при упругом рассеянии не может измениться модуль волнового вектора:  $k = k_0$ . Выясним теперь, при каких условиях это равенство совместимо с условием рассеяния  $\mathbf{k} = \mathbf{k}_0 + \mathbf{B}$ . Другими словами, выясним, каким должен быть волновой вектор  $\mathbf{k}_0$ , чтобы волна могла куда-нибудь рассеяться. Для этого возведем равенство  $\mathbf{k}_0 + \mathbf{B} = \mathbf{k}$  в квадрат, учитывая, что  $k^2 = k_0^2$ . В результате получим:

$$(\mathbf{k}_0\mathbf{B}) = -\frac{B^2}{2}. \quad (2.39)$$

Упругое рассеяние волны с волновым вектором  $\mathbf{k}_0$  возможно, только если найдется такой вектор обратной решетки  $\mathbf{B}$ , при котором будет удовлетворяться это равенство.

**Математическое отступление.** Рассмотрим плоскость в трехмерном пространстве, не проходящую через начало координат. Проведем нормаль к этой плоскости.

Очевидно, что все радиус-векторы, концы которых лежат на плоскости, имеют одинаковую проекцию на ее нормаль  $\mathbf{p}$ . Эта проекция равна расстоянию  $p$  от начала координат до плоскости. В векторной форме это утверждение можно записать в виде:

$$\frac{\mathbf{r}\mathbf{p}}{p} = p \quad \text{или} \quad \mathbf{r}\mathbf{p} = p^2. \quad (2.40)$$

Итак, уравнение вида  $\mathbf{r}\mathbf{p} = p^2$  определяет геометрическое место точек плоскости. Плоскость проходит через точку, заданную радиус-вектором  $\mathbf{p}$ , перпендикулярно этому радиус-вектору.

Таким образом, уравнение (2.39) — это уравнение плоскости в  $\mathbf{k}$ -пространстве. Роль вектора-нормали  $\mathbf{p}$ , через конец которого проведена плоскость, играет при этом вектор  $-\mathbf{B}/2$ . Поскольку  $\mathbf{B}$  — любой вектор обратной решетки, вместо  $-\mathbf{B}/2$  можно писать  $\mathbf{B}/2$ . Геометрическое место концов волновых векторов  $\mathbf{k}_0$  рассеивающихся волн — это совокупность плоскостей, перпендикулярных направлениям  $\mathbf{OB}$  и лежащих на расстоянии  $B/2$  от начала координат  $O$  (см. рис.2.6). Волновые вектора  $\mathbf{k}_0$ , не кончающиеся на этих плоскостях, соответствуют волнам, которые проходят через кристалл, не рассеиваясь упруго.

Совместим один из узлов обратной решетки с началом координат в  $\mathbf{k}$ -пространстве. Для каждого вектора обратной решетки, соединяющего этот узел (начало координат) с другим узлом, построим плоскость, перпендикулярную этому вектору и проходящую через его середину. Объединение

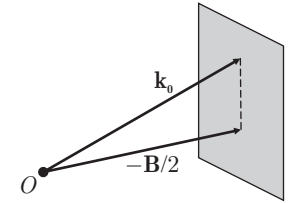


Рис. 2.6. Волновой вектор падающей волны, удовлетворяющий условию упругого рассеяния (2.40)



таких плоскостей и образует множество «рассеивающихся» волновых векторов: волна рассеивается, если конец ее волнового вектора лежит на какой-либо из этих плоскостей.

## 2.8. Зона Бриллюэна

Определенное таким образом семейство плоскостей ограничивает вокруг узла обратной решетки, совмещенного с началом координат, некоторый объем, имеющий форму многогранника.

В случае простой кубической решетки многогранник является кубом. Его грани образованы плоскостями, проведенными через середины векторов, соединяющих данный узел решетки с его ближайшими соседями. Все остальные плоскости проходят дальше от узла и не участвуют в формировании граней куба.

В случае объемноцентрированной кубической решетки получается более сложный многогранник (см. ниже). Его грани образуются плоскостями, проведенными между данным узлом и его ближайшими и вторыми ближайшими соседями.

Построенный многогранник называется первой зоной Бриллюэна.

Если из всего объединения плоскостей удалить грани первой зоны Бриллюэна, то оставшееся множество ограничит новый, больший многогранник вокруг начала координат. Часть этого многогранника, не входящая в первую зону Бриллюэна, образует вторую зону Бриллюэна. Иными словами, вторая зона Бриллюэна — это часть обратного пространства между первым и вторым многогранниками.

На следующем шаге мы получим третью зону Бриллюэна и т. д.

На рис. 2.7 изображены четыре зоны Бриллюэна для двумерной прямоугольной обратной решетки. Строятся они так же, как и в трехмерном случае, только вместо плоскостей проводятся прямые.

Первая зона Бриллюэна плоской квадратной решетки получается как область, заключенная между взаимно перпендикулярными прямыми, проходящими через середины кратчайших векторов обратной решетки  $\mathbf{G}_1$  и еще трех векторов, эквивалентных  $\mathbf{G}_1$  по симметрии (см. рис. 2.7, а). Таким образом, для построения первой зоны Бриллюэна нужны четыре вектора обратной решетки (на рис. 2.7, а два вектора  $\mathbf{G}_1$  отмечены жирными линиями, а два пунктирными). Если постоянная решетки равна  $a$ , то эти четыре вектора равны:  $\pm \frac{2\pi}{a} \hat{\mathbf{k}}_x$  и  $\pm \frac{2\pi}{a} \hat{\mathbf{k}}_y$ , где  $\hat{\mathbf{k}}_x$  и  $\hat{\mathbf{k}}_y$  — единичные вектора вдоль осей  $k_x$  и  $k_y$ .

Вторая зона Бриллюэна строится при помощи вектора  $\mathbf{G}_2$  и еще трех векторов, эквивалентных  $\mathbf{G}_2$  по симметрии. Аналогичным путем при помощи вектора  $\mathbf{G}_3$  строится и третья зона Бриллюэна. Вторая и третья зоны, состоящие из одинаковых участков, изображены на рис. 2.7, б.

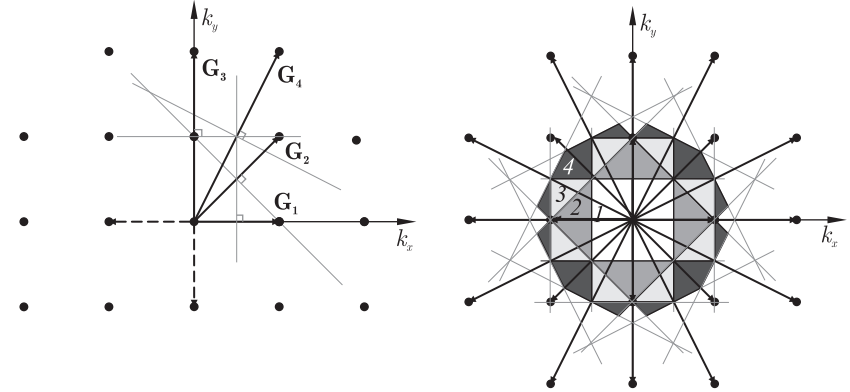


Рис. 2.7. а) Построение в  $\mathbf{k}$ -пространстве первых четырех зон Бриллюэна для случая плоской квадратной решетки. Четыре наименьших вектора обратной решетки обозначены через  $\mathbf{G}_1$ ,  $\mathbf{G}_2$ ,  $\mathbf{G}_3$  и  $\mathbf{G}_4$ . Проведены прямые через середины векторов  $\mathbf{G}_1$ ,  $\mathbf{G}_2$ ,  $\mathbf{G}_3$  и  $\mathbf{G}_4$  перпендикулярно к ним. б) Проводя, кроме уже указанных в (а), все эквивалентные им по симметрии прямые линии, мы получим области  $\mathbf{k}$ -пространства, образующие первые четыре зоны Бриллюэна. Числа 1, 2, 3, 4 написаны на участках, относящихся к соответствующей по номеру зоне

Чтобы определить границы той или иной зоны Бриллюэна, мы должны рассмотреть несколько не эквивалентных векторов обратной решетки. Например, границы крайнего справа верхнего участка третьей зоны (рис. 2.7, б) образованы перпендикулярами через середины трех векторов  $\mathbf{G}$ , а именно:  $\frac{2\pi}{a} \hat{\mathbf{k}}_x$ ,  $\frac{4\pi}{a} \hat{\mathbf{k}}_y$ ,  $\frac{2\pi}{a} (\hat{\mathbf{k}}_x + \hat{\mathbf{k}}_y)$ .

Теперь, после того как мы ввели понятие зон Бриллюэна, мы можем переформулировать условие рассеяния рентгеновской волны следующим образом: волна рассеивается, если конец ее волнового вектора лежит на границе некоторой зоны Бриллюэна.

В дальнейшем мы для простоты часто будем рассматривать одномерные решетки и нам придется иметь дело с одномерными зонами Бриллюэна.

Пусть прямая решетка имеет период  $a$ .

Понятие обратной решетки возникло из разложения периодической функции по плоским волнам. В одномерном случае разложение периодической функции в ряд Фурье имеет вид:

$$f(x) = \sum_n f_n \exp\left(\frac{2\pi n i}{a} x\right) = \sum_G f_G e^{iGx}. \quad (2.41)$$

Таким образом, обратную решетку в одномерном случае образуют точки  $G_n = 2\pi n/a$ , эти точки и играют роль векторов обратной решетки.

Зоны Бриллюэна строятся так же, как и в трехмерном случае, но вместо плоскостей ставятся точки на середине отрезков, соединяющих узлы обратной решетки. Первая зона — отрезок  $[-\pi/a, \pi/a]$ , вторая — объединение отрезков  $[-2\pi/a, -\pi/a]$  и  $[\pi/a, 2\pi/a]$ , и т. д. (рис. 2.8).

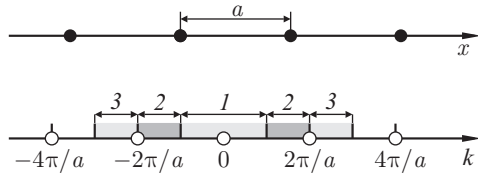


Рис. 2.8. Зоны Бриллюэна в одномерном случае

Первую зону Бриллюэна можно рассматривать как примитивную ячейку обратной решетки, обладающую точечной симметрией этой решетки, так как она содержит один атом. Действительно, если мы построим первую зону Бриллюэна вокруг каждого узла обратной решетки (поместив начало координат в этот узел), то такие зоны целиком заполняют все пространство, не перекрываясь при этом друг с другом. Отсюда, в частности, следует, что объем первой зоны Бриллюэна равен объему примитивной ячейки обратной решетки.

Некоторые свойства зон Бриллюэна:

1. Объем всех зон Бриллюэна одинаков и равен объему примитивной ячейки обратной решетки. Отсюда в частности следует, что зона Бриллюэна содержит столько разрешенных значений волновых векторов, сколько примитивных ячеек содержится в объеме кристалла (см. (2.35)).

2. Если к вектору обратного пространства  $\mathbf{k}$ , конец которого лежит в первой зоне Бриллюэна, прибавить любой вектор обратной решетки, то полученный вектор выйдет за пределы первой зоны.

3. И наоборот, из любого вектора  $\mathbf{k}$ , конец которого не лежит в первой зоне Бриллюэна, можно путем вычитания некоторого вектора обратной решетки получить вектор, конец которого находится в пределах первой зоны.

Таким образом, все зоны Бриллюэна являются репликами первой зоны: их можно получить, транслируя точки, принадлежащие первой зоне, на вектора обратной решетки.

Большинство распространенных полупроводников имеет гранецентрированную кубическую (ГЦК) решетку Браве, поэтому посмотрим, как устроена первая зона Бриллюэна у этой решетки.

Обратная решетка в этом случае является объемноцентрированной кубической (ОЦК). Первая зона Бриллюэна изображена на рис. 2.9. Получается она, согласно описанному выше алгоритму, следующим образом.

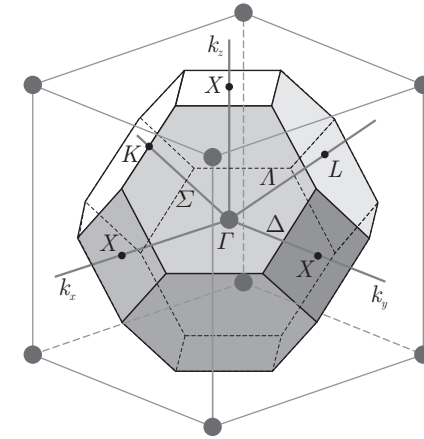


Рис. 2.9. Зона Бриллюэна для гранецентрированной кубической решетки. Видно, что обратная решетка является объемноцентрированной кубической решеткой

Проведем сначала плоскости, делящие пополам отрезки между центральным узлом в кубической элементарной ячейке и его вторыми ближайшими соседями — центральными узлами в прилегающих кубических ячейках. Эти плоскости ограничат вокруг центрального узла куб, совпадающий с кубической ячейкой. По углам этого куба расположены ближайшие соседи центрального узла. Плоскости, проведенные между центральным узлом и этими соседями, при пересечении с кубом образуют шестиугольные грани, а на гранях куба «вырежут» квадраты.

Некоторые точки и линии высокой симметрии в первой зоне Бриллюэна ГЦК-решетки имеют общепринятые обозначения.

Центр зоны Бриллюэна обозначается греческой буквой  $\Gamma$ .

Точки на поверхности обозначаются латинскими буквами:  $X$  — центр квадратной грани [направление (100) в  $\mathbf{k}$ -пространстве и эквивалентные ему, всего 6 точек];  $L$  — центр шестиугольной грани [направления (111) и эквивалентные, 8 точек];  $K$  — середина ребра между соседними шестиугольными гранями [направление (110) и эквивалентные, 12 точек].

Линии, соединяющие центр зоны с точками на поверхности, обозначаются греческими буквами:  $\Gamma X$  —  $\Delta$  [направления типа (100)],  $\Gamma L$  —  $\Lambda$  [направления типа (111)],  $\Gamma K$  —  $\Sigma$  [направления типа (110)].

**Построение Эвальда.** Выше мы нашли условие, при котором рентгеновская волна упруго рассеивается кристаллом: ее волновой вектор  $\mathbf{k}_0$  должен лежать на границе какой-либо зоны Бриллюэна. Алгебраически это условие записывается в форме (2.39).

Если волна рассеивается, то возникает вопрос: в каких направлениях происходит рассеяние? Мы выяснили, что волновой вектор рассеян-

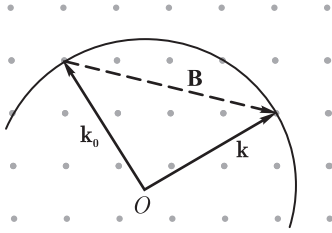


Рис. 2.10. Построение Эвальда.  $\mathbf{k}_0$  — волновой вектор падающей волны;  $\mathbf{k}$  — волновой вектор одной из рассеянных волн;  $\mathbf{B}$  — вектор обратной решетки

ной волны  $\mathbf{k}$  может отличаться от волнового вектора падающей волны  $\mathbf{k}_0$  только на произвольный вектор обратной решетки кристалла. При этом длины волновых векторов падающей и рассеянной волн равны:  $k = k_0$  (так как частота рассеянной волны должна быть равна частоте падающей).

Найти направления рассеяния, определяемые этими условиями, можно графически.

Совместим конец волнового вектора падающей волны  $\mathbf{k}_0$  с узлом обратной решетки (рис. 2.10). Изобразим в обратном пространстве сферу радиуса  $k_0$  с центром в точке  $O$  — начале вектора  $\mathbf{k}_0$ . Множество векторов, соединяющих точку  $O$  с точками, лежащими на поверхности сферы, — это множество волновых векторов  $\mathbf{k}$ , удовлетворяющих условию упругого рассеяния,  $k = k_0$ . Но рассеянными волнам соответствуют лишь те из этих векторов, для которых выполняется второе условие рассеяния: они должны отличаться от  $\mathbf{k}_0$  на произвольный вектор обратной решетки. Вектор обратной решетки — это вектор, соединяющий два некоторых узла обратной решетки. Поскольку конец вектора  $\mathbf{k}_0$  совмещен с узлом обратной решетки, то и конец волнового вектора рассеянной волны должен совпадать с некоторым узлом обратной решетки и лежать при этом на поверхности сферы. Таким образом, волновые вектора рассеянных волн — это вектора, соединяющие точку  $O$  с узлами обратной решетки, лежащими на сфере.

Если волновой вектор падающей волны не удовлетворяет условию рассеяния (2.39), то есть не лежит на границе некоторой зоны Бриллюэна, то сфера не пересекается с узлами обратной решетки, за исключением конца вектора  $\mathbf{k}_0$ , и рассеяния не происходит.

Описанное графическое построение называется построением Эвальда.

## Задачи

**Задача 2.1.** Рассмотрим электрон, представляющий собой в момент  $t = 0$  гауссов одномерный волновой пакет  $f(x) = f_0 \exp(-x^2/d^2)$ , где  $f_0$  — нормировочная постоянная,  $d$  — размер пакета. Опишите распыление этого пакета со временем.

*Указание.* Перейдите к фурье-представлению и воспользуйтесь уравнением (2.18).

**Задача 2.2.** Получите условие дифракции рентгеновских лучей на системе параллельных атомных плоскостей, расстояние между которыми равно  $d$ . Сравните его с уравнением (2.39).

*Ответ.* Это условие Брегга (иногда называемое условием Брегга–Вульфа):

$$2d \sin \theta = n\lambda, \quad n \in \mathbb{N},$$

где  $\lambda$  — длина волны рентгеновского излучения;  $\theta$  — угол между направлением падения лучей и атомными плоскостями.

**Задача 2.3.** Докажите, что обратной к кубической гранецентрированной решетке будет кубическая объемноцентрированная (и наоборот).

*Указание.* Выпишите явный вид векторов обратной решетки в декартовой системе координат. В качестве базисных векторов ГЦК решетки можно взять, например,

$$\mathbf{a}_1 = \frac{a}{2}(\mathbf{e}_y + \mathbf{e}_z), \quad \mathbf{a}_2 = \frac{a}{2}(\mathbf{e}_x + \mathbf{e}_z), \quad \mathbf{a}_3 = \frac{a}{2}(\mathbf{e}_x + \mathbf{e}_y).$$

## Глава 3

### КОЛЕБАНИЯ КРИСТАЛЛИЧЕСКОЙ РЕШЕТКИ

#### 3.1. Введение

Кристаллическая структура — это равновесное состояние системы атомов, отвечающее минимуму потенциальной энергии. В состоянии покоя сумма сил, действующих на каждый атом кристалла со стороны других атомов, равна нулю.

Если вывести эту систему из положения равновесия, то в кристалле возникнут сложные колебания. Эти колебания, в частности, всегда имеются при конечной температуре, когда кристаллическая структура обладает определенной (тепловой) энергией, то есть не находится в состоянии статического равновесия.

Рассмотрим колебания решетки в рамках классической механики <sup>1)</sup>.

При смещении атома относительно других атомов кристалла возникает сила, которая стремится вернуть его в равновесное положение. Если смещения невелики, мы можем разложить зависимость силы от смещений в ряд и ограничиться линейными по смещениям членами. Тогда колебания кристаллической решетки будут линейными, то есть будут описываться системой линейных дифференциальных уравнений.

Такая система уравнений обладает важным свойством: если есть несколько решений, то их линейная комбинация также является решением, иными словами, сумма двух возможных колебаний — это тоже колебание.

Эта система может быть решена, если известна зависимость силы, действующей на атом, от его смещения, а основные характеристики линейных колебаний могут быть предсказаны на основании одних только свойств симметрии кристалла.

Чтобы показать главные черты линейных колебаний кристаллической решетки, мы рассмотрим простейший случай одномерного кристалла — линейную цепочку атомов.

<sup>1)</sup> Вопрос о колебаниях решетки с точки зрения квантовой механики изложен в Приложении к § 4 настоящей главы.

#### 3.2. Одномерная цепочка с одним атомом в примитивной ячейке

Рассмотрим одномерную периодическую цепочку атомов — одномерный кристалл с одним атомом в элементарной ячейке. Пусть период этой цепочки будет равен  $a$ . Тогда в состоянии равновесия координата  $n$ -го атома цепочки  $x_n$  будет равна  $na$ .

Обозначим через  $u_n$  смещение  $n$ -го атома из положения равновесия. Будем считать, что атомы взаимодействуют только с ближайшими соседями. Сила, с которой  $(n + 1)$ -й атом действует на  $n$ -й, зависит от разности смещений этих двух атомов  $u_{n+1} - u_n$ . При небольших смещениях ( $|u_{n+1} - u_n| \ll a$ ) эту силу можно считать пропорциональной разности смещений:  $F_{n,n+1} = \gamma(u_{n+1} - u_n)$ , где  $\gamma$  — коэффициент пропорциональности. Удобно представить, что атомы связаны друг с другом пружинками с жесткостью  $\gamma$ .

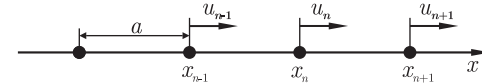


Рис. 3.1. Одномерная цепочка с одним атомом в элементарной ячейке

На рис. 3.1 пружинка между  $n$ -м и  $(n + 1)$ -м атомами растянута, так что ее упругая сила действует на  $n$ -й атом в положительном направлении. Растянутая пружинка между  $(n - 1)$ -м и  $n$ -м атомом действует на  $n$ -й атом в отрицательном направлении с силой  $F_{n,n-1} = -\gamma(u_n - u_{n-1})$ .

Запишем закон Ньютона для  $n$ -го атома цепочки:

$$M \frac{d^2 u_n}{dt^2} = F_{n,n+1} + F_{n,n-1} = \gamma(u_{n+1} - u_n) - \gamma(u_n - u_{n-1}), \quad (3.1)$$

где  $M$  — масса атома. При выводе (3.1) мы учли, что производные по времени от смещения атома и от его координаты равны.

Первое слагаемое в правой части — это сила, действующая на  $n$ -й атом со стороны  $n + 1$ -го, второе — сила, действующая со стороны  $n - 1$ -го атома.

После упрощения получим

$$M \frac{d^2 u_n}{dt^2} = \gamma [u_{n+1} + u_{n-1} - 2u_n]. \quad (3.2)$$

Система таких уравнений, записанных для каждого атома, полностью описывает колебания цепочки.

Если рассматривать только длинноволновые колебания, то есть колебания с длиной волны, много большей периода цепочки  $a$ , то можно заменить разность  $u_{n+1} - u_n$  на  $(\partial u_n / \partial x)a$ , а величину, стоящую в пра-

вой части (3.2), — на  $\gamma a^2(\partial^2 u / \partial x^2)$ . В результате получим волновое уравнение

$$\frac{\partial^2 u}{\partial t^2} = a^2 \frac{\gamma}{M} \frac{\partial^2 u}{\partial x^2}, \quad (3.3)$$

решением которого являются волны  $u = A \exp(ikx - i\omega t)$  с линейным законом дисперсии  $\omega = s|k|$  (звуковые волны). Здесь  $s$  — скорость звука:

$$s = a \sqrt{\frac{\gamma}{M}}. \quad (3.4)$$

Ниже мы решим задачу о колебаниях одномерной цепочки точно, то есть рассмотрим колебания со всеми возможными длинами волн. При этом мы установим границы применимости (3.3).

Будем искать колебания, зависящие от времени по гармоническому закону,

$$u_n = C_n e^{-i\omega t}. \quad (3.5)$$

Здесь  $\omega$  — частота колебаний, одна и та же для всех атомов (такие колебания называются гармоническими),  $C_n$  — комплексная амплитуда колебаний  $n$ -го атома. Напомним, что колебания описываются вещественной частью (3.5), но технически удобно пользоваться комплексным решением.

Такая подстановка — стандартный метод решения линейных систем уравнений с постоянными коэффициентами. В силу линейности уравнений колебание с произвольной временной зависимостью может быть разложено в интеграл (ряд) Фурье по гармоническим колебаниям.

Из (3.2) для амплитуды  $C_n$  получаем уравнение

$$-M\omega^2 C_n = \gamma [C_{n+1} + C_{n-1} - 2C_n]. \quad (3.6)$$

Эти уравнения образуют бесконечную систему линейных уравнений. Если применить к цепочке граничные условия Борна–Кармана, то система (3.6) станет конечной. (Заметим, что в одномерном случае условия Борна–Кармана эквивалентны тому, что цепочка достаточно большой длины  $L$  замкнута в кольцо.) Тогда, приравнявая ее определитель нулю, мы можем найти частоты собственных колебаний, а затем, решая систему уравнений для каждой из найденных частот, и соответствующие амплитуды.

Но мы поступим иначе. Будем искать решение в виде плоской волны

$$C_n = A e^{ikx_n}. \quad (3.7)$$

Подставив это выражение в (3.6), получим

$$-M\omega^2 e^{ikx_n} = \gamma (e^{ikx_{n+1}} + e^{ikx_{n-1}} - 2e^{ikx_n}). \quad (3.8)$$

Разделим последнее уравнение на  $\exp(ikx_n)$  и воспользуемся тем, что  $x_{n+1} = x_n + a$ ,  $x_{n-1} = x_n - a$ , тогда

$$-M\omega^2 = \gamma (e^{ika} + e^{-ika} - 2). \quad (3.9)$$

Таким образом, подстановка в виде плоской волны оказалась верной: мы избавились от номера атома  $n$  и получили уравнение, связывающее  $\omega$  и  $k$ , то есть уравнение, которое определяет закон дисперсии волн.

Поскольку

$$\frac{e^{ika} + e^{-ika}}{2} = \cos ka, \quad (3.10)$$

то

$$\omega^2 = \frac{2\gamma}{M} (1 - \cos ka) = \frac{4\gamma}{M} \sin^2 \frac{ka}{2}, \quad (3.11)$$

и мы получаем закон дисперсии для упругих колебаний бесконечной одномерной цепочки

$$\omega = \sqrt{\frac{4\gamma}{M}} \left| \sin \frac{ka}{2} \right|. \quad (3.12)$$

Итак, мы пришли к выводу, что при колебании одномерной цепочки смещения атомов описываются плоской гармонической волной

$$u_n = \operatorname{Re} [A e^{i(kx_n - \omega t)}] = |A| \cos(kx_n - \omega t + \varphi). \quad (3.13)$$

Точнее, колебания представляют собой произвольную линейную комбинацию таких волн. Здесь  $\varphi$  — фаза комплексной амплитуды  $A = |A| \exp(i\varphi)$ .

Напомним, что смещение — это вещественная величина, которая описывается вещественной частью комплексной плоской гармонической волны, что явно записано в (3.13). В дальнейшем при описании вещественных колебаний комплексной плоской волной мы будем для краткости опускать обозначение вещественной части.

Волновой вектор  $k$  в плоской волне (3.13) может, вообще говоря, быть любым. Но вследствие дискретности цепочки ( $x_n$  может принимать лишь дискретный набор значений  $na$ ) плоские волны, волновые вектора которых отличаются друг от друга на произвольный вектор обратной решетки  $2\pi l/a$ , описывают одно и то же колебание. (Здесь  $l$  — любое целое число.)

Действительно, так как  $x_n = na$ , то

$$\exp \left[ i \left( k + \frac{2\pi l}{a} \right) x_n \right] = \exp \left[ i \left( kx_n + \frac{2\pi l}{a} na \right) \right] = e^{ikx_n}. \quad (3.14)$$

Поэтому достаточно рассматривать волновые вектора, лежащие в первой зоне Бриллюэна  $-\pi/a < k < \pi/a$ . Граничные значения волнового вектора  $\pm\pi/a$  соответствуют одному и тому же колебанию с минимальной длиной волны  $\lambda = 2\pi/k = 2a$ . При такой длине волны соседние атомы цепочки движутся в противофазе. Ясно, что короче длина волны быть уже не может.

График зависимости  $\omega(k)$  для одномерной цепочки с одним атомом в примитивной ячейке изображен на рис. 3.2.

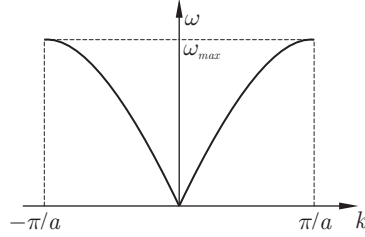


Рис. 3.2. Закон дисперсии колебаний цепочки с одним атомом в примитивной ячейке

Обсудим теперь особенности закона дисперсии (3.12).

Важным его свойством является то, что частота волн, распространяющихся по цепочке, ограничена частотой  $\omega_{\max} = \sqrt{4\gamma/M}$ . Чтобы оценить эту частоту, необходимо знать порядок величины постоянной  $\gamma$ .

Сила  $F$  равна произведению  $\gamma$  на смещение  $u$ , поэтому

$$[\gamma] = \frac{[F]}{[u]} = \frac{\text{энергия}}{\text{длина}^2}. \quad (3.15)$$

Характерная длина — межатомное расстояние  $a$  имеет порядок  $1 \text{ \AA} = 10^{-8} \text{ см}$ . Характерная энергия — это энергия, которую приобретает атом при смещении на расстояние порядка  $a$ . Ее можно оценить как энергию химической связи, которая по порядку величины равна 10 эВ. Таким образом,

$$\gamma \sim \frac{10^{-12} \text{ эрг}}{10^{-16} \text{ см}^2} = 10^4 \frac{\text{эрг}}{\text{см}^2}. \quad (3.16)$$

В качестве массы для оценки можно подставить величину  $10M_p$ , где  $M_p \approx 2 \cdot 10^{-24} \text{ г}$  — масса протона.

Для  $\omega_{\max}$  получаем

$$\omega_{\max} \sim \sqrt{\frac{4 \cdot 1,6 \cdot 10^4}{2 \cdot 10^{-23}}} \approx 6 \cdot 10^{13} \text{ с}^{-1}, \quad \nu = \frac{\omega}{2\pi} \sim 10^{13} \text{ Гц}. \quad (3.17)$$

Найдем длину волны электромагнитного излучения на этой частоте

$$\lambda = \frac{c}{\nu} \approx \frac{3 \cdot 10^{10}}{10^{13}} = 3 \cdot 10^{-3} \text{ см} = 30 \text{ мкм}. \quad (3.18)$$

Электромагнитные волны с такой длиной волны принадлежат инфракрасному диапазону.

При  $ka/2 \ll 1$ , когда длина волны  $\lambda = 2\pi/k$  много больше  $a$ ,  $\sin(ka/2) \approx ka/2$ , поэтому в согласии с (3.4)

$$\omega(k) \approx \sqrt{\frac{4\gamma}{M}} \left| \frac{ka}{2} \right| = a \sqrt{\frac{\gamma}{M}} |k| = s|k|, \quad s = a \sqrt{\frac{\gamma}{M}}. \quad (3.19)$$

Таким образом, длинноволновые колебания — это звуковые волны с линейным законом дисперсии  $\omega = s|k|$ . Выше мы уже получали такой результат, заменив точное уравнение цепочки (3.2) волновым уравнением (3.3). Это и неудивительно: длинные волны «не чувствуют» дискретной структуры цепочки, цепочка ведет себя как непрерывная упругая среда. По этой причине скорость звука  $s$  зависит только от макроскопических характеристик цепочки: от линейной плотности  $M/a$  и упругой постоянной цепочки  $\gamma \cdot a$  — коэффициента пропорциональности между относительным удлинением цепочки и возникающей при этом силой натяжения:

$$s = \sqrt{\frac{\gamma \cdot a}{M/a}}. \quad (3.20)$$

Рассмотренные нами колебания одномерной цепочки называют *акустическими*, поскольку при  $k \rightarrow 0$  ( $\lambda \rightarrow \infty$ ) они соответствуют звуковым волнам.

Ниже мы увидим, что в цепочке с двумя (и более) атомами в элементарной ячейке наряду с акустическими могут распространяться волны и другого типа.

При квантовомеханическом описании каждому колебанию соответствует квазичастица с импульсом  $p = \hbar k$  и энергией  $\mathcal{E} = \hbar \omega$ . Квазичастицы, соответствующие упругим колебаниям кристаллической решетки, называются *фононами*. Фононы, соответствующие акустическим колебаниям, также называются акустическими.

Оценим максимальную энергию акустического фонона в одномерной цепочке:

$$\mathcal{E} = \hbar \omega \sim 10^{-27} \text{ эрг} \cdot \text{с} \cdot 10^{14} \text{ с}^{-1} = 10^{-13} \text{ эрг} \approx 50 \text{ мэВ}. \quad (3.21)$$

Экспериментальные значения  $\hbar \omega_{\max}$  в реальных кристаллах составляют 30–40 мэВ.

Эта величина намного меньше большинства характерных электронных энергий ( $\sim 1 \text{ эВ}$ ) и близка к тепловой энергии при комнатной температуре ( $k_B T \approx 25 \text{ мэВ}$ , здесь  $k_B$  — постоянная Больцмана).

### 3.3. Одномерная цепочка с двумя атомами в примитивной ячейке

Исследуем теперь колебания цепочки, элементарная ячейка которой состоит из двух атомов с разными массами:  $M_1$  и  $M_2$ , для определенности положим  $M_1 < M_2$ . Период цепочки (расстояние между узлами ее решетки Браве), как и прежде, обозначим через  $a$  (рис. 3.3). Для простоты будем считать, что «пружинки», соединяющие атомы, имеют одинаковую жесткость  $\gamma$ .



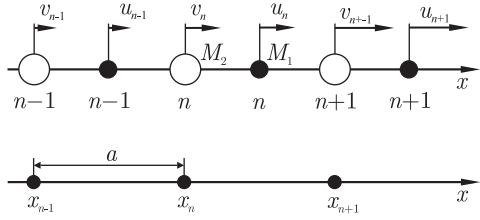


Рис. 3.3. Одномерная цепочка с двумя атомами в примитивной ячейке и ее решетка Браве

Запишем закон Ньютона для двух атомов  $n$ -й ячейки:

$$M_1 \frac{d^2 u_n}{dt^2} = \gamma(v_{n+1} - u_n) - \gamma(u_n - v_n) = \gamma(v_{n+1} + v_n - 2u_n), \quad (3.22)$$

$$M_2 \frac{d^2 v_n}{dt^2} = \gamma(u_n - v_n) - \gamma(v_n - u_{n-1}) = \gamma(u_n + u_{n-1} - 2v_n).$$

Здесь  $u_n$  и  $v_n$  — смещения более легкого и более тяжелого атомов в  $n$ -й ячейке из положения равновесия соответственно.

Будем, как и в случае цепочки с одним атомом в примитивной ячейке, искать решение в виде плоской гармонической волны

$$\begin{aligned} u_n &= A e^{i(kx_n - \omega t)}, \\ v_n &= B e^{i(kx_n - \omega t)}. \end{aligned} \quad (3.23)$$

Амплитуды колебаний легкого и тяжелого атомов  $A$  и  $B$  в общем случае являются разными как по абсолютной величине, так и по фазе.

После подстановки (3.23) в (3.22) получим линейную однородную систему уравнений для  $A$  и  $B$

$$\begin{aligned} -M_1 \omega^2 A &= \gamma(B e^{ika} + B - 2A), \\ -M_2 \omega^2 B &= \gamma(A + A e^{-ika} - 2B). \end{aligned} \quad (3.24)$$

Перепишем ее в стандартном виде:

$$\begin{aligned} (2\gamma - M_1 \omega^2)A - \gamma(e^{ika} + 1)B &= 0, \\ \gamma(e^{-ika} + 1)A + (2\gamma - M_2 \omega^2)B &= 0. \end{aligned} \quad (3.25)$$

Такая система имеет ненулевые решения лишь в том случае, когда ее определитель равен нулю. Приравняв нулю определитель (3.25), получим уравнение, связывающее  $\omega$  и  $k$ , то есть дисперсионное уравнение

$$M_1 M_2 \omega^4 - 2\gamma(M_1 + M_2)\omega^2 + 2\gamma^2(1 - \cos ka) = 0. \quad (3.26)$$

Это уравнение удобно переписать, используя приведенную массу атомов примитивной ячейки  $\mu$ :

$$\frac{1}{\mu} = \frac{1}{M_2} + \frac{1}{M_1} = \frac{M_2 + M_1}{M_1 M_2}, \quad (3.27)$$

как

$$\omega^4 - 2\frac{\gamma}{\mu}\omega^2 + \frac{4\gamma^2}{M_1 M_2} \sin^2 \frac{ka}{2} = 0. \quad (3.28)$$

Его решения имеют вид

$$\omega^2 = \frac{\gamma}{\mu} \pm \sqrt{\frac{\gamma^2}{\mu^2} - \frac{4\gamma^2}{M_1 M_2} \sin^2 \frac{ka}{2}}, \quad (3.29)$$

или

$$\omega^2 = \frac{\gamma}{\mu} \left( 1 \pm \sqrt{1 - \frac{4\mu^2}{M_1 M_2} \sin^2 \frac{ka}{2}} \right). \quad (3.30)$$

Величина  $4\mu^2/(M_1 M_2)$  при любых  $M_1, M_2$  не превосходит единицы, поэтому подкоренное выражение всегда неотрицательно.

Итак, для каждого волнового вектора  $k$  существуют две частоты  $\omega$ , удовлетворяющие дисперсионному уравнению. Точнее, есть две непрерывные функции  $\omega(k)$ , которые отличаются знаком перед корнем. Говорят, что существуют две ветви колебаний.

Исследуем обе ветви.

Напомним, что волновые векторы, отличающиеся на вектор обратной решетки, описывают одно и то же колебание. (Вследствие этого функция  $\omega(k)$  периодична с периодом обратной решетки  $2\pi/a$ , а в трехмерном случае обладает трансляционной симметрией обратной решетки). Поэтому мы считаем, что волновой вектор лежит в пределах первой зоны Бриллюэна  $-\pi/a < k < \pi/a$ .

**Решение со знаком «-».** В точке  $k = 0$

$$\omega^2(0) = \frac{\gamma}{\mu} - \sqrt{\frac{\gamma^2}{\mu^2}} = 0. \quad (3.31)$$

На границе зоны Бриллюэна

$$\omega^2\left(\frac{\pi}{a}\right) = \frac{\gamma}{\mu} - \sqrt{\frac{\gamma^2}{\mu^2} - \frac{4\gamma^2}{M_1 M_2}} = \frac{2\gamma}{M_2}. \quad (3.32)$$

При  $ka \ll 1$  (длинные волны)

$$\begin{aligned} \omega^2(k) &\approx \frac{\gamma}{\mu} \left[ 1 - \sqrt{1 - \frac{4\mu^2}{M_1 M_2} \left(\frac{ka}{2}\right)^2} \right] \approx \frac{\gamma}{\mu} \left[ 1 - \left( 1 - \frac{2\mu^2}{M_1 M_2} \left(\frac{ka}{2}\right)^2 \right) \right] = \\ &= \frac{2\mu\gamma}{M_1 M_2} \left(\frac{ka}{2}\right)^2 = a^2 \frac{\gamma}{2(M_1 + M_2)} k^2, \end{aligned} \quad (3.33)$$



другими словами,

$$\omega(k) \approx a \sqrt{\frac{\gamma}{2(M_1 + M_2)}} |k| = s |k|, \quad s = a \sqrt{\frac{\gamma}{2(M_1 + M_2)}}. \quad (3.34)$$

Мы видим, что в длинноволновом пределе закон дисперсии этой ветви линеен, то есть, как и в случае цепочки с одним атомом в примитивной ячейке, описывает акустические колебания. По этой причине вся ветвь (решение со знаком «-») называется акустической ветвью (рис. 3.4).

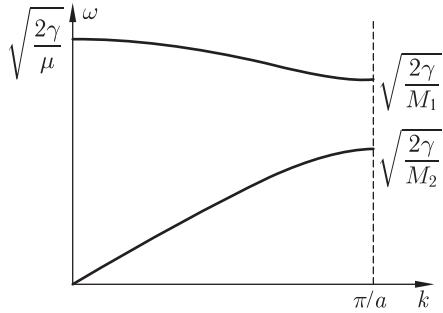


Рис. 3.4. Закон дисперсии колебаний цепочки с двумя атомами в примитивной ячейке

Выражение для скорости звука имеет такой же вид, что и соответствующее выражение для цепочки с одним атомом в ячейке (3.20), и зависит от тех же макроскопических характеристик: линейной плотности и упругой постоянной цепочки:

$$s = \sqrt{\frac{\gamma \cdot a/2}{(M_1 + M_2)/a}}. \quad (3.35)$$

Линейная плотность двухатомной цепочки равна  $(M_1 + M_2)/a$ , а упругая постоянная —  $\gamma \cdot a/2$  (так как длина одной пружинки в наших обозначениях равна  $a/2$ ).

Это и неудивительно. Изучая цепочку с одним атомом в примитивной ячейке, мы уже видели, что длинноволновые акустические колебания можно получить, рассматривая цепочку как непрерывную упругую среду. Атомы ячейки при таких колебаниях движутся вместе, как единое целое, поэтому структура примитивной ячейки не играет роли, а важны лишь макроскопические, усредненные характеристики цепочки.

Тот факт, что атомы ячейки при длинноволновых акустических колебаниях движутся вместе, можно получить и непосредственно, решая систему (3.25). Эта система разрешима, когда ее определитель равен нулю, определитель же равен нулю, когда  $\omega$  и  $k$  связаны законом дисперсии. При этом уравнения системы уже не являются независимыми, и мы можем взять любое из них, чтобы найти отношение амплитуд  $A$  и  $B$ .

Из первого уравнения системы (3.25) получаем

$$\frac{B}{A} = \frac{2\gamma - M_1\omega^2}{\gamma(e^{ika} + 1)}, \quad (3.36)$$



Рис. 3.5. Амплитуды колебаний атомов цепочки в случае длинноволновых акустических колебаний

откуда в пределе длинноволновых акустических колебаний ( $k \rightarrow 0$ ,  $\omega = s|k| \rightarrow 0$ ) следует, что  $B/A \rightarrow 1$ , то есть  $A = B$ : это значит, что атомы движутся в фазе с одинаковыми амплитудами (рис. 3.5).

**Решение со знаком «+».** В точке  $k = 0$

$$\omega^2(0) = \frac{\gamma}{\mu} + \sqrt{\frac{\gamma^2}{\mu^2}} = \frac{2\gamma}{\mu}. \quad (3.37)$$

На границе зоны Бриллюэна

$$\omega^2\left(\frac{\pi}{a}\right) = \frac{\gamma}{\mu} + \sqrt{\frac{\gamma^2}{\mu^2} - \frac{4\gamma^2}{M_1 M_2}} = \frac{2\gamma}{M_1}. \quad (3.38)$$

Групповая скорость этой ветви  $\partial\omega/\partial k$  равна нулю как на границе зоны Бриллюэна, так и при  $k = 0$ .

Эта ветвь целиком лежит выше акустической ветви: ее минимальная частота  $\sqrt{2\gamma/M_1}$  больше максимальной частоты акустических колебаний  $\sqrt{2\gamma/M_2}$ . Таким образом, в цепочке могут распространяться волны с частотами от 0 до  $\sqrt{2\gamma/M_2}$  и от  $\sqrt{2\gamma/M_1}$  до  $\sqrt{2\gamma/\mu}$ . Интервал частот  $(\sqrt{2\gamma/M_2}, \sqrt{2\gamma/M_1})$  является «запрещенной зоной»: волн с такими частотами не существует. Относительная ширина этого интервала тем больше, чем больше отношение масс  $M_2/M_1$ .

Чтобы понять, что представляют собой длинноволновые колебания этой ветви, найдем отношение амплитуд колебаний  $B/A$  при  $k = 0$  с помощью (3.36):

$$\frac{B}{A} = \frac{2\gamma - M_1\omega^2}{\gamma(e^{ika} + 1)} = \frac{2\gamma - M_1(2\gamma/\mu)}{2\gamma} = -\frac{M_1}{M_2}. \quad (3.39)$$

Мы видим, что атомы в каждой ячейке движутся в противофазе, то сближаясь, то удаляясь друг от друга, причем одновременно во всех ячейках (если  $k = 0$ ). Амплитуда движения легкого атома больше амплитуды тяжелого в  $M_2/M_1$  раз, то есть центр тяжести ячейки остается на месте (рис. 3.6).



Рис. 3.6. Амплитуды атомов цепочки в случае длинноволновых оптических колебаний

Если атомы заряжены, то при колебаниях такого типа каждая ячейка представляет собой переменный дипольный момент. Дипольные мо-

менты взаимодействуют с электромагнитным полем, и колебания легко возбуждаются электромагнитными волнами соответствующих частот. В связи с этим, вся ветвь колебаний называется оптической.

При длинноволновых акустических колебаниях атомы ячейки движутся в фазе и никакого дипольного момента не возникает. Поэтому акустические колебания взаимодействуют с электромагнитным полем слабо.

Энергия длинноволнового оптического фонона имеет тот же порядок величины, что и энергия фонона акустического колебания с максимальной частотой, которую мы оценили в 0,05 эВ. Энергии оптических фононов большинства полупроводниковых кристаллов лежат в диапазоне 0,03–0,1 эВ.

Рассмотрим теперь, как колеблются атомы, когда длина волны минимальна, то есть когда волновой вектор лежит на границе зоны Бриллюэна.

В случае акустических колебаний  $\omega^2 = 2\gamma/M_2$  коэффициент при  $B$  во втором уравнении системы (3.25) обращается в нуль, откуда следует, что  $A = 0$ .

В случае оптических колебаний  $\omega^2 = 2\gamma/M_1$ , и из первого уравнения (3.25) следует, что  $B = 0$ .

Таким образом, в случае акустических волн при  $k = \pi/a$  колеблются тяжелые атомы, а легкие неподвижны, в случае же оптических, наоборот, колеблются легкие, а тяжелые стоят на месте. Отметим также, что на границе зоны Бриллюэна групповая скорость  $\partial\omega/\partial k$  равна нулю. Это утверждение справедливо для обеих ветвей колебаний.

Обобщим теперь полученные результаты. Нетрудно показать, что если примитивная ячейка одномерной цепочки содержит  $l$  атомов, то спектр колебаний состоит из  $l$  ветвей, одна из которых акустическая, а остальные — оптические.

Мы рассматривали бесконечную цепочку, не накладывая никаких ограничений на длины волн упругих колебаний. В результате мы пришли к выводу, что в цепочке могут распространяться колебания с любыми волновыми векторами, лежащими в первой зоне Бриллюэна. (Было показано, что из-за дискретности цепочки волновые векторы, отличающиеся на произвольный вектор обратной решетки, описывают одни и те же колебания. Поэтому можно брать волновой вектор из любой зоны Бриллюэна. Естественней всего описывать колебание наименьшим волновым вектором, то есть вектором из первой зоны Бриллюэна.)

Чтобы иметь дело не с непрерывным, а с дискретным набором волновых векторов, можно потребовать, чтобы отклонение атомов от равновесия было периодической функцией  $u(x_n) = u(x_n + L)$ , иными словами, поставить граничные условия Борна–Кармана. Период  $L$  должен быть кратным постоянной решетки цепочки.

Условиям Борна–Кармана удовлетворяют только гармонические колебания с «разрешенными» волновыми векторами  $k_n = 2\pi n/L$ . Нетруд-

но подсчитать, что в зоне Бриллюэна размещается  $L/a$  разрешенных волновых векторов, то есть ровно столько, сколько примитивных ячеек укладывается на длине  $L$ . (Волновым векторам  $-\pi/a$  и  $\pi/a$  соответствует одно и то же колебание, и поэтому считаем эти два значения за одно.) Мы уже упоминали выше об этом свойстве зоны Бриллюэна.

Так как колебание однозначно определяется волновым вектором и ветвью, то различных колебаний столько, сколько атомов содержит цепочка. Это общее свойство линейных колебательных систем: количество независимых колебаний (нормальных мод) равно числу степеней свободы системы.

### 3.4. Трехмерный кристалл

Мы рассмотрели колебания в одномерной цепочке. Подобным же образом могут быть описаны и колебания решетки трехмерного кристалла.

Предположим, что примитивная ячейка кристалла состоит из  $l$  атомов. Каждый атом ячейки будем обозначать индексом  $\nu$ , который принимает  $l$  различных значений. Любой атом кристалла однозначно определяется радиус-вектором  $\mathbf{r}_n$ , который задает положение ячейки (соответствующего узла решетки Браве), и индексом  $\nu$ , который характеризует положение атома внутри ячейки (тип атома).

Этот радиус-вектор пробегает значения

$$\mathbf{r}_n = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3, \quad (3.40)$$

где  $n_1, n_2, n_3$  — целые числа, а  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  — основные периоды решетки.

Обозначим смещения атомов при колебаниях посредством  $\mathbf{U}_\nu(n)$ , где индекс  $\nu$  указывает номер атома в примитивной ячейке ( $\nu = 1, 2, \dots, l$ ). Функция Лагранжа кристаллической решетки как механической системы частиц, совершающих малые колебания вокруг своих положений равновесия, в рамках гармонического приближения имеет вид

$$L = \frac{1}{2} \sum_{n\nu} M_\nu \dot{\mathbf{U}}_\nu^2(n) - \frac{1}{2} \sum_{\substack{n\nu \\ n'\nu'}} \Lambda_{ik}^{\nu\nu'}(\mathbf{n} - \mathbf{n}') U_{\nu i}(\mathbf{n}) U_{\nu' k}(\mathbf{n}'). \quad (3.41)$$

Здесь  $\mathbf{n} = (n_1, n_2, n_3)$ ,  $M_\nu$  — массы атомов;  $i, k$  — векторные индексы, пробегающие значения  $x, y, z$  (напомним, что под дважды повторяющимся индексом подразумевается суммирование).  $\Lambda_{ik}^{\nu\nu'}$  — это силовые постоянные, характеризующие взаимодействие между атомами. Коэффициенты  $\Lambda$  зависят только от разностей  $\mathbf{n} - \mathbf{n}'$ , поскольку силы взаимодействия атомов, благодаря трансляционной симметрии кристалла, могут зависеть лишь от относительного положения атомов в пространстве. Матрица силовых постоянных  $\Lambda_{ik}^{\nu\nu'}$  представляет собой вторую

производную от потенциальной энергии по смещениям. Следовательно, она обладает следующим свойством симметрии:

$$\Lambda_{ik}^{\nu\nu'}(\mathbf{n}) = \Lambda_{ki}^{\nu\nu'}(-\mathbf{n}). \quad (3.42)$$

Это свойство вытекает непосредственно из вида функции Лагранжа (3.41). Важно отметить, что коэффициенты  $\Lambda$  связаны друг с другом определенными соотношениями, выражающими тот факт, что при операциях групп симметрии кристалла (параллельном переносе решетки или при повороте решетки как целого) не возникает никаких действующих на атомы сил. Например, при параллельном переносе все смещения  $\mathbf{U}_\nu(\mathbf{n}) = \text{const}$ , и поэтому должно быть

$$\sum_{\mathbf{n}\nu'} \Lambda_{ik}^{\nu\nu'}(\mathbf{n}) = 0. \quad (3.43)$$

Зная функцию Лагранжа (3.41), нетрудно записать уравнения движения, описывающие малые колебания атомов <sup>1)</sup>,

$$M_\nu \ddot{U}_{\nu i} = - \sum_{\mathbf{n}'\nu'} \Lambda_{ik}^{\nu\nu'}(\mathbf{n} - \mathbf{n}') U_{\nu'k}(\mathbf{n}'). \quad (3.44)$$

Правая часть (3.44) — это сила  $\mathbf{F}_{\mathbf{n}\nu}$ , действующая на  $\nu$ -й атом в  $\mathbf{n}$ -й элементарной ячейке при смещении атомов из положения равновесия.

Будем искать решения уравнений (3.44) в виде монохроматической плоской волны

$$\mathbf{U}_\nu(\mathbf{n}, t) = \mathbf{e}_{\nu\alpha}(\mathbf{k}) \exp[i(\mathbf{k}\mathbf{r}_n - \omega_j(\mathbf{k})t)]. \quad (3.45)$$

Частота колебаний одинакова для всех атомов кристалла. Амплитуда колебаний  $\mathbf{e}_{\nu\alpha}$  (она комплексная) зависит от типа атома (индекса  $\nu$ ), то есть она различна лишь для разных атомов в одной и той же ячейке, но одинакова для всех однотипных атомов. Векторы  $\mathbf{e}_{\nu\alpha}$  определяют как величину амплитуды колебаний, так и их направление (поляризацию). Индекс  $j$  обозначает ветвь колебаний.

Подставив (3.45) в (3.44), получим

$$\omega^2 M_\nu e_{\nu\alpha i} \exp(i\mathbf{k}\mathbf{r}_n) = \sum_{\mathbf{n}'\nu'} \Lambda_{ik}^{\nu\nu'}(\mathbf{n} - \mathbf{n}') e_{\nu' i} \exp(i\mathbf{k}\mathbf{r}_{n'}). \quad (3.46)$$

Разделив обе части равенства на  $\exp(i\mathbf{k}\mathbf{r}_n)$  и заменив суммирование по  $\mathbf{n}'$  суммированием по  $(\mathbf{n} - \mathbf{n}')$ , находим

$$\sum_{\nu'} \Lambda_{ik}^{\nu\nu'}(\mathbf{k}) e_{\nu' k} - \omega^2 M_\nu e_{\nu i} = 0. \quad (3.47)$$

Здесь введено обозначение

$$\Lambda_{ik}^{\nu\nu'}(\mathbf{k}) = \sum_{\mathbf{n}'} \Lambda_{ik}^{\nu\nu'}(\mathbf{n}) \exp(-i\mathbf{k}\mathbf{r}_n). \quad (3.48)$$

<sup>1)</sup> См. Ландау Л. Д., Лифшиц Е. М. Механика. — М.: Наука, 1988.

Система (3.47) линейных однородных алгебраических уравнений для амплитуд имеет отличные от нуля решения, только если определитель равен нулю:

$$\det |\omega^2 M_\nu \delta_{\nu\nu'} \delta_{ik} - \Lambda_{ik}^{\nu\nu'}(\mathbf{k})| = 0. \quad (3.49)$$

Поскольку векторные индексы  $i, k$  пробегают по три значения, а индексы  $\nu, \nu'$  — по  $l$  значений, то порядок определителя равен  $3l$ . Таким образом, (3.49) — это алгебраическое уравнение степени  $3l$  относительно  $\omega^2$ . Каждое из  $3l$  решений этого уравнения определяет частоту  $\omega$  как функцию волнового вектора  $\mathbf{k}$ , то есть закон дисперсии волн, а уравнение (3.49) называется дисперсионным уравнением. Таким образом, для каждого заданного значения волнового вектора  $\mathbf{k}$  частота  $\omega$  может иметь в общем случае  $3l$  различных значений. Можно сказать, что частота есть многозначная функция волнового вектора, обладающая  $3l$  ветвями:  $\omega = \omega_j(\mathbf{k})$ , где индекс  $j$ , как уже было отмечено выше, нумерует ветви колебаний.

В силу симметрии механических уравнений движения (3.44) по отношению к изменению знака времени, если возможно распространение некоторой волны (3.45), то возможно распространение такой же волны и в противоположном направлении. Но такое изменение направления распространения эквивалентно изменению знака волнового вектора  $\mathbf{k}$ . Следовательно, функция  $\omega(\mathbf{k})$  должна быть четной:

$$\omega(\mathbf{k}) = \omega(-\mathbf{k}). \quad (3.50)$$

Волновой вектор колебаний решетки обладает следующим важным свойством. Вектор  $\mathbf{k}$  входит в выражение (3.45) только через экспоненциальный множитель  $e^{i\mathbf{k}\mathbf{r}_n}$ , который не меняется при замене

$$\mathbf{k} \rightarrow \mathbf{k} + \mathbf{b}, \quad \mathbf{b} = p_1 \mathbf{b}_1 + p_2 \mathbf{b}_2 + p_3 \mathbf{b}_3, \quad (3.51)$$

где  $\mathbf{b}$  — любой вектор обратной решетки ( $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$  — ее основные периоды,  $p_1, p_2, p_3$  — целые числа). Таким образом, волновой вектор колебаний решетки физически неоднозначен: значения  $\mathbf{k}$ , отличающиеся на вектор  $\mathbf{b}$ , физически эквивалентны. В обратной решетке функция  $\omega(\mathbf{k})$  периодична:

$$\omega(\mathbf{k} + \mathbf{b}) = \omega(\mathbf{k}). \quad (3.52)$$

Поэтому для каждой ветви колебаний достаточно рассматривать значения вектора  $\mathbf{k}$ , лежащие в некотором конечном интервале в первой зоне Бриллюэна. Если выбрать оси координат по трем основным периодам обратной решетки, то можно ограничиться областью

$$-(1/2)b_i < k_i < (1/2)b_i. \quad (3.53)$$

Среди  $3l$  ветвей спектра колебаний решетки должны быть такие, которые при больших (по сравнению с постоянной решетки) длинах волн соответствуют обычным упругим (то есть звуковым) волнам в кристалле. Как известно из теории упругости, в кристалле, рассматриваемом как сплошная среда, могут распространяться волны трех

типов с различными законами дисперсии, причем для всех трех типов  $\omega$  описывается однородной функцией первого порядка от компонент вектора  $\mathbf{k}$ , которая обращается в нуль при  $\mathbf{k} = 0$ <sup>1)</sup>. Следовательно, среди  $3l$  ветвей функции  $\omega(\mathbf{k})$  должны существовать три, в которых при малых  $\mathbf{k}$  закон дисперсии имеет вид

$$\omega = s|\mathbf{k}| \equiv sk, \quad (3.54)$$

где  $s$  — скорость звука в кристалле. Эти три типа волн называются акустическими, они характеризуются тем, что (при малых  $\mathbf{k}$ ) решетка колеблется в целом как сплошная среда. В пределе  $\mathbf{k} \rightarrow 0$  эти малые колебания переходят в простое параллельное смещение всей решетки. Следует заметить, что скорость звука  $s$  зависит от направления распространения волны, то есть от направления  $\mathbf{k}$ . В случае длинноволновых акустических колебаний амплитуды всех атомов примитивной ячейки примерно одинаковы.

В сложных решетках, содержащих более одного атома в ячейке, существуют еще  $3l - 3$  типа волн. В этих ветвях спектра частота не обращается в нуль при  $\mathbf{k} = 0$  и стремится при  $\mathbf{k} \rightarrow 0$  к постоянному пределу. Эти колебания решетки называют оптическими. В этом случае атомы в каждой ячейке движутся друг относительно друга, причем в предельном случае  $\mathbf{k} = 0$  центр тяжести ячейки остается в покое.

Не все  $3l - 3$  предельные частоты оптических колебаний (частоты при  $\mathbf{k} = 0$ ) должны быть различными. При определенных свойствах симметрии кристалла предельные частоты некоторых из оптических ветвей спектра могут совпадать или, как говорят, быть вырожденными.

Функция  $\omega(\mathbf{k})$  с невырожденной предельной частотой может быть разложена вблизи  $\mathbf{k} = 0$  в ряд по степеням компонент вектора  $\mathbf{k}$ . В силу четности функции  $\omega(\mathbf{k})$  такое разложение может содержать только четные степени  $k_i$ , так что его первые члены имеют вид

$$\omega = \omega_0 + \frac{1}{2}\gamma_{ik}k_ik_k, \quad (3.55)$$

где  $\omega_0$  — предельная частота,  $\gamma_{ik}$  — постоянные величины.

Если же предельные частоты нескольких ветвей совпадают, то функции  $\omega(\mathbf{k})$  в этих ветвях вообще не могут быть разложены по степеням  $\mathbf{k}$ , поскольку точка  $\mathbf{k} = 0$  является для них особой точкой (точкой ветвления). Можно лишь утверждать, что вблизи  $\mathbf{k} = 0$  разность  $\omega - \omega_0$  будет (в зависимости от симметрии кристалла) однородной функцией компонент  $\mathbf{k}$  либо первого, либо второго порядка.

По направлению смещений атомов относительно волнового вектора акустические колебания можно разделить на продольное (ЛА) и два поперечных (ТА). Строго говоря, смещения атомов при этих колебаниях параллельны или перпендикулярны вектору  $\mathbf{k}$  только при распространении волны в направлениях с высокой симметрией, например [100]

для кристаллов кубической сингонии. Как правило, скорость звука у продольного колебания больше, чем у поперечных.

У кристаллов со структурой алмаза или цинковой обманки примитивная ячейка содержит два атома. Соответственно, кроме трех акустических ветвей колебаний, эти кристаллы обладают тремя оптическими ветвями, из которых также можно выделить продольную (ЛО) и две поперечных (ТО) ветви.

Количество разрешенных волновых векторов в зоне Бриллюэна равно  $N = V/v_0$ , то есть числу примитивных ячеек в нормировочном объеме кристалла  $V = L^3$  ( $v_0$  — объем примитивной ячейки). Действительно, плотность разрешенных волновых векторов в обратном пространстве равна  $V/(2\pi)^3$ , таким образом, в объеме обратного пространства  $\Delta^3k$  содержится  $\Delta^3k \cdot V/(2\pi)^3$  разрешенных волновых векторов. Объем зоны Бриллюэна — объем примитивной ячейки обратной решетки — равен  $(2\pi)^3/v_0$ , и для числа разрешенных состояний получаем  $(2\pi)^3/v_0 \cdot V/(2\pi)^3 = V/v_0 = N$ .

Число ветвей равно  $3l$ , поэтому полное число колебаний равно  $3lN$  — утроенному числу атомов кристалла в объеме  $L^3$ , то есть числу степеней свободы механической системы.

**Замечание.** Следует еще раз отметить, что изложенная выше теория колебаний кристаллической решетки основана на так называемом гармоническом приближении, в котором учитываются лишь квадратичные по смещениям атомов члены в потенциальной энергии. Только в этом приближении различные монохроматические волны (3.45) не взаимодействуют друг с другом, но свободно распространяются по решетке. При учете же следующих, ангармонических, членов появляются различного рода процессы распада и рассеяния этих волн друг на друге.

### 3.5. Фононы

Рассмотрим вопрос о том, как выглядит картина колебаний кристаллической решетки с точки зрения квантовой теории<sup>1)</sup>.

В квантовой механике вместо волн (3.45), в которых атомы испытывают определенные смещения, вводится понятие о *фононах* как о некоторых квазичастицах, распространяющихся по решетке и обладающих определенными энергиями и направлениями движения (см. Приложение в конце главы). Согласно квантовой теории фонон (в рамках гармонического приближения) представляет собой линейный гармонический осциллятор.

Поскольку энергия осциллятора есть целое кратное от  $\hbar\omega_{\alpha\mathbf{k}}$  (где  $\omega_{\alpha\mathbf{k}}$  — частота классической волны), то энергия фонона связана с частотой  $\omega_{\alpha\mathbf{k}}$  соотношением

$$\varepsilon_{\alpha}(\mathbf{k}) = \hbar\omega_{\alpha\mathbf{k}}. \quad (3.56)$$

<sup>1)</sup> См. Ландау Л. Д., Лифшиц Е. М. Теория упругости. — М.: Наука, 1987.

<sup>1)</sup> Детали изложены в Приложении к § 4 (см. конец настоящей главы).

Подобное же соотношение имеет место для световых квантов — фотонов. Что же касается волнового вектора  $\mathbf{k}$ , то он определяет квазиимпульс фонона  $\mathbf{p}$ :

$$\mathbf{p} = \hbar \mathbf{k}. \quad (3.57)$$

Это величина во многом аналогична обычному импульсу. В то же время между ними имеется существенное отличие, связанное с тем, что квазиимпульс является величиной, определенной лишь с точностью до постоянного вектора вида  $\hbar \mathbf{b}$ . Значения  $\mathbf{p}$ , отличающиеся на такую величину, физически эквивалентны.

Скорость фонона определяется групповой скоростью соответствующих классических волн

$$\mathbf{v} = \frac{\partial \omega}{\partial \mathbf{k}}. \quad (3.58)$$

Согласно (3.56) и (3.57) выражения для групповой скорости можно записать в виде

$$\mathbf{v} = \frac{\partial \varepsilon(\mathbf{p})}{\partial \mathbf{p}}. \quad (3.59)$$

Эта формула полностью аналогична обычному соотношению между энергией и импульсом частиц (см. главы 4 и 6).

В кристаллической решетке может быть возбуждено одновременно сколько угодно одинаковых фононов<sup>1)</sup>. Другими словами, в каждом квантовом состоянии фононов может находиться любое их число (в классической картине этому случаю отвечает произвольная интенсивность волн). Это значит, что фононный газ подчиняется статистике Бозе–Эйнштейна. Фононный газ обладает еще одной существенной особенностью — число частиц  $N$  в нем является переменной величиной, а не заданной постоянной, как в обычном газе (здесь имеется полная аналогия с фотонным газом).

Таким образом, число фононов  $N$  должно само определяться из условий теплового равновесия. Тогда, потребовав выполнения условия минимальности свободной энергии газа  $F$  (при заданных температуре  $T$  и объеме  $V$ ), получим в качестве одного из необходимых условий

$$\frac{\partial F}{\partial N} = 0. \quad (3.60)$$

Но поскольку

$$\left( \frac{\partial F}{\partial N} \right)_{T,V} = \mu, \quad (3.61)$$

то химический потенциал газа фононов будет равен нулю,

$$\mu = 0. \quad (3.62)$$

<sup>1)</sup> Количество фононов неограничено лишь в линейном, гармоническом приближении.

Поэтому в тепловом равновесии среднее число фононов в каждом квантовом состоянии (с квазиимпульсом  $\mathbf{p}$  и энергией  $\varepsilon_\alpha(\mathbf{k})$ ) определяется функцией распределения Планка

$$\bar{n}_{\alpha\mathbf{k}} = \bar{n}(\varepsilon_\alpha(\mathbf{k})) = \frac{1}{\exp\left(\frac{\varepsilon_\alpha(\mathbf{k})}{k_B T}\right) - 1} = \frac{1}{\exp\left(\frac{\hbar\omega_{\alpha\mathbf{k}}}{k_B T}\right) - 1}. \quad (3.63)$$

Здесь  $k_B$  — постоянная Больцмана. Нормальные колебания решетки ведут себя как набор независимых гармонических осцилляторов. Роль координаты осциллятора играет при этом амплитуда колебания, числом фононов определяется уровень энергии осциллятора.

На каждое колебание приходится средняя энергия  $\mathcal{E}_\alpha(\mathbf{k})\bar{n}_{\alpha\mathbf{k}} = \hbar\omega_{\alpha\mathbf{k}}\bar{n}_{\alpha\mathbf{k}}$ . Строго говоря, к этой энергии надо прибавить энергию основного состояния колебаний (энергию нулевых колебаний). Как и у обычного гармонического осциллятора, она равна  $\frac{1}{2}\hbar\omega_{\alpha\mathbf{k}}$ . Но энергией нулевых колебаний кристалл обладает всегда, и мы просто примем ее за начало отсчета.

При высоких температурах, когда  $k_B T \gg \hbar\omega$ , число фононов пропорционально температуре,

$$\bar{n}_{\alpha\mathbf{k}} \approx \frac{k_B T}{\hbar\omega_{\alpha\mathbf{k}}}. \quad (3.64)$$

Средняя энергия колебания при этом равна  $k_B T$ . Это известный результат классической статистической механики для средней энергии гармонического осциллятора. Таким образом, пока температура превосходит энергию фонона, квантовые эффекты не играют роли.

Они играют существенную роль при низких температурах. Если  $k_B T \ll \hbar\omega$ , то среднее число фононов экспоненциально мало:

$$\bar{n}_\alpha(\mathbf{k}) \approx \exp\left(-\frac{\hbar\omega_{\alpha\mathbf{k}}}{k_B T}\right) \ll 1. \quad (3.65)$$

Можно сказать, что колебания, частота которых превосходит величину  $k_B T/\hbar$ , «вымерзают». Энергия колебания не может быть меньше энергии одного фонона  $\hbar\omega_{\alpha\mathbf{k}}$ , а энергия фонона намного превышает характерную тепловую энергию  $k_B T$ , поэтому такие колебания практически не возбуждаются.

### 3.6. Энергия колебаний и теплоемкость кристаллической решетки

Энергию колебаний и теплоемкость решетки будем рассчитывать для единичного объема кристалла, то есть положим нормировочный объем равным единице:  $V = L^3 = 1$ .

Чтобы вычислить полную энергию колебаний кристаллической решетки, нужно просуммировать среднюю энергию всех типов колебаний (всех состояний фононов):

$$E = \sum_{\alpha, \mathbf{k}} \mathcal{E}_{\alpha}(\mathbf{k}) \bar{n}_{\alpha \mathbf{k}} = \sum_{\alpha, \mathbf{k}} \hbar \omega_{\alpha \mathbf{k}} \bar{n}_{\alpha \mathbf{k}} = \sum_{\alpha, \mathbf{k}} \frac{\hbar \omega_{\alpha \mathbf{k}}}{\exp\left(\frac{\hbar \omega_{\alpha \mathbf{k}}}{k_B T}\right) - 1}. \quad (3.66)$$

Проще всего это сделать при высоких температурах, когда для частот всех колебаний выполняется неравенство  $\hbar \omega_{\alpha \mathbf{k}} \ll k_B T$  (классический предел). Тогда средняя энергия, приходящаяся на каждое колебание, равна  $k_B T$ . Всего колебаний  $3lN = 3lN$ , для полной энергии  $E$  получаем

$$E = 3lN k_B T = \frac{3l}{v_0} k_B T. \quad (3.67)$$

Так как  $N$  — это число примитивных ячеек кристалла в единице объема, то  $N = 1/v_0$ , где  $v_0$  — объем примитивной ячейки.

Теплоемкость решетки при высоких температурах постоянна (закон Дюлонга и Пти):

$$C_V = 3lN k_B. \quad (3.68)$$

При низких температурах расчет оказывается сложнее. Чтобы точно вычислить энергию решетки, то есть просуммировать (3.66), необходимо знать дисперсионные зависимости для всех ветвей колебаний. И даже при условии, что зависимости эти известны, аналитическое выражение для энергии получить практически невозможно.

Поэтому для нахождения энергии и теплоемкости решетки применяют различные приближения.

**3.6.1. Модель Эйнштейна.** В модели Эйнштейна предполагается, что частоты всех фононов одинаковы,  $\omega_{\alpha \mathbf{k}} = \omega_1$ . Тогда для энергии получаем

$$E_{\mathcal{E}} = 3lN \frac{\hbar \omega_1}{\exp\left(\frac{\hbar \omega_1}{k_B T}\right) - 1}. \quad (3.69)$$

При высоких температурах  $k_B T \gg \hbar \omega_1$ , эта зависимость приводит к выражению (3.67) для энергии и закону Дюлонга и Пти (3.68) для теплоемкости.

При низких температурах  $k_B T \ll \hbar \omega_1$  энергия колебаний и теплоемкость экспоненциально уменьшаются:

$$E_{\mathcal{E}} = 3lN \hbar \omega_1 \exp\left(-\frac{\hbar \omega_1}{k_B T}\right), \quad (3.70)$$

$$C_V = 3lN \frac{\hbar^2 \omega_1^2}{k_B T^2} \exp\left(-\frac{\hbar \omega_1}{k_B T}\right) = \frac{\hbar \omega_1}{k_B T^2} E_{\mathcal{E}}. \quad (3.71)$$

Модель Эйнштейна хорошо описывает вклад в энергию и теплоемкость оптических ветвей фононов, у которых частота так слабо зависит

от волнового вектора, что ее можно считать постоянной. Чтобы учесть только оптические ветви, частоту которых мы полагаем равной  $\omega_1$ , нужно вместо  $3l$  подставлять число этих ветвей. В общем случае частоты разных оптических ветвей могут сильно отличаться друг от друга и их вклад в энергию и теплоемкость следует учитывать отдельно.

**3.6.2. Модель Дебая.** Опыт показывает, что теплоемкость действительно падает с уменьшением температуры, но не экспоненциально, а пропорционально  $T^3$ . Дело в том, что при любых сколь угодно низких температурах в кристалле найдутся колебания, энергия фонона которых меньше  $k_B T$ . Это длинноволновые акустические колебания, частота которых меньше  $k_B T / \hbar$ . Они вносят основной вклад в энергию при низких температурах. Колебания с большими частотами (оптические и более коротковолновые акустические) «выморожены», и фононов этих колебаний экспоненциально мало.

Произведем простую оценку. Вклад в энергию вносят фононы, энергия которых меньше  $k_B T$ . Пусть скорость звука  $\alpha$ -й акустической ветви равна  $s_{\alpha}$  и не зависит от направления волнового вектора:  $\omega = s_{\alpha} |k|$ . Тогда вклад в энергию дают колебания с волновыми векторами, меньшими  $k_{\max} = k_B T / (\hbar s_{\alpha})$ . Плотность разрешенных значений волновых векторов в  $k$ -пространстве кристалла равна  $V / (2\pi)^3$ , поэтому внутри сферы радиуса  $k_{\max}$  содержится

$$\frac{4\pi}{3} k_{\max}^3 \cdot \frac{V}{(2\pi)^3} = V \frac{k_{\max}^3}{2\pi^2} \cdot \frac{1}{3}$$

разрешенных значений волновых векторов. Это число колебаний одной акустической ветви, которые вносят существенный вклад в энергию. На каждое такое колебание приходится энергия порядка  $k_B T$ . Для энергии колебаний одной акустической ветви получаем

$$E_{\alpha} \sim k_B T \frac{k_{\max}^3}{2\pi^2} = \frac{(k_B T)^4}{2\pi^2 (\hbar s_{\alpha})^3}. \quad (3.72)$$

Поскольку мы вычисляем энергию и теплоемкость единицы объема кристалла, то в (3.72) мы положили  $V = 1$ .

Таким образом, вклад одной акустической ветви в теплоемкость пропорционален  $T^3$ :

$$C_{V\alpha} = \frac{dE}{dT} \sim \frac{2k_B^4 T^3}{\pi^2 (\hbar s_{\alpha})^3}. \quad (3.73)$$

Чтобы получить полную энергию и теплоемкость, следует сложить вклады от трех акустических ветвей:

$$C_V \approx \frac{2k_B^4 T^3}{\pi^2 \hbar^3} \sum_{\alpha} \frac{1}{s_{\alpha}^3}, \quad (3.74)$$

где через  $s_{\alpha}$  обозначена скорость звука  $\alpha$ -й акустической ветви.

Мы проделали достаточно грубую оценку, поэтому к численным коэффициентам в последних двух выражениях не стоит относиться серьезно. Тем не менее, эта оценка дает правильную зависимость энергии и теплоемкости от температуры и скорости звука.

Рассчитаем теперь энергию решетки при низких температурах более аккуратно.

Формула (3.66) имеет вид суммы по различным колебаниям (различным состояниям фононов) определенной величины, которая зависит только от энергии фонона:

$$\sum_{\alpha, \mathbf{k}} f(\mathcal{E}_\alpha(\mathbf{k})), \quad f(\mathcal{E}) = \frac{\mathcal{E}}{\exp\left(\frac{\mathcal{E}}{k_B T}\right) - 1}. \quad (3.75)$$

Такие суммы встречаются довольно часто. Так как  $f$  зависит только от энергии фонона, то от суммы по состояниям можно перейти к интегралу по энергии:

$$\sum_{\alpha, \mathbf{k}} f(\mathcal{E}_\alpha(\mathbf{k})) = \int f(\mathcal{E}) \rho(\mathcal{E}) d\mathcal{E}. \quad (3.76)$$

Здесь  $\rho(\mathcal{E})$  — плотность состояний фононов. Напомним, что  $\rho(\mathcal{E})d\mathcal{E}$  — это число состояний квазичастиц (фононов) в единице объема с энергиями от  $\mathcal{E}$  до  $\mathcal{E} + d\mathcal{E}$ , то есть число различных колебаний с такими энергиями.

Суммарная плотность состояний складывается из плотности состояний разных ветвей:  $\rho(\mathcal{E}) = \sum_{\alpha} \rho_{\alpha}(\mathcal{E})$ . Плотность состояний ветви определяется ее законом дисперсии  $\omega_{\alpha}(\mathbf{k})$ . Аналитически получить законы дисперсии и плотности состояний фононов реальных кристаллов практически невозможно.

Однако при низких температурах энергия и теплоемкость определяются длинноволновыми акустическими фононами. Плотность состояний акустических фононов нам известна, мы получили ее в качестве примера, когда вводили само понятие плотности состояний (см. (2.12)). Если для  $\alpha$ -й акустической ветви  $\omega = s_{\alpha}|k|$ , то

$$\rho_{\alpha}(\mathcal{E}) = \frac{\mathcal{E}^2}{2\pi^2(s_{\alpha}\hbar)^3}. \quad (3.77)$$

Плотность состояний длинноволновых колебаний всех акустических ветвей определяется как сумма по трем акустическим ветвям:

$$\rho(\mathcal{E}) = \frac{\mathcal{E}^2}{2\pi^2\hbar^3} \sum_{\alpha} \frac{1}{s_{\alpha}^3} = \frac{3\mathcal{E}^2}{2\pi^2(S\hbar)^3}, \quad (3.78)$$

где  $S$  — «усредненная» скорость звука:

$$\frac{3}{S^3} = \sum_{\alpha=1}^3 \frac{1}{s_{\alpha}^3}. \quad (3.79)$$

Линейный закон дисперсии  $\omega = s|k|$  и соответствующая плотность состояний верны только при малых  $k$ . При больших значениях волнового вектора закон дисперсии и плотность состояний имеют более сложный вид.

Однако при низких температурах вклад в энергию и теплоемкость вносят только длинноволновые фононы, а при высоких температурах вид плотности состояний неважен, так как в этом случае на каждое колебание приходится энергия  $k_B T$ . Чтобы получить выражение, которое давало бы правильные предельные зависимости при низких и высоких температурах, Дебай предложил считать, что закон дисперсии  $\omega = s|k|$  выполняется и при больших  $k$ . Максимальное значение волнового вектора  $k_D$  при этом выбирается таким образом, чтобы в шаре радиуса  $k_D$  содержалось столько разрешенных значений волновых векторов, сколько их содержится в зоне Бриллюэна, то есть  $N = 1/v_0$ . Иными словами, объем этого шара должен быть равен объему зоны Бриллюэна  $(2\pi)^3/v_0$ , откуда

$$k_D^3 = 6\pi^2 N = \frac{6\pi^2}{v_0}. \quad (3.80)$$

Таким образом, сохраняя число акустических колебаний, мы заменяем первую зону Бриллюэна сферой, а реальный закон дисперсии — линейным. Фонон с волновым вектором  $k_D$  имеет энергию  $\mathcal{E}_D = \hbar s k_D = (6\pi^2 N)^{1/3} \hbar s$ . Соответствующая  $\mathcal{E}_D$  температура

$$\theta = \frac{\mathcal{E}_D}{k_B} = (6\pi^2 N)^{1/3} \frac{\hbar s}{k_B}, \quad (3.81)$$

называется температурой Дебая.

В таком приближении мы можем вычислить вклад акустических ветвей в энергию и теплоемкость решетки

$$\begin{aligned} E &= \int_0^{\mathcal{E}_D} \mathcal{E} \cdot \bar{n}(\mathcal{E}) \cdot \rho(\mathcal{E}) d\mathcal{E} = \int_0^{\mathcal{E}_D} \frac{\mathcal{E}}{\exp\left(\frac{\mathcal{E}}{k_B T}\right) - 1} \cdot \frac{3\mathcal{E}^2 d\mathcal{E}}{2\pi^2 s^3 \hbar^3} = \\ &= \frac{3k_B^4 T^4}{2\pi^2 s^3 \hbar^3} \int_0^{\theta/T} \frac{x^3 dx}{e^x - 1}. \end{aligned} \quad (3.82)$$

При низких температурах  $T \ll \theta$  верхний предел интеграла много больше единицы. Благодаря экспоненте в знаменателе интеграл сходится очень быстро, что позволяет положить верхний предел равным



бесконечности. Значение такого интеграла известно:

$$\int_0^{\infty} \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15}. \quad (3.83)$$

Для энергии акустических колебаний при низких температурах получаем

$$E = \frac{\pi^2}{10} \frac{k_B^4 T^4}{s^3 \hbar^3} = \frac{3\pi^4}{5} N k_B T \left( \frac{T}{\theta} \right)^3, \quad (3.84)$$

откуда следует, что теплоемкость решетки при низких температурах пропорциональна  $T^3$ :

$$C_V = \frac{dE}{dT} = \frac{2\pi^2}{5} \frac{k_B^4 T^3}{s^3 \hbar^3} = \frac{12\pi^4}{5} N k_B \left( \frac{T}{\theta} \right)^3. \quad (3.85)$$

При высоких температурах  $T \gg \theta$  верхний предел интегрирования мал, поэтому можно считать, что  $\exp(x) - 1 \approx x$ , таким образом,

$$\int_0^{\theta/T} \frac{x^3 dx}{e^x - 1} \approx \frac{1}{3} \left( \frac{\theta}{T} \right)^3, \quad (3.86)$$

$$E = 3N k_B T, \quad (3.87)$$

$$C_V = 3N k_B. \quad (3.88)$$

Это закон Дюлонга и Пти, в котором вместо полного числа колебаний  $3N$  подставлено число колебаний акустических ветвей  $3N$ . (При высоких температурах на каждое колебание приходится средняя энергия  $k_B T$ , полное число акустических колебаний равно  $3N$ , поэтому вклад акустических ветвей в энергию равен  $3N k_B T$ .)

В пределе низких и высоких температур модель Дебая дает точные значения для вклада акустических ветвей в энергию и теплоемкость. В области же промежуточных температур  $T \sim \theta$  эта модель лишь аппроксимирует реальную зависимость энергии и теплоемкости от температуры.

Температура Дебая разделяет две температурные области. В области низких температур на энергию и теплоемкость решетки сильное влияние оказывают квантовые эффекты («вымораживание» высокочастотных колебаний). В области высоких температур эти эффекты незначительны и теплоемкость может быть вычислена в классическом приближении. Для большинства кристаллов температура Дебая лежит в интервале от 100 до 300 К.

Чтобы получить полную энергию и теплоемкость кристаллической решетки, следует к вкладу акустических колебаний прибавить вклад оптических ветвей, для которого хорошим приближением является модель Эйнштейна. При низких температурах этот вклад пренебрежимо мал. При высоких температурах вклады всех ветвей в энергию и теплоемкость одинаковы.

### 3.7. Экспериментальные методы исследования закона дисперсии фононов

Наиболее удобным методом исследования закона дисперсии фононов является комбинационное рассеяние разного типа частиц (волн) на колебаниях кристаллической решетки. Это неупругое рассеяние: частица (волна), взаимодействуя с колебаниями решетки, изменяет не только направление движения, но и энергию (частоту).

Рассмотрим комбинационное рассеяние света. Пусть на кристалл падает пучок монохроматического света с волновым вектором  $\kappa_0$  и частотой  $\Omega_0 = c\kappa_0$  (рис. 3.7).

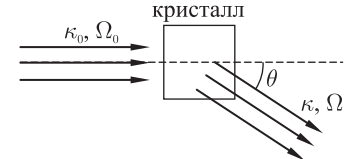


Рис. 3.7. Схема неупругого рассеяния фотона с волновым вектором  $\kappa_0$  на угол  $\theta$

Если исследовать спектр света, рассеянного кристаллом в определенном направлении, то в простейшем случае он будет иметь вид, показанный на рис. 3.8.

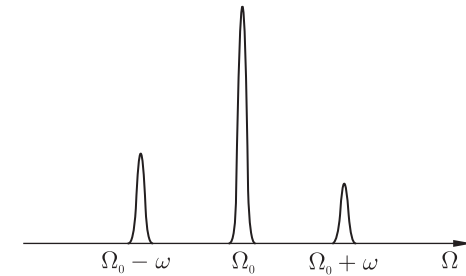


Рис. 3.8. Спектр комбинационного рассеяния света кристаллом. Слева от центральной линии ( $\Omega_0$ ) — стоксов пик, соответствующий испусканию фонона, справа — антистоксов, соответствующий поглощению фонона

Помимо высокого центрального пика, расположенного на частоте падающего света  $\Omega_0$  (упругое рассеяние), появляются еще два сателлита, сдвинутых влево и вправо на частоту длинноволнового оптического фонона  $\omega$ :

$$\Omega = \Omega_0 \pm \omega. \quad (3.89)$$

Левый пик  $\Omega = \Omega_0 - \omega$  называют стоксовым, правый  $\Omega = \Omega_0 + \omega$  — антистоксовым.

Изменение частоты при рассеянии невелико, так как частота оптического фонона меньше частоты света в десятки раз. Действительно, характерная энергия кванта света  $\hbar\Omega_0$  равна 1 эВ, а энергия оптического фонона по нашей оценке (см. (3.21)) составляет около 50 мэВ.

Если в кристалле имеется несколько ветвей оптических фононов, то в спектре комбинационного рассеяния будет наблюдаться несколько пар сателлитов.

Частоты, на которых располагаются линии спектра рассеянного света, представляют собой «комбинации» частоты падающего света  $\Omega_0$  и частоты фонона  $\omega$ . Из-за этого рассеяние и называется комбинационным.

Впервые эффект комбинационного рассеяния был обнаружен и объяснен Г. С. Ландсбергом и Л. И. Мандельштамом в 1928 г. В экспериментах использовались кристаллы кварца и исландского шпата. В том же году Ч. В. Раман обнаружил комбинационное рассеяние в жидкостях, поэтому эффект комбинационного рассеяния называют также эффектом Рамана, а рассеяние — рамановским. В 1930 г. Раману за обнаружение комбинационного рассеяния была присуждена Нобелевская премия.

Комбинационное рассеяние наблюдается не только в твердых телах. Например, при рассеянии на молекулах газа сдвиг линий комбинационного рассеяния будет определяться частотой колебаний атомов в молекуле. Если у молекулы имеется несколько колебательных степеней свободы, то в спектре рассеянного излучения будет наблюдаться несколько линий.

Возникает вопрос, почему при комбинационном рассеянии на кристалле получаются столь узкие пики (линии) излучения. Ведь оптические фононы обладают дисперсией и их частоты занимают достаточно широкий непрерывный интервал. Дело в том, что свет, рассеиваемый в данном направлении, взаимодействует только со строго определенным колебанием решетки. Проще всего показать это, описывая рассеяние на квантовом языке, хотя тот же результат можно получить и в рамках классической физики.

Явление комбинационного рассеяния заключается в том, что фотон падающего на кристалл света либо испускает фонон (стоксов процесс), либо поглощает его (антистоксов процесс). Эти процессы можно проиллюстрировать соответствующими фейнмановскими диаграммами <sup>1)</sup> (рис. 3.9).

Энергия и импульс должны сохраняться, то есть при рассеянии импульс и энергия фотона уменьшаются (стоксов процесс) или увели-

<sup>1)</sup> Фейнмановские диаграммы дают представление о процессах, происходящих с участием частиц (волн). Они позволяют по рисунку написать выражение для вероятности соответствующего процесса. Мы однако не будем здесь описывать эту процедуру. Отметим только, что линии в диаграммах обозначают частицы (волны), в точках соединения линий должны выполняться законы сохранения (3.90). Подробнее о фейнмановских диаграммах можно узнать из книг: *Лифшиц Е. М., Питаевский Л. П. Статистическая физика. Часть 2.* — М.: Наука, 1978. 448 с.; *Абрикосов А. А., Горьков Л. П., Дзялошинский И. Е. Методы квантовой теории поля в статистической физике.* — М.: Наука, 1962.

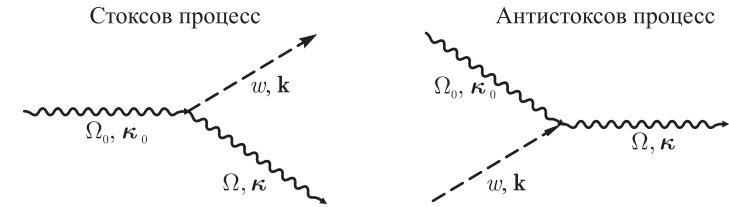


Рис. 3.9. Диаграммы рассеяния фотона, соответствующие испусканию фонона с волновым вектором  $\mathbf{k}$  (стоксов процесс) и поглощению фонона с волновым вектором  $\mathbf{k}$  (антистоксов процесс)

чиваются (антистоксов процесс) на импульс и энергию фонона соответственно.

Фотон падающего на кристалл света имеет энергию  $\hbar\Omega_0$  и импульс  $\hbar\kappa_0$ , рассеянный фотон — энергию  $\hbar\Omega$  и импульс  $\hbar\kappa$ . Энергия и импульс фонона равны  $\hbar\omega$  и  $\hbar\mathbf{k}$ , где  $\omega$  — частота, а  $\mathbf{k}$  — волновой вектор фонона.

Таким образом,

$$\hbar\kappa_0 \pm \hbar\mathbf{k} = \hbar\kappa, \quad (3.90)$$

$$\hbar\Omega_0 \pm \hbar\omega(\mathbf{k}) = \hbar\Omega. \quad (3.91)$$

Здесь знак «—» соответствует стоксовому процессу, «+» — антистоксовому.

Выражения (3.90) и (3.91) можно сократить на постоянную Планка, тогда

$$\kappa_0 \pm \mathbf{k} = \kappa, \quad (3.92)$$

$$\Omega_0 \pm \omega(\mathbf{k}) = \Omega. \quad (3.93)$$

Тот факт, что постоянная Планка не входит в эти уравнения, свидетельствует о том, что к такому же результату можно было прийти, описывая рассеяние языком классической физики.

Частота света и частота колебаний решетки представляют собой функции соответствующих волновых векторов:  $\Omega = c\kappa$  и  $\omega = \omega(\mathbf{k})$ . Поэтому, если задать направление и частоту падающего света и направление рассеяния, а также ветвь колебаний решетки, уравнения (3.92) и (3.93) будут однозначно определять волновой вектор колебания решетки  $\mathbf{k}$  и изменение частоты рассеянного света.

Как уже говорилось, характерные энергии оптических фононов ( $\sim 50$  мэВ) много меньше характерных энергий фотонов видимого света ( $\sim 1$  эВ). Другими словами, частота и соответственно длина волнового вектора фотона при рассеянии меняются мало, и  $\kappa \approx \kappa_0$ .

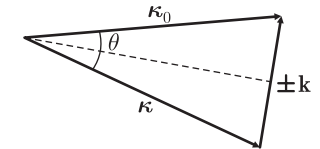


Рис. 3.10. Геометрический вывод формулы для модуля волнового вектора фонона, участвующего в рассеянии:  
 $k = 2\kappa_0 \sin \theta/2$

Поэтому, как видно из рис. 3.10, модуль волнового вектора фонона, участвующего в рассеянии, приблизительно равен  $2\kappa_0 \sin \theta/2$ , где  $\theta$  — угол рассеяния. Максимального значения  $2\kappa_0$  он достигает при  $\theta = \pi$ , то есть при рассеянии света назад.

Длина волны света видимого диапазона по порядку величины равна  $1 \text{ мкм} = 10^4 \text{ \AA}$ . Вследствие этого в комбинационном рассеянии света участвуют только длинноволновые фононы, волновой вектор которых ( $k \sim 10^{-4} \text{ \AA}^{-1}$ ) очень мал по сравнению с размерами зоны Бриллюэна ( $\pi/a \sim 1 \text{ \AA}^{-1}$ ).

Частота оптических фононов с такими волновыми векторами практически не отличается от  $\omega_0 = \omega_{\text{опт}}(\mathbf{k} = 0)$ , см. рис. 3.11, а, поэтому смещение стоксовой и антистоксовой линий при комбинационном рассеянии на оптических фононах не зависит от направления рассеяния и равно  $\omega_0$ . Таким образом, из спектра комбинационного рассеяния можно определить лишь одну точку дисперсионной зависимости оптических фононов.

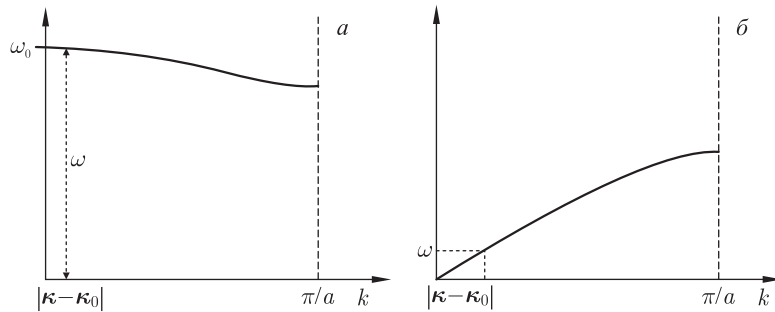


Рис. 3.11. Закон дисперсии оптических (а) и акустических (б) фононов. Отмечены частоты и волновые вектора фононов (пунктирные линии) участвующих в комбинационном рассеянии света

Мандельштам и Ландсберг в своих первых экспериментах по комбинационному рассеянию предполагали обнаружить рассеяние на акустических, а не на оптических колебаниях решетки. Изменение частоты при таком рассеянии намного меньше, чем при рассеянии на оптических фононах (рис. 3.11, б) и может быть обнаружено лишь при высоком спектральном разрешении измерительных приборов.

Рассеяние на акустических фононах действительно наблюдается и носит название рассеяния Мандельштама–Бриллюэна.

Закон дисперсии акустических колебаний при малых волновых векторах линеен:  $\omega_{\text{ак}} = sk$ , где  $s$  — скорость звука. Поэтому изменение частоты света при рассеянии на угол  $\theta$  равно

$$\omega = sk = s \cdot 2\kappa_0 \sin \theta/2 = 2\Omega_0 \frac{s}{c} \sin \theta/2. \quad (3.94)$$

Отсюда видно, что относительное изменение частоты  $\omega/\Omega_0 \sim s/c$  очень мало. Скорость света в кристалле равна  $10^8 \text{ м/с}$ , скорость звука —  $10^3 \text{ м/с}$ , поэтому  $s/c \sim 10^{-5}$ . Отметим также, что в отличие от рассеяния на оптических фононах при рассеянии на акустических фононах изменение частоты света зависит от угла рассеяния, см. рис. 3.11, б. Это позволяет измерять дисперсию акустических фононов.

В спектре комбинационного рассеяния (рис. 3.8) высота стоксового пика больше, чем антистоксового. Этому явлению легко дать качественное объяснение: чтобы поглотить фонон, нужно, чтобы он в кристалле присутствовал, а испустить фонон можно без помощи других фононов. Поэтому при низких температурах, когда фононов мало, интенсивность антистоксовой линии намного меньше интенсивности стоксовой.

Как показал Эйнштейн, имеющиеся в кристалле фононы «помогают» излучению фононов (вынужденное излучение). Вынужденное излучение будет более подробно рассмотрено ниже в применении к свету, а пока лишь скажем, что вероятность рассеяния с поглощением фонона пропорциональна числу уже имеющихся фононов данного типа  $N$ , а вероятность обратного процесса — рассеяния с испусканием фонона — пропорциональна  $N + 1$ . Единица в последнем выражении как раз и соответствует испусканию фонона «на пустом месте» (спонтанному излучению).

Если пренебречь малым (по сравнению с частотой света) изменением частоты при рассеянии, то коэффициенты пропорциональности в этих вероятностях можно считать одинаковыми. В состоянии термодинамического равновесия число фононов данного типа с частотой  $\omega$  описывается распределением Бозе–Эйнштейна

$$N = \frac{1}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1}. \quad (3.95)$$

Отсюда для отношения интенсивностей стоксовой и антистоксовой линий получаем

$$\frac{W_{\text{антистокс}}}{W_{\text{стокс}}} \approx \frac{N}{N + 1} = \exp\left(-\frac{\hbar\omega}{k_B T}\right). \quad (3.96)$$

Это отношение позволяет по данным рассеяния бесконтактным методом измерить температуру кристалла. Интенсивности стоксовой и антистоксовой линий сильно различаются при низких температурах, когда  $k_B T < \hbar\omega$ . При  $k_B T \gg \hbar\omega$ , в классическом пределе, интенсивности обеих линий равны.

Итак, мы рассмотрели комбинационное рассеяние света на оптических и акустических колебаниях решетки. Один из выводов этого рассмотрения состоит в том, что в рассеянии участвуют лишь длинноволновые фононы. Поэтому с помощью комбинационного рассеяния света видимого диапазона можно получить информацию только о длинновол-

новой части фононного спектра — определить частоты  $\omega_0 = \omega(\mathbf{k} = 0)$  для оптических фононов и скорость звука  $s$  для акустических.

Как с помощью комбинационного рассеяния исследовать весь фононный спектр, в частности область больших волновых векторов? Для этого нужно изучать рассеяние частиц (или квазичастиц), длина волны которых сравнима с постоянной решетки. В рассеянии таких частиц могут участвовать коротковолновые фононы.

Электромагнитное излучение с такими длинами волн принадлежит рентгеновскому диапазону. Мы уже рассматривали упругое рассеяние рентгеновских волн на неподвижной идеальной кристаллической решетке. Энергия кванта такого излучения по порядку величины равна  $10^4$  эВ. Напомним, что энергия оптического фонона составляет величину порядка  $0,03$ – $0,1$  эВ. Заметить такое изменение на фоне  $10^4$  эВ очень трудно.

По этой причине для изучения коротковолнового фононного спектра в качестве рассеивающихся частиц используют нейтроны. Нейтрон электрически нейтрален, он слабо взаимодействует с электронной системой кристалла. Энергия нейтрона, обладающего длиной волны порядка постоянной решетки, сравнима с энергией фонона:

$$\mathcal{E}_n = \frac{\hbar^2 k^2}{2M_n} = \frac{\hbar^2 \cdot 4\pi^2}{2M_n \lambda^2} \sim \frac{10^{-54} \cdot 40}{2 \cdot 10^{-24} \cdot 10^{-16}} \sim 0,1 \text{ эВ}. \quad (3.97)$$

Такое значение для  $\mathcal{E}_n$  связано с тем, что масса нейтрона  $M_n$  близка к массе атома  $M$ .

Энергии нейтронов, испускаемых при ядерных реакциях, на много порядков больше, поэтому, чтобы использовать нейтроны для исследования фононного спектра, приходится их замедлять.

Зная энергию падающего нейтрона и нейтрона, рассеянного в определенном направлении, мы можем определить энергию и импульс фонона, воспользовавшись законами сохранения энергии и импульса

$$\kappa = \kappa_0 + \mathbf{k}, \quad (3.98)$$

$$\frac{\hbar^2 \kappa_0^2}{2M} = \frac{\hbar^2 \kappa^2}{2M} + \hbar\omega(\mathbf{k}). \quad (3.99)$$

Изменяя энергию нейтронов и направление рассеяния, можно исследовать закон дисперсии фононов.

Отметим, что в спектре рассеянных нейтронов (как и в спектре рассеянных фотонов) будут также присутствовать пики, вызванные рассеянием с участием нескольких фононов.

Необходимо сделать одно замечание о законе сохранения импульса (3.98).

Из-за дискретности кристаллической решетки волновые векторы  $\mathbf{k}$  и  $\mathbf{k} + \mathbf{G}$  соответствуют одному и тому же фонону. (Здесь  $\mathbf{G}$  — произвольный вектор обратной решетки). Мы условились описывать колебания волновыми векторами, лежащими в первой зоне Бриллюэна,

так как любой вектор можно перенести в первую зону, прибавляя к нему векторы обратной решетки.

В комбинационном рассеянии света участвуют фононы с малыми волновыми векторами, которые автоматически оказываются в первой зоне Бриллюэна.

Волновой вектор нейтрона сравним с размерами первой зоны Бриллюэна. В связи с этим может возникнуть вопрос, не распространяется ли ограничение волновых векторов фононов первой зоной Бриллюэна на закон сохранения импульса (3.98).

Анализ показывает, что это не так: формально в рассеянии может участвовать фонон с произвольным волновым вектором. Мы можем рассматривать частоту фононов как периодическую функцию  $\mathbf{k}$ , не ограничиваясь первой зоной Бриллюэна. При этом надо помнить, что волновые векторы, отличающиеся на вектор постоянной решетки, описывают одно и то же колебание.

Есть и другая возможность. Можно считать, что  $\mathbf{k}$  обязательно лежит в первой зоне Бриллюэна, но вместо (3.98) записывать закон сохранения импульса в следующей форме:

$$\kappa = \kappa_0 + \mathbf{k} + \mathbf{G}, \quad (3.100)$$

где  $\mathbf{G}$  — произвольный вектор обратной решетки. Можно считать, что мы учитываем закон сохранения импульса для трех частиц: нейтрона, фонона и кристалла как целого. При этом кристалл как целое может отдать или забрать не любой импульс, а лишь импульс, который соответствует произвольному вектору обратной решетки  $\mathbf{G}$ .

## Приложение.

### Операторы рождения и уничтожения фононов

В этом параграфе мы покажем, каким образом введенные в § 3.4 понятия фононов появляются при квантовании колебаний кристаллической решетки. Получающиеся при этом выражения являются очень важными, на них основан математический аппарат, применяемый при изучении элементарных процессов взаимодействия фононов с электронами (см. главу 4).

Аналогично случаю электромагнитного поля произвольное колебательное движение кристаллической решетки может быть представлено в виде суперпозиции бегущих плоских волн<sup>1)</sup>. Если рассматривать объем кристалла как большой, но конечный, то волновой вектор  $\mathbf{k}$  будет пробегать ряд дискретных значений. Смещения атомов кристалла

<sup>1)</sup> Ландау Л. Д., Лифшиц Е. М. Теория поля. — М.: Наука, 1973. 504 с.

$\mathbf{U}_\nu(\mathbf{n}, t)$  записываются в виде сумм

$$\mathbf{U}_\nu(\mathbf{n}, t) = \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{3l} \sum_{\mathbf{k}} \left\{ a_{\mathbf{k}\alpha}(t) \mathbf{e}_\nu^{(\alpha)}(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}_n} + a_{\mathbf{k}\alpha}^*(t) \mathbf{e}_\nu^{(\alpha)*}(\mathbf{k}) e^{-i\mathbf{k}\mathbf{r}_n} \right\}, \quad (3.101)$$

где  $N$  — число элементарных ячеек в кристалле;  $l$  — число атомов в элементарной ячейке. Суммирование производится по всем (неэквивалентным) значениям  $\mathbf{k}$  и по всем ветвям спектра колебаний.

Векторы  $\mathbf{e}_\nu^{(\alpha)}$  в (3.101) — это векторы поляризации колебаний. Они удовлетворяют не только уравнениям (3.47), но и определенным условиям нормировки. Эти условия можно записать в виде <sup>1)</sup>

$$\sum_{\nu=1}^l \frac{M_\nu}{M} \mathbf{e}_\nu^{(\alpha)}(\mathbf{k}) \left[ \mathbf{e}_\nu^{(\alpha')}(\mathbf{k}) \right]^* = \delta_{\alpha\alpha'}, \quad (3.102)$$

где  $M = \sum M_\nu$  — суммарная масса атомов в одной элементарной ячейке. Подставим выражение для смещения  $\mathbf{U}_\nu(\mathbf{n}, t)$  в уравнение (3.44), и тогда для коэффициентов  $a_{\mathbf{k}\alpha}(t)$  получаем уравнение

$$\ddot{a}_{\mathbf{k}\alpha} + \omega_\alpha^2(\mathbf{k}) a_{\mathbf{k}\alpha} = 0. \quad (3.103)$$

Колебательная энергия решетки (сумма кинетической  $T$  и потенциальной  $V$  энергий) выражается через смещения и скорости атомов формулой

$$E = \frac{1}{2} \sum_{\mathbf{n}\nu} M_\nu \dot{\mathbf{U}}_\nu^2(\mathbf{n}) + \frac{1}{2} \sum_{\substack{\mathbf{n}\mathbf{n}' \\ \nu\nu'}} \Lambda_{ik}^{\nu\nu'}(\mathbf{n} - \mathbf{n}') U_{\nu i}(\mathbf{n}) U_{\nu' k}(\mathbf{n}') \equiv T + V. \quad (3.104)$$

Подставим в (3.104) выражение для смещения  $\mathbf{U}_\nu(\mathbf{n}, t)$  из (3.101). Все члены получающихся сумм, которые содержат множители  $\exp[\pm i(\mathbf{k} \pm \mathbf{k}')\mathbf{r}_n]$  с  $\mathbf{k} \pm \mathbf{k}' \neq 0$ , обращаются в нуль при суммировании по  $\mathbf{n}$  в силу того, что

$$\sum_{\mathbf{n}} e^{i\mathbf{q}\mathbf{r}_n} = \begin{cases} N & \text{при } \mathbf{q} = 0, \\ 0 & \text{при } \mathbf{q} \neq 0, \end{cases} \quad (3.105)$$

где  $\mathbf{q}$  пробегает все неэквивалентные значения. Учитывая условие (3.102), а также соотношение <sup>2)</sup>

$$e_\nu^{(\alpha)}(-\mathbf{k}) = \left[ e_\nu^{(\alpha)}(\mathbf{k}) \right]^*, \quad (3.106)$$

<sup>1)</sup> Ландау Л.Д., Лифшиц Е.М. Статистическая физика. Часть 1. — М.: Наука, 1976. 584 с.

<sup>2)</sup> Там же.

можно привести кинетическую энергию в (3.104) к виду

$$T = \sum_{\alpha\mathbf{k}} M\omega_\alpha^2 \left\{ a_{\mathbf{k}\alpha} a_{\mathbf{k}\alpha}^* + \frac{1}{2} (a_{\mathbf{k}\alpha} a_{-\mathbf{k}\alpha} + a_{\mathbf{k}\alpha}^* a_{-\mathbf{k}\alpha}^*) \right\}. \quad (3.107)$$

Потенциальную энергию в (3.104) с помощью уравнений движения (3.44) можно привести к виду

$$V = -\frac{1}{2} \sum_{\mathbf{n}\nu} M_\nu \ddot{\mathbf{U}}_\nu(\mathbf{n}) \mathbf{U}_\nu(\mathbf{n}). \quad (3.108)$$

Затем подставим сюда выражение для  $\mathbf{U}_\nu(\mathbf{n})$  из (3.101) и преобразуем это выражение аналогично тому, как это было сделано для кинетической энергии  $T$ . В результате, для потенциальной энергии получаем

$$V = \sum_{\alpha\mathbf{k}} M\omega_\alpha^2 \left\{ a_{\mathbf{k}\alpha} a_{\mathbf{k}\alpha}^* - \frac{1}{2} (a_{\mathbf{k}\alpha} a_{-\mathbf{k}\alpha} + a_{\mathbf{k}\alpha}^* a_{-\mathbf{k}\alpha}^*) \right\}. \quad (3.109)$$

Складывая обе части  $T$  и  $V$ , найдем полную энергию

$$E = \sum_{\alpha\mathbf{k}} 2M\omega_\alpha^2(\mathbf{k}) |a_{\mathbf{k}\alpha}|^2. \quad (3.110)$$

Таким образом, полная энергия колебаний решетки выражается в виде суммы энергий, связанных с каждой из волн в отдельности.

Произведем далее замену переменных, в результате которой уравнения движения решетки получат вид канонических уравнений механики. Для этого вводим новые переменные  $Q_{\mathbf{k}\alpha}$  и  $P_{\mathbf{k}\alpha}$  согласно

$$\begin{aligned} Q_{\mathbf{k}\alpha} &= \sqrt{M} (a_{\mathbf{k}\alpha} + a_{\mathbf{k}\alpha}^*), \\ P_{\mathbf{k}\alpha} &= -i\omega_\alpha(\mathbf{k}) \sqrt{M} (a_{\mathbf{k}\alpha} - a_{\mathbf{k}\alpha}^*) = \dot{Q}_{\mathbf{k}\alpha}. \end{aligned} \quad (3.111)$$

Обратные к (3.111) соотношения таковы:

$$\begin{aligned} a_{\mathbf{k}\alpha} &= \frac{1}{2\sqrt{M}} \left( Q_{\mathbf{k}\alpha} + \frac{iP_{\mathbf{k}\alpha}}{\omega_\alpha(\mathbf{k})} \right), \\ a_{\mathbf{k}\alpha}^* &= \frac{1}{2\sqrt{M}} \left( Q_{\mathbf{k}\alpha} - \frac{iP_{\mathbf{k}\alpha}}{\omega_\alpha(\mathbf{k})} \right). \end{aligned} \quad (3.112)$$

Подставив (3.112) в (3.110), получим функцию Гамильтона для решетки

$$H = \frac{1}{2} \sum_{\alpha\mathbf{k}} [P_{\mathbf{k}\alpha}^2 + \omega_\alpha^2(\mathbf{k}) Q_{\mathbf{k}\alpha}^2]. \quad (3.113)$$

Легко убедиться в том, что уравнения Гамильтона  $\partial H / \partial P_{\mathbf{k}\alpha} = \dot{Q}_{\mathbf{k}\alpha}$  совпадают с равенством  $P_{\mathbf{k}\alpha} = \dot{Q}_{\mathbf{k}\alpha}$  (см. (3.111)), а из уравнения  $\partial H / \partial Q_{\mathbf{k}\alpha} = -\dot{P}_{\mathbf{k}\alpha}$  находим уравнение движения решетки

$$\ddot{Q}_{\mathbf{k}\alpha} + \omega_\alpha^2(\mathbf{k}) Q_{\mathbf{k}\alpha} = 0. \quad (3.114)$$

Таким образом, мы представили функцию Гамильтона (3.113) в виде суммы независимых членов. Каждое слагаемое имеет вид гамильтоновой функции одномерного гармонического осциллятора и описывает определенное нормальное колебание решетки. Нормальное колебание в данном случае представляет собой бегущую волну, которая характеризуется определенными номерами колебательной ветви  $\alpha$  и квазиволновым вектором  $\mathbf{k}$ . Такой способ описания классического колебательного движения позволяет указать путь перехода к квантовой теории аналогично тому, как производится переход от классического описания свободного электромагнитного поля к квантовой картине фотонов <sup>1)</sup>.

Произведем теперь квантование колебаний решетки. Для этого мы должны рассматривать канонические переменные — обобщенные координаты  $Q_{\mathbf{k}\alpha}$  и обобщенные импульсы  $P_{\mathbf{k}\alpha}$  — как операторы, удовлетворяющие перестановочным соотношениям

$$\hat{P}_{\mathbf{k}\alpha} \hat{Q}_{\mathbf{k}\alpha} - \hat{Q}_{\mathbf{k}\alpha} \hat{P}_{\mathbf{k}\alpha} = -i\hbar. \quad (3.115)$$

В результате функция Гамильтона становится оператором. Таким образом, задача о собственных значениях оператора энергии системы сводится к задаче об уровнях энергии системы линейных осцилляторов. Они хорошо известны из квантовой механики <sup>2)</sup>:

$$E = \sum_{\alpha \mathbf{k}} \hbar \omega_{\alpha}(\mathbf{k}) \left( n_{\mathbf{k}\alpha} + \frac{1}{2} \right), \quad n_{\mathbf{k}\alpha} = 0, 1, 2, \dots \quad (3.116)$$

Полученное для энергии выражение и дает возможность ввести понятие о фононах. Таким образом, возбужденное состояние решетки можно рассматривать как совокупность элементарных возбуждений (квазичастиц), каждое из которых имеет энергию  $\hbar \omega_{\alpha}(\mathbf{k})$ , являющуюся функцией квазиимпульса  $\mathbf{k}$ . В этом случае квантовые числа  $n_{\mathbf{k}\alpha}$  становятся числами заполнения различных состояний квазичастиц, или просто фононов.

**Замечание.** Слагаемое в (3.116) при  $n_{\mathbf{k}\alpha} = 0$  принято называть энергией нулевых колебаний. Это слагаемое следует включить в энергию основного состояния кристалла. Существование энергии нулевых колебаний решетки не приводит к каким-либо принципиальным затруднениям (в отличие от квантовой электродинамики, где сумма  $\sum \hbar \omega$  расходится).

Из квантовой механики известно, что для гармонического осциллятора величины  $\omega_{\alpha}(\mathbf{k}) \hat{Q}_{\mathbf{k}\alpha} \pm i \hat{P}_{\mathbf{k}\alpha}$  имеют отличные от нуля матричные

<sup>1)</sup> См. Берестецкий В. Б., Лифшиц Е. М., Питаевский Л. П. Квантовая электродинамика. — М.: Наука, 1980.

<sup>2)</sup> См. Ландау Л. Д., Лифшиц Е. М. Квантовая механика. — М.: Наука, 1974.

элементы только для переходов с изменением чисел  $n_{\mathbf{k}\alpha}$  на единицу <sup>1)</sup>. Введем операторы

$$\begin{aligned} \hat{C}_{\mathbf{k}\alpha} &= \frac{1}{\sqrt{2\hbar\omega_{\alpha}(\mathbf{k})}} \left[ \omega_{\alpha}(\mathbf{k}) \hat{Q}_{\mathbf{k}\alpha} + i \hat{P}_{\mathbf{k}\alpha} \right], \\ \hat{C}_{\mathbf{k}\alpha}^+ &= \frac{1}{\sqrt{2\hbar\omega_{\alpha}(\mathbf{k})}} \left[ \omega_{\alpha}(\mathbf{k}) \hat{Q}_{\mathbf{k}\alpha} - i \hat{P}_{\mathbf{k}\alpha} \right], \end{aligned} \quad (3.117)$$

отличные от нуля матричные элементы которых будут равны

$$\langle n_{\mathbf{k}\alpha} - 1 | \hat{C}_{\mathbf{k}\alpha} | n_{\mathbf{k}\alpha} \rangle = \langle n_{\mathbf{k}\alpha} | \hat{C}_{\mathbf{k}\alpha}^+ | n_{\mathbf{k}\alpha} - 1 \rangle = \sqrt{n_{\mathbf{k}\alpha}}. \quad (3.118)$$

С помощью определения (3.117) и правил (3.115) находим следующие соотношения коммутации между операторами  $\hat{C}_{\mathbf{k}\alpha}$  и  $\hat{C}_{\mathbf{k}\alpha}^+$ :

$$\hat{C}_{\mathbf{k}\alpha} \hat{C}_{\mathbf{k}\alpha}^+ - \hat{C}_{\mathbf{k}\alpha}^+ \hat{C}_{\mathbf{k}\alpha} = 1, \quad (3.119)$$

а операторы соответствующие разным  $\mathbf{k}$  и  $\alpha$  — коммутируют. При действии оператора  $\hat{C}_{\mathbf{k}\alpha}$  на волновую функцию состояния с  $n_{\mathbf{k}\alpha}$  фононами получается волновая функция состояния с  $n_{\mathbf{k}\alpha} - 1$  фононом. И наоборот, при действии оператора  $\hat{C}_{\mathbf{k}\alpha}^+$  на функцию состояния с  $n_{\mathbf{k}\alpha}$  фононами, получаем функцию состояния с  $n_{\mathbf{k}\alpha} + 1$  фононом. Таким образом, в смысле воздействия на функции чисел заполнения операторы  $\hat{C}_{\mathbf{k}\alpha}$  и  $\hat{C}_{\mathbf{k}\alpha}^+$  играют роль операторов уничтожения и рождения фононов. При этом коммутационное правило (3.119), как и следовало ожидать, отвечает статистике Бозе.

Гамильтониан решетки, выраженный через операторы  $\hat{C}_{\mathbf{k}\alpha}$  и  $\hat{C}_{\mathbf{k}\alpha}^+$ , имеет вид

$$\hat{H} = \frac{1}{2} \sum_{\alpha \mathbf{k}} \hbar \omega_{\alpha}(\mathbf{k}) (\hat{C}_{\mathbf{k}\alpha}^+ \hat{C}_{\mathbf{k}\alpha} + \hat{C}_{\mathbf{k}\alpha} \hat{C}_{\mathbf{k}\alpha}^+) = \sum_{\alpha \mathbf{k}} \hbar \omega_{\alpha}(\mathbf{k}) \left( \hat{C}_{\mathbf{k}\alpha}^+ \hat{C}_{\mathbf{k}\alpha} + \frac{1}{2} \right). \quad (3.120)$$

Таким образом, произведение  $\hat{C}_{\mathbf{k}\alpha}^+ \hat{C}_{\mathbf{k}\alpha}$  в представлении фононных чисел заполнения есть диагональный оператор с матричными элементами, равными  $n_{\mathbf{k}\alpha}$ . Он, таким образом, является оператором числа фононов в состоянии  $\mathbf{k}, \alpha$  <sup>2)</sup>.

В результате мы можем заключить, что вместе с величинами  $\hat{C}_{\mathbf{k}\alpha}$  становятся операторами также и векторы смещения  $\mathbf{U}_{\alpha}(\mathbf{n})$ :

$$\hat{\mathbf{U}}_{\nu}(\mathbf{n}) = \sqrt{\frac{\hbar}{2MN}} \sum_{\alpha \mathbf{k}} \frac{1}{\sqrt{\omega_{\alpha}(\mathbf{k})}} \left[ \hat{C}_{\mathbf{k}\alpha} \mathbf{e}_{\nu}^{(\alpha)}(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}_n} + \hat{C}_{\mathbf{k}\alpha}^+ \mathbf{e}_{\nu}^{(\alpha)*}(\mathbf{k}) e^{-i\mathbf{k}\mathbf{r}_n} \right]. \quad (3.121)$$

<sup>1)</sup> См. Ландау Л. Д., Лифшиц Е. М. Квантовая механика. — М.: Наука, 1974.

<sup>2)</sup> Часто, в литературе, операторы рождения и уничтожения фононов обозначаются через  $\hat{a}_{\mathbf{k}\alpha}^+$  и  $\hat{a}_{\mathbf{k}\alpha}$  соответственно.

Для дальнейшего будет удобно преобразовать это выражение. Для этого заменим во втором слагаемом (3.121) индекс суммирования  $\mathbf{k}$  на  $-\mathbf{k}$  и учтем соотношение (3.106) для векторов  $\mathbf{e}_\nu^{(\alpha)}(\mathbf{k})$ . Вынося за скобки общий коэффициент при операторах  $\hat{C}_{\mathbf{k}\alpha}$  и  $\hat{C}_{-\mathbf{k}\alpha}^+$ , получаем

$$\hat{U}_\nu(\mathbf{n}) = \sqrt{\frac{\hbar}{2MN}} \sum_{\alpha\mathbf{k}} \frac{1}{\sqrt{\omega_\alpha(\mathbf{k})}} \mathbf{e}_\nu^{(\alpha)}(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}_n} \left( \hat{C}_{\mathbf{k}\alpha} + \hat{C}_{-\mathbf{k}\alpha}^+ \right). \quad (3.122)$$

В заключение остановимся на особенностях квантовомеханического описания длинноволновых акустических колебаний. Для перехода к этому описанию нет необходимости пользоваться уравнениями движения дискретной решетки, можно действовать непосредственно в рамках теории упругости. В данном случае следует исходить из выражения для полной энергии упругой сплошной среды, которая равна сумме потенциальной и кинетической энергий,

$$E = \frac{1}{2} \int d^3r \left( \rho \dot{\mathbf{U}}^2 + C_{ilmn} \frac{\partial U_i}{\partial r_l} \frac{\partial U_m}{\partial r_n} \right). \quad (3.123)$$

Здесь  $\rho = MN/V$  — плотность кристалла;  $V$  — нормировочный объем кристалла;  $\mathbf{U}(\mathbf{r})$  — вектор смещения;  $C_{ilmn}$  — тензор упругих модулей кристалла (см. конец гл. 1).

Поскольку вся дальнейшая процедура квантования длинноволновых акустических колебаний аналогична проделанной выше, мы, не останавливаясь на деталях, приведем конечные результаты.

Оператор упругого смещения  $\hat{\mathbf{U}}(\mathbf{r}, t)$  сплошной среды, аналогично (3.122), имеет вид

$$\hat{\mathbf{U}}(\mathbf{r}, t) = \sqrt{\frac{\hbar}{2\rho V}} \sum_{\alpha\mathbf{k}} \frac{1}{\sqrt{\omega_\alpha(\mathbf{k})}} \mathbf{e}^{(\alpha)}(\mathbf{k}) e^{i\mathbf{k}\mathbf{r}} \left( \hat{C}_{\mathbf{k}\alpha} + \hat{C}_{-\mathbf{k}\alpha}^+ \right). \quad (3.124)$$

Здесь  $\alpha$  пробегает три значения, соответствующие трем акустическим колебаниям.

Легко написать выражение для оператора пространственной производной от вектора смещения

$$\frac{\partial \hat{U}_l}{\partial r_m} = i \sqrt{\frac{\hbar}{2\rho V}} \sum_{\alpha\mathbf{k}} \frac{1}{\sqrt{\omega_\alpha(\mathbf{k})}} \mathbf{e}_l^{(\alpha)}(\mathbf{k}) k_m e^{i\mathbf{k}\mathbf{r}} \left( \hat{C}_{\mathbf{k}\alpha} + \hat{C}_{-\mathbf{k}\alpha}^+ \right). \quad (3.125)$$

Наконец, оператор Гамильтона системы длинноволновых акустических фононов имеет вид, аналогичный (3.120),

$$\hat{H} = \sum_{\alpha\mathbf{k}} \hbar \omega_\alpha(\mathbf{k}) \left( \hat{n}_{\mathbf{k}\alpha} + \frac{1}{2} \right). \quad (3.126)$$

Здесь  $\hat{n}_{\mathbf{k}\alpha} = \hat{C}_{\mathbf{k}\alpha}^+ \hat{C}_{\mathbf{k}\alpha}$  — оператор чисел заполнения фононов.

## Задачи

**Задача 3.1.** Чем отличается форма оптических и акустических колебаний в длинноволновом пределе ( $ka \ll 1$ ).

*Решение.* В длинноволновом пределе ( $ka \ll 1$ ), согласно (3.25) и (3.36), имеем

$$\frac{B}{A} = 1 - \frac{M_1 \omega^2}{2\gamma},$$

где  $A$  и  $B$  — амплитуды колебаний меньшего и большего атомов.

а) Для акустических колебаний, в длинноволновом пределе,  $\omega = s|k|$  и стремится к нулю. Тогда  $B/A = 1$ . Это значит, что амплитуды колебаний легкого и тяжелого атомов одинаковы. Следовательно, оба атома колеблются в фазе.

б) Для оптической ветви при  $k = 0$  (см. (3.37))

$$\omega^2(0) = \frac{2\gamma}{\mu}.$$

Тогда

$$\frac{B}{A} = \frac{2\gamma - M_1(2\gamma/\mu)}{2\gamma} = -\frac{M_1}{M_2}.$$

Мы получаем, что атомы в каждой ячейке движутся в противофазе. Причем, амплитуда движения легкого атома больше амплитуды тяжелого атома в  $M_2/M_1$  раз.

**Задача 3.2.** Каково отличие между акустической и оптической ветвей на границе зоны Бриллюэна?

*Решение.* Согласно (3.25) и (3.36) найдем отношение амплитуд колебаний легкого и тяжелого атомов:

$$\frac{A}{B} = \frac{\gamma(e^{ika} + 1)}{2\gamma - M_1 \omega^2}.$$

а) Для акустической ветви, на границе зоны Бриллюэна

$$\frac{A}{B} = \frac{\gamma(e^{i\pi} + 1)}{2\gamma - M_1 \omega^2} = 0.$$

Это значит, что легкие атомы совсем не колеблются, а колеблются только тяжелые с массой  $M_2$ .

б) Для оптической ветви на границе зоны Бриллюэна, согласно (3.38),

$$\omega^2\left(\frac{\pi}{a}\right) = \frac{2\gamma}{M_2},$$

следовательно,  $B = 0$ . В результате, в случае оптических колебаний при  $k = \frac{\pi}{a}$  колеблются легкие атомы, а тяжелые стоят на месте.

**Задача 3.3.** Найти скорость звука для акустических колебаний в длинноволновом пределе.

*Решение.* Согласно (3.30) и (3.33), получаем, что при  $ka \ll 1$

$$\omega^2(k) \approx \frac{2\mu\gamma}{M_1 M_2} \left( \frac{ka}{2} \right)^2 = a^2 \frac{\gamma}{2(M_1 + M_2)} k^2.$$

Следовательно, скорость звука  $s$  равна

$$s = a \sqrt{\frac{\gamma}{2(M_1 + M_2)}}.$$

**Задача 3.4.** Найти плотность состояний оптических и акустических колебаний в длинноволновой области спектра ( $ka \ll 1$ ).

*Решение.* Для случая, когда есть два атома в элементарной ячейке, мы имеем:

- 1) одну продольную акустическую ветвь LA,
- 2) две поперечные акустические ветви TA,
- 3) одну продольную оптическую ветвь LO,
- 4) две поперечные оптические ветви TO.

При  $ka \ll 1$ , согласно (3.33), имеем

$$\omega_{\text{ак}} = a \sqrt{\frac{\gamma}{2(M_1 + M_2)}} k = sk.$$

Аналогично для оптической ветви при  $ka \ll 1$  имеем

$$\omega_{\text{опт}} = \sqrt{\frac{2\gamma}{\mu}} \left( 1 - \frac{\mu^2}{8M_1 M_2} k^2 a^2 \right) = \omega_0 \left( 1 - \frac{M_1 M_2}{8(M_1 + M_2)^2} k^2 a^2 \right),$$

где  $\omega_0 = \sqrt{\frac{2\gamma}{\mu}}$ .

Согласно § 2.3, трехмерная плотность состояний равна

$$\rho^{3D}(\varepsilon) = \frac{k^2}{2\pi^2 \left| \frac{d\varepsilon}{dk} \right|}.$$

Аналогично одномерная плотность состояний равна

$$\rho^{1D}(\varepsilon) = \frac{1}{\pi \left| \frac{d\varepsilon}{dk} \right|}.$$

В результате, для акустических фононов в длинноволновом пределе получаем

$$\rho_{\text{ак}}^{3D} = \frac{\varepsilon^2}{2\pi^2 (s\hbar)^3}; \quad \rho_{\text{ак}}^{1D} = \frac{1}{\pi\hbar s} = \text{const}.$$

Здесь  $\varepsilon = \hbar\omega$ .

Аналогично для оптических фононов в длинноволновом пределе находим

$$\rho_{\text{опт}}^{3D} = \frac{4\sqrt{2} (M_1 + M_2)^3}{\pi^2 a^3 (M_1 M_2)^{3/2} \varepsilon_0} \left( 1 - \frac{\varepsilon}{\varepsilon_0} \right)^{1/2};$$

$$\rho_{\text{опт}}^{1D} = \frac{2\sqrt{2} (M_1 + M_2)}{\sqrt{2} \pi (M_1 M_2)^{1/2} a \varepsilon_0} \frac{1}{\left( 1 - \frac{\varepsilon}{\varepsilon_0} \right)^{1/2}},$$

где  $\varepsilon_0 = \hbar\omega_0$ .

**Задача 3.5.** Оцените, какую энергию имеет нейтрон с длиной волны де Бройля порядка постоянной решетки?

*Ответ.*

$$\varepsilon_H = \frac{\hbar^2 k^2}{2M} = \frac{\hbar^2}{2M} \frac{4\pi^2}{\lambda^2} \approx 10^{-13} \text{ эрг} \approx 0,1 \text{ эВ}.$$



## Глава 4

## ЭЛЕКТРОНЫ В ИДЕАЛЬНОМ КРИСТАЛЛЕ

## 4.1. Введение

В свободном атоме электронные состояния определяются взаимодействием электронов с ядром атома и друг с другом. Как известно, связанные состояния локализованы, а их энергетический спектр состоит из дискретных уровней.

В твердом (и жидком) теле, состоящем из большого числа атомов, электрон взаимодействует с электронами и ядрами всех атомов, которые входят в кристалл. Благодаря этому взаимодействию и осуществляется химическая связь: полная энергия конденсированного состояния (твердого или жидкого) оказывается меньше суммарной энергии отдельных атомов. В образовании химической связи участвуют в основном те электроны, которые в изолированном атоме составляли его внешнюю оболочку. Как мы увидим, состояния этих электронов в кристалле существенно отличаются от их состояния в атоме. В частности, эти состояния уже не являются локализованными, и их энергетический спектр состоит из непрерывных зон, а не из дискретных уровней.

Электронные состояния внутренних атомных оболочек при образовании кристалла меняются мало, так как радиус их локализации в атоме мал по сравнению как с радиусом атома, так и межатомным расстоянием в кристалле, а их энергия связи в атоме много больше энергии взаимодействия с ядрами и электронами других атомов. В большинстве случаев эти состояния по-прежнему можно считать локализованными, а энергетический спектр — дискретным.

Строго говоря, состояние кристалла полностью описывается волновой функцией, которая зависит от координат всех частиц, его составляющих: как электронов, так и атомных ядер. Однако масса ядер намного больше массы электронов. Поэтому, если мы интересуемся электронными степенями свободы, то координаты ядер можно считать точно определенными (точнее — координаты ионных остовов, то есть ядер с электронами внутренних оболочек, сильно связанных с ядром). Если пока не принимать во внимание колебания кристаллической решетки, то ядра будут располагаться в ее узлах. Задача таким образом сводится к решению уравнения Шредингера для волновой функции многих электронов  $\Psi(\mathbf{r}_1, \mathbf{r}_1, \dots, t)$ , находящихся в периодическом поле

ионных остовов. При этом необходимо учесть достаточно сильное взаимодействие электронов друг с другом, поэтому решить такую задачу из первых принципов практически невозможно.

Оказывается, однако, что хорошие результаты дает одноэлектронное приближение. В этом приближении считается, что взаимодействие одного, выделенного, электрона со всеми остальными можно приближенно описать с помощью эффективного потенциала, который действует на выделенный электрон со стороны всех остальных. Иными словами, предполагается, что потенциальная энергия электрона в кристалле,  $U(\mathbf{r})$ , складывается из двух слагаемых:  $U(\mathbf{r}) = U_a(\mathbf{r}) + V(\mathbf{r})$ . Первое слагаемое  $U_a$  — это энергия взаимодействия данного электрона с ядрами атомов, второе слагаемое  $V$  — эффективная энергия его взаимодействия с другими электронами кристалла.

Таким образом, состояние каждого электрона описывается уравнением Шредингера с эффективным гамильтонианом

$$\hat{H} = \frac{\hat{p}^2}{2m} + U(\mathbf{r}). \quad (4.1)$$

Эту задачу можно решать самосогласованно: находится состояние каждого электрона в эффективном потенциале, после чего можно вычислить сам эффективный потенциал для каждого состояния. При этом вычисленный потенциал должен совпадать с исходным эффективным потенциалом. Описанный метод называется методом Хартри–Фока, он, в частности, применяется для расчета состояний многоэлектронных атомов.

Нас сейчас, однако, интересуют общие свойства электронных состояний в кристалле. Поэтому нам не важен конкретный вид эффективной потенциальной энергии  $U(\mathbf{r})$ , тем более, что уравнение Шредингера с произвольной  $U(\mathbf{r})$  все равно не решается аналитически. Важно то, что  $U(\mathbf{r})$  обладает симметрией кристалла, в частности,  $U(\mathbf{r})$  — периодическая функция, обладающая периодом кристаллической решетки. Эта периодичность определяет общий вид волновых функций стационарных состояний электронов.

## 4.2. Теорема Блоха

Пусть волновая функция частицы  $\psi(\mathbf{r})$  удовлетворяет стационарному уравнению Шредингера

$$\hat{H}\psi = E\psi, \quad \text{где} \quad \hat{H} = \frac{\hat{p}^2}{2m} + U(r), \quad (4.2)$$

где  $U(\mathbf{r})$  — периодическая функция, обладающая периодом кристаллической решетки.

Теорема Блоха утверждает, что состояния электрона в таком потенциале определяются двумя квантовыми числами. Одно квантовое

число непрерывное — это волновой вектор  $\mathbf{k}$ <sup>1)</sup>. Второе принимает дискретные значения — это так называемый номер зоны  $n$ . Понятно, что энергия состояния тоже зависит от этих двух чисел:  $E = E_{n\mathbf{k}}$ . При фиксированном номере зоны энергия  $E_{n\mathbf{k}}$  является непрерывной функцией волнового вектора  $\mathbf{k}$ .

Сама волновая функция, соответствующая квантовым числам  $n, \mathbf{k}$ , имеет вид

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}). \quad (4.3)$$

Здесь  $u_{n\mathbf{k}}(\mathbf{r})$  — периодическая функция с периодом, равным периоду решетки кристалла. Эта функция зависит от  $n$  и  $\mathbf{k}$  как от параметров и называется блоховской амплитудой.

Таким образом, волновая функция электрона в периодическом потенциале сама не является периодической, за исключением случая, когда экспонента  $e^{i\mathbf{k}\mathbf{r}}$  периодична. Она представляет собой «коротковолновое» колебание  $u_{n\mathbf{k}}(\mathbf{r})$  с периодом, равным периоду решетки, модулированное «длинноволновой» плоской волной  $e^{i\mathbf{k}\mathbf{r}}$ . Это и неудивительно: волновая функция свободного электрона также является не константой, а плоской волной. Как видим, состояние электрона в периодическом потенциале имеет много общего с состоянием свободного электрона. Отличие заключается в умножении плоской волны на периодическую блоховскую амплитуду  $u_{n\mathbf{k}}(\mathbf{r})$ . Вид вещественной части волновой функции при  $k = 0$  и при малых  $k$  ( $ka \ll 1$ ,  $a$  — постоянная решетки) показан на рис. 4.1.

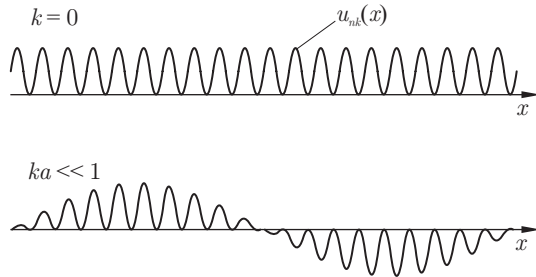


Рис. 4.1. Зависимость вещественной части волновой функции от координаты при  $k = 0$  (а) и при  $ka \ll 1$  (б)

Заметим, что теорема Блоха ничего не говорит о блоховской амплитуде  $u_{n\mathbf{k}}$ , за исключением того, что она периодическая. Вид  $u_{n\mathbf{k}}$  зависит от конкретного вида потенциала  $U(\mathbf{r})$ .

<sup>1)</sup> В кристалле  $\mathbf{k}$  следует называть квазиволновым, а не волновым вектором, так как он обладает некоторыми особенностями, не свойственными волновому вектору свободной частицы; например, при столкновении сумма квазиимпульсов сохраняется с точностью до произвольного вектора обратной решетки (см. главу 2).

До сих пор мы имели дело с координатной частью стационарной волновой функции  $\psi(\mathbf{r})$ . Аналогия со свободным электроном становится еще более ясной, если рассмотреть зависимость волновой функции от времени

$$\Psi_{n\mathbf{k}}(\mathbf{r}, t) = \psi_{n\mathbf{k}}(\mathbf{r}) \exp\left(-\frac{i}{\hbar} E_{n\mathbf{k}} t\right) = u_{n\mathbf{k}} e^{i\mathbf{k}\mathbf{r} - i\omega_{n\mathbf{k}} t}. \quad (4.4)$$

Здесь, как обычно,  $\omega_{n\mathbf{k}} = E_{n\mathbf{k}}/\hbar$ . Итак, волновая функция электрона в периодическом потенциале представляет собой свободно распространяющуюся плоскую гармоническую волну, «заполненную» блоховской амплитудой  $u_{n\mathbf{k}}$ . Таким образом, электроны проходят через идеальный кристалл не рассеиваясь, подобно свободным частицам. Скорость электрона — это скорость волнового пакета, составленного из таких плоских гармонических волн, или групповая скорость

$$\mathbf{v}_g = \frac{\partial \omega_{n\mathbf{k}}}{\partial \mathbf{k}} = \frac{\partial E_{n\mathbf{k}}}{\hbar \partial \mathbf{k}} = \frac{\partial E_{n\mathbf{k}}}{\partial \mathbf{p}}. \quad (4.5)$$

Введя квазиимпульс электрона в кристаллической решетке  $\mathbf{p} = \hbar \mathbf{k}$ , мы получили такое же выражение для скорости электрона через энергию, как и в случае свободного пространства. Однако зависимость энергии от волнового вектора, в отличие от свободного электрона, нам пока неизвестна. Более строгий вывод формулы (4.5) см. ниже.

Рассматривая упругие колебания решетки, мы пришли к выводу, что волновые векторы, отличающиеся друг от друга на произвольный вектор обратной решетки, соответствуют одному и тому же колебанию. А так как любой волновой вектор можно перенести в первую зону Бриллюэна, если прибавить к нему некоторый вектор обратной решетки, то для описания всех колебаний достаточно взять волновые вектора только из первой зоны Бриллюэна.

Подобным же образом обстоит дело и с волновыми векторами блоховских функций электронов в кристалле. Пусть волновой вектор  $\mathbf{k}$  лежит вне первой зоны Бриллюэна. Тогда его можно представить в виде  $\mathbf{k} = \mathbf{k}' + \mathbf{G}$ , где  $\mathbf{k}'$  — волновой вектор, лежащий в первой зоне Бриллюэна,  $\mathbf{G}$  — вектор обратной решетки. Для блоховской волновой функции с волновым вектором  $\mathbf{k}$  получаем следующее выражение:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}'\mathbf{r}} (e^{i\mathbf{G}\mathbf{r}} u_{n\mathbf{k}}) = e^{i\mathbf{k}'\mathbf{r}} u_{n'\mathbf{k}}. \quad (4.6)$$

Нетрудно видеть, что выражение, стоящее в скобках в правой части (4.6), есть периодическая функция с периодом кристаллической решетки. Другими словами, эту функцию можно рассматривать как блоховскую амплитуду  $u_{n'\mathbf{k}}$ , которая соответствует некоторому номеру зоны  $n'$ , отличному от  $n$ . Таким образом, любую волновую функцию электрона в кристалле можно записать в форме (4.3), используя волновые вектора только из первой зоны Бриллюэна.

**Доказательство теоремы Блоха.** Запишем уравнение Шредингера для электрона в периодическом потенциале

$$\left[ \frac{\hat{p}^2}{2m_0} + U(\mathbf{r}) \right] \Psi = E\Psi. \quad (4.7)$$

Здесь  $\hat{p} = -i\hbar\nabla$  — оператор импульса;  $m_0$  — масса свободного электрона.

Будем искать  $\Psi$  в виде разложения в ряд Фурье

$$\Psi = \sum_{\mathbf{q}} c_{\mathbf{q}} e^{i\mathbf{q}\mathbf{r}}, \quad (4.8)$$

где  $\mathbf{q}$  принимает все возможные значения. (Чтобы написать разложение в виде суммы, а не интеграла, необходимо использовать граничные условия Борна–Кармана.)

Неизвестными величинами теперь являются коэффициенты Фурье  $c_{\mathbf{q}}$ .

Оператор кинетической энергии действует на ряд Фурье следующим образом:

$$\frac{\hat{p}^2}{2m_0} \Psi = \sum_{\mathbf{q}} c_{\mathbf{q}} \frac{\hbar^2 q^2}{2m_0} e^{i\mathbf{q}\mathbf{r}}. \quad (4.9)$$

Действие оператора потенциальной энергии  $U$  сводится к простому умножению.  $U(\mathbf{r})$  — периодическая функция с периодом решетки кристалла, поэтому ее можно разложить в ряд Фурье по векторам обратной решетки:

$$U(\mathbf{r}) = \sum_{\mathbf{G}} U_{\mathbf{G}} e^{i\mathbf{G}\mathbf{r}}. \quad (4.10)$$

Умножая  $U$  на  $\Psi$ , получаем

$$U(\mathbf{r})\Psi(\mathbf{r}) = \sum_{\mathbf{q}, \mathbf{G}} c_{\mathbf{q}} U_{\mathbf{G}} e^{i(\mathbf{q}+\mathbf{G})\mathbf{r}}. \quad (4.11)$$

Введя новый вектор  $\mathbf{q}' = \mathbf{q} + \mathbf{G}$ , перепишем это выражение в следующем виде:

$$U(\mathbf{r})\Psi(\mathbf{r}) = \sum_{\mathbf{q}', \mathbf{G}} c_{\mathbf{q}'-\mathbf{G}} U_{\mathbf{G}} e^{i\mathbf{q}'\mathbf{r}} = \sum_{\mathbf{q}, \mathbf{G}} e^{i\mathbf{q}\mathbf{r}} \left( \sum_{\mathbf{G}} c_{\mathbf{q}-\mathbf{G}} U_{\mathbf{G}} \right). \quad (4.12)$$

При переходе к последнему равенству мы заменили  $\mathbf{q}'$  на  $\mathbf{q}$ , поскольку неважно, как именно обозначается индекс суммирования.

Подставляя (4.9) и (4.12) в уравнение Шредингера (4.7), получим

$$\sum_{\mathbf{q}} e^{i\mathbf{q}\mathbf{r}} \left[ \frac{\hbar^2 q^2}{2m_0} c_{\mathbf{q}} + \sum_{\mathbf{G}} c_{\mathbf{q}-\mathbf{G}} U_{\mathbf{G}} \right] = \sum_{\mathbf{q}} E c_{\mathbf{q}} e^{i\mathbf{q}\mathbf{r}}. \quad (4.13)$$

Но два ряда Фурье равны тогда и только тогда, когда равны между собой соответствующие коэффициенты Фурье, поэтому для любого  $\mathbf{q}$

$$\frac{\hbar^2 q^2}{2m_0} c_{\mathbf{q}} + \sum_{\mathbf{G}} c_{\mathbf{q}-\mathbf{G}} U_{\mathbf{G}} = E c_{\mathbf{q}}. \quad (4.14)$$

Мы получили бесконечную систему линейных уравнений для определения коэффициентов  $c_{\mathbf{q}}$ . Энергия электронного состояния входит в эту систему как параметр; значения энергии, при которых система имеет решения, являются разрешенными. Если бы система была конечной, мы могли бы найти разрешенные значения энергии обычным способом, приравняв определитель нулю. После этого, подставляя разрешенные значения энергии в систему, можно было бы найти наборы коэффициентов  $c_{\mathbf{q}}$ , описывающие состояния с этими значениями энергии.

С бесконечной системой все сложнее. Однако данная система обладает интересным свойством. Заметим, что уравнения этой системы связывают коэффициент  $c_{\mathbf{q}}$  только с коэффициентами типа  $c_{\mathbf{q}+\mathbf{G}}$ , где  $\mathbf{G}$  — произвольный вектор обратной решетки.

Таким образом, исходная система разбивается на бесконечное множество бесконечных подсистем, каждая из которых связывает между собой компоненты Фурье, принадлежащие определенному подмножеству. Каждое такое подмножество однозначно задается вектором  $\mathbf{q}_1$ , лежащим в первой зоне Бриллюэна, и состоит из коэффициентов типа  $c_{\mathbf{q}_1+\mathbf{G}}$ , где  $\mathbf{G}$  — произвольный вектор обратной решетки. Между собой эти подмножества не пересекаются и, в силу всего сказанного, являются независимыми.

Следовательно, и независимые решения системы имеют вид

$$\begin{aligned} \Psi &= c_{\mathbf{q}_1} e^{i\mathbf{q}_1\mathbf{r}} + c_{\mathbf{q}_1+\mathbf{G}_1} e^{i(\mathbf{q}_1+\mathbf{G}_1)\mathbf{r}} + c_{\mathbf{q}_1+\mathbf{G}_2} e^{i(\mathbf{q}_1+\mathbf{G}_2)\mathbf{r}} + \dots = \\ &= e^{i\mathbf{q}_1\mathbf{r}} \left[ \sum_{\mathbf{G}} c_{\mathbf{q}_1+\mathbf{G}} e^{i\mathbf{G}\mathbf{r}} \right]. \end{aligned} \quad (4.15)$$

Нетрудно заметить, что мы пришли к выражению типа (4.3), так как сумма в квадратных скобках является периодической функцией, то есть выполняет роль блоховской амплитуды  $u_{\mathbf{q}_1}(\mathbf{r})$ .

Так как подсистема уравнений, определяющая множество коэффициентов  $\{c_{\mathbf{q}_1+\mathbf{G}}\}$ , бесконечна, то она будет разрешимой при бесконечном наборе значений параметра  $E$ ; эти значения мы пронумеруем индексом  $n$ . Таким образом, для любого  $\mathbf{q}_1$  из первой зоны Бриллюэна мы имеем бесконечное множество разрешенных значений  $E_{n\mathbf{q}_1}$ , каждому из которых соответствует (вообще говоря, не одно) решение уравнения Шредингера, имеющее вид

$$\Psi = e^{i\mathbf{q}_1\mathbf{r}} u_{n\mathbf{q}_1}(\mathbf{r}), \quad (4.16)$$

дискретный индекс  $n$  называется номером зоны. С точностью до обозначений ( $\mathbf{q}_1$  вместо  $\mathbf{k}$ ) это выражение совпадает с выражением (4.3).

Итак, теорема Блоха доказана. Электронные состояния в периодическом поле определяются двумя квантовыми числами: дискретным номером зоны  $n$  и непрерывным волновым вектором  $\mathbf{k}$ , лежащим в первой зоне Бриллюэна. Волновая функция состояния  $\{n, \mathbf{k}\}$  имеет вид (4.3).

Чтобы определить динамику электронов в кристалле необходимо доказать еще две теоремы.

### 4.3. Теорема о скорости и ускорении

Пусть электрон находится в состоянии с квантовыми числами  $n$  и  $\mathbf{k}$ . Теорема о скорости утверждает, что квантовомеханическая средняя скорость электрона в блоховском состоянии  $(n, \mathbf{k})$  равна

$$\langle \mathbf{v} \rangle = \frac{1}{\hbar} \frac{\partial E_{n,\mathbf{k}}}{\partial \mathbf{k}}, \quad (4.17)$$

где  $E_{n,\mathbf{k}}$  — энергия электрона в состоянии  $\{n, \mathbf{k}\}$ . Под угловыми скобками мы будем понимать квантовомеханическое среднее. Согласно общей формуле для средних значений в квантовой механике средняя скорость электрона в состоянии  $(n, \mathbf{k})$  имеет вид

$$\langle \mathbf{v} \rangle = \int \psi_{n,\mathbf{k}}^*(\mathbf{r}) \hat{\mathbf{v}} \psi_{n,\mathbf{k}}(\mathbf{r}) d^3r. \quad (4.18)$$

Оператор скорости (по определению равен)

$$\hat{\mathbf{v}} = \frac{\hat{\mathbf{p}}}{m_0} = -\frac{i\hbar}{m_0} \nabla.$$

Интегрирование в (4.18) ведется по всему кристаллу. Волновая функция  $\psi_{n,\mathbf{k}}$  согласно теореме Блоха дается выражением (4.3). Если подействовать оператором скорости  $\hat{\mathbf{v}}$  на блоховскую функцию (4.3), то получаем:

$$\hat{\mathbf{v}} \psi_{n,\mathbf{k}} = \frac{\hbar \mathbf{k}}{m_0} e^{i\mathbf{k}\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) + e^{i\mathbf{k}\mathbf{r}} \left( -\frac{i\hbar}{m_0} \nabla u_{n\mathbf{k}}(\mathbf{r}) \right). \quad (4.19)$$

Если бы в (4.19) было только первое слагаемое, то для средней скорости получили то же соотношение, что и в классической механике:  $\langle \mathbf{v} \rangle = \mathbf{p}/m_0$ .

Однако при наличии второго члена в (4.19) выражение для средней скорости электрона принимает другой вид. Ниже, при вычислении средней скорости, мы воспользуемся условием нормировки для волновых функций

$$\int \psi_{n,\mathbf{k}}^*(\mathbf{r}) \psi_{n,\mathbf{k}}(\mathbf{r}) d^3r = 1. \quad (4.20)$$

Подставим (4.19) в (4.18), тогда получим

$$\langle \mathbf{v} \rangle = \frac{\hbar \mathbf{k}}{m_0} - \frac{i\hbar}{m_0} \int u_{n\mathbf{k}}^* \nabla u_{n\mathbf{k}} d^3r \equiv \frac{\hbar \mathbf{k}}{m_0} + \int u_{n\mathbf{k}}^* \frac{\hat{\mathbf{p}}}{m_0} u_{n\mathbf{k}} d^3r. \quad (4.21)$$

Оказывается, интеграл в (4.21) можно выразить через энергию электрона. Для доказательства этого утверждения рассмотрим две леммы.

**Лемма 1.** Уравнение для блоховской амплитуды. Рассмотрим уравнение Шредингера

$$\left[ \frac{\hat{\mathbf{p}}^2}{2m_0} + U(\mathbf{r}) \right] \psi_{n,\mathbf{k}}(\mathbf{r}) = E_{n,\mathbf{k}} \psi_{n,\mathbf{k}}(\mathbf{r}). \quad (4.22)$$

Подставим в (4.22) выражение для волновой функции в виде (4.3):  $\psi_{n,\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r})$ . Тогда уравнение Шредингера для блоховской амплитуды имеет вид

$$\left[ \frac{(\hat{\mathbf{p}} + \hbar \mathbf{k})^2}{2m_0} + U(\mathbf{r}) \right] u_{n\mathbf{k}} = E_{n,\mathbf{k}} u_{n\mathbf{k}}. \quad (4.23)$$

**Лемма 2.** Теорема Фейнмана. Рассмотрим стационарное уравнение Шредингера

$$\hat{H} \psi = E \psi. \quad (4.24)$$

Пусть гамильтониан  $\hat{H}$  зависит от некоторого произвольного параметра  $\alpha$ . Тогда от этого же параметра  $\alpha$  будут зависеть энергия  $E$  и волновые функции  $\psi$ . Тогда утверждается следующее тождество:

$$\int \psi^*(\mathbf{r}) \frac{\partial \hat{H}}{\partial \alpha} \psi(\mathbf{r}) d^3r = \frac{\partial E}{\partial \alpha}, \quad (4.25)$$

что означает, что среднее значение производной гамильтониана по параметру равно производной от энергии по этому параметру.

**Доказательство.** В квантовой механике среднее значение энергии в состоянии с волновой функцией  $\psi$  равно

$$E = \int \psi^*(\mathbf{r}) \hat{H} \psi(\mathbf{r}) d^3r. \quad (4.26)$$

Продифференцируем (4.26) по параметру  $\alpha$ :

$$\frac{\partial E}{\partial \alpha} = \int \frac{\partial \psi^*(\mathbf{r})}{\partial \alpha} \hat{H} \psi(\mathbf{r}) d^3r + \int \psi^*(\mathbf{r}) \hat{H} \frac{\partial \psi(\mathbf{r})}{\partial \alpha} d^3r + \int \psi^*(\mathbf{r}) \frac{\partial \hat{H}}{\partial \alpha} \psi(\mathbf{r}) d^3r. \quad (4.27)$$

Докажем, что первые два члена в (4.27) в сумме дают ноль:

$$\begin{aligned} & \int \frac{\partial \psi^*(\mathbf{r})}{\partial \alpha} \hat{H} \psi(\mathbf{r}) d^3r + \int \psi^*(\mathbf{r}) \hat{H} \frac{\partial \psi(\mathbf{r})}{\partial \alpha} d^3r = \\ & = E \left\{ \int \frac{\partial \psi^*}{\partial \alpha} \psi d^3r + \int \psi^* \frac{\partial \psi}{\partial \alpha} d^3r \right\} = E \frac{\partial}{\partial \alpha} \int |\psi|^2 d^3r = 0. \end{aligned} \quad (4.28)$$

При выводе (4.28) мы воспользовались уравнением (4.24) и свойством самосопряженности оператора гамильтониана  $\hat{H}$ :

$$\int \psi^* H \frac{\partial \psi}{\partial \alpha} d^3 r = \int \frac{\partial \psi}{\partial \alpha} \hat{H} \psi^* d^3 r. \quad (4.29)$$

Таким образом, уравнение (4.27) дает доказательство теоремы Фейнмана (4.25).

Теперь вернемся к уравнению (4.23) для блоховской амплитуды  $u_{n\mathbf{k}}$ . Применим к этому уравнению теорему Фейнмана. В этом случае роль волновой функции играет  $u_{n\mathbf{k}}$ , а в качестве параметра рассмотрим волновой вектор  $\mathbf{k}$ . Перепишем (4.23) в виде

$$\hat{H} u_{n\mathbf{k}} = E_{n\mathbf{k}} u_{n\mathbf{k}}; \quad \hat{H} = \frac{(\hat{\mathbf{p}} + \hbar \mathbf{k})^2}{2m_0} + U(\mathbf{r}). \quad (4.30)$$

Очевидно, что

$$\frac{\partial \hat{H}}{\partial \mathbf{k}} = \frac{\hat{\mathbf{p}} + \hbar \mathbf{k}}{m_0} \hbar. \quad (4.31)$$

Тогда, согласно теореме Фейнмана (4.25), получим

$$\hbar \int u_{n\mathbf{k}}^* \frac{\partial \hat{H}}{\partial \mathbf{k}} u_{n\mathbf{k}} d^3 r = \frac{\partial E_{n\mathbf{k}}}{\partial \mathbf{k}}. \quad (4.32)$$

Сравнивая левую часть (4.32) с выражением (4.21) для средней скорости  $\langle \mathbf{v} \rangle$  видим, что

$$\langle \mathbf{v} \rangle = \frac{1}{\hbar} \frac{\partial E_{n\mathbf{k}}}{\partial \mathbf{k}}. \quad (4.33)$$

Это есть точное выражение для средней квантовомеханической скорости электрона в кристалле. В дальнейшем для нас будут важны экстремумы энергии как функции  $\mathbf{k}$  в зоне  $n$ . В точках минимума или максимума энергии скорость  $\langle \mathbf{v} \rangle$  равна нулю, то есть электрон покоится.

Итак, мы доказали теорему о скорости. Ниже аналогично рассмотрим теорему об ускорении.

При наличии внешнего электрического поля состояние электрона становится нестационарным. Следовательно, необходимо использовать для электрона временное уравнение Шредингера

$$i\hbar \frac{\partial \psi}{\partial t} = \left[ \frac{\hat{\mathbf{p}}^2}{2m_0} + U(\mathbf{r}) - e\mathbf{E}\mathbf{r} \right] \psi. \quad (4.34)$$

Последний член в квадратных скобках учитывает влияние электрического поля на электрон. Если выбрать волновую функцию электрона в виде

$$\psi = u_{n\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}(t)\mathbf{r}} \exp\left(-\frac{i}{\hbar} \int_0^t E_{n\mathbf{k}} dt'\right) \quad (4.35)$$

и подставить в уравнение Шредингера (4.34), возникают слагаемые от дифференцирования блоховской амплитуды  $u_{n\mathbf{k}}$  по времени. Можно показать, что волновая функция (4.35) точно удовлетворяет уравнению (4.34), если использовать соотношение <sup>1)</sup>

$$\frac{d\hbar \mathbf{k}}{dt} = e\mathbf{E}. \quad (4.36)$$

Составим классическую функцию Гамильтона от переменных  $\mathbf{r}$  и  $\mathbf{p}$ :  $H(\mathbf{r}, \mathbf{p})$ , где  $\mathbf{p} = \hbar \mathbf{k}$  — квазиимпульс электрона,  $\mathbf{r}$  — радиус-вектор электрона,

$$H(\mathbf{r}, \mathbf{p}) = E_{n\mathbf{k}} + V(\mathbf{r}). \quad (4.37)$$

Здесь  $E_{n\mathbf{k}}$  играет роль аналога кинетической энергии электрона (но  $E_{n\mathbf{k}} \neq \mathbf{p}^2/2m_0$ ),

$$V(\mathbf{r}) = -e\mathbf{E}\mathbf{r} \quad (4.38)$$

есть потенциальная энергия электрона во внешнем поле.

Теорема о скорости утверждает, что

$$\dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{p}}. \quad (4.39)$$

А теорема об ускорении утверждает, что

$$\dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{r}}. \quad (4.40)$$

Следовательно (4.39) и (4.40) — это не что иное, как уравнения Гамильтона

$$\dot{\mathbf{r}} = \frac{\partial H}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{r}}. \quad (4.41)$$

Итак, при изучении классической динамики электронов из уравнений движения полностью выпал потенциал кристаллической решетки. В уравнения Гамильтона мы подставляем энергию  $E_{n\mathbf{k}}$ , которая играет роль кинетической энергии. В качестве потенциальной энергии используется энергия электрона во внешнем поле.

**Вывод.** Мы можем рассматривать электрон в кристалле под действием внешних электрических полей как свободный электрон в поле  $V(\mathbf{r})$ , как и в классической механике, с функцией Гамильтона  $H = E_{n\mathbf{k}} + V(\mathbf{r})$ . Но разница между электроном в кристалле и свободным электроном состоит в том, что вместо кинетической энергии для свободного электрона  $\mathbf{p}^2/2m_0$  мы пишем  $E_{n\mathbf{k}}$ .

Как правило, кинетическая энергия  $E_{n\mathbf{k}}$  хорошо известна, и поэтому с электроном в кристалле можно обращаться как со свободным

<sup>1)</sup> Строгое доказательство см. в книге: Киттель Ч. Квантовая теория твердых тел. — М.: Наука, 1967. 492 с.

электроном при наличии внешнего потенциала  $V(\mathbf{r})$ . Это вывод, который следует из теоремы Блоха<sup>1)</sup>.

**Примечание** (к утверждению о том, что мы можем в кристалле рассматривать электрон как свободный и что только кинетическая энергия  $E_{n\mathbf{k}}$  иначе зависит от импульса). Мы этот вывод сделали для случая, когда внешнее поле  $E$  является однородным и постоянным:  $d\mathbf{p}/dt = e\mathbf{E}$ . Но если поле неоднородное, то есть зависит от координаты, тогда  $d\mathbf{p}/dt = \mathbf{F}$ , а  $\mathbf{F} = -\partial V/\partial \mathbf{r}$ , где  $V$  — потенциальная энергия электрона во внешних полях. Мы доказали теорему об ускорении, когда  $\mathbf{F} = \text{const}$ . Но если сила  $\mathbf{F}$  слабо меняется на длине постоянной решетки, то все, что было сказано выше, приближенно будет правильно. Наше рассмотрение, в котором пренебрегаем производной от блоховской амплитуды, справедливо для слабых полей. Для сильных электрических полей эти результаты не применимы. В сильных полях под действием внешнего поля происходят межзонные переходы (зине-ровский пробой). По тем же причинам наше рассмотрение верно, если частоты внешних воздействий малы по сравнению с  $(E_{n\mathbf{k}} - E_{n'\mathbf{k}})/\hbar$ .

Итак, мы получим систему уравнений для описания движения электрона во внешних полях

$$\frac{d\mathbf{p}}{dt} = -\frac{\partial V}{\partial \mathbf{r}}, \quad \mathbf{v} = \frac{1}{\hbar} \frac{\partial E_{n\mathbf{k}}}{\partial \mathbf{k}}. \quad (4.42)$$

Здесь  $V$  — потенциальная энергия электрона во внешних полях.

Подход, использующий систему уравнений (4.42), называется **методом эффективной массы**.

Решетка кристалла присутствует в (4.42) только в зависимости энергии  $E_{n\mathbf{k}}$  от квазиимпульса. Поэтому зависимость энергии электрона  $E_{n\mathbf{k}}$  от квазиимпульса становится очень важной.

Для полупроводников главную роль играют экстремумы функции  $E_{n\mathbf{k}}$ . Мы дальше опускаем индекс  $n$  и не будем оговаривать, о какой зоне идет речь. Исследуем энергию  $E_{\mathbf{k}}$  вблизи экстремума, причем волновой вектор  $\mathbf{k}$ , как уже было показано, изменяется в пределах первой зоны Бриллюэна. Рассмотрим кубический кристалл. Пусть, для простоты, минимум энергии оказывается в центре зоны Бриллюэна (для многих полупроводников  $A_3B_5$  минимум энергии действительно расположен в точке  $\Gamma(\mathbf{k} = 0)$  зоны Бриллюэна). Тогда мы можем разложить энергию  $E_{\mathbf{k}}$  вблизи минимума в ряд Тейлора

$$E_{\mathbf{k}} = E_0 + \left( \frac{\partial E_{\mathbf{k}}}{\partial k_x} \right)_{k=0} k_x + \left( \frac{\partial E_{\mathbf{k}}}{\partial k_y} \right)_{k=0} k_y + \left( \frac{\partial E_{\mathbf{k}}}{\partial k_z} \right)_{k=0} k_z + \frac{1}{2!} \sum_{ij} \left( \frac{\partial^2 E_{\mathbf{k}}}{\partial k_i \partial k_j} \right)_{k=0} k_i k_j + \dots \quad (4.43)$$

Так как мы предполагаем, что минимум энергии расположен в точке  $\Gamma(k = 0)$ , то первые производные обращаются в нуль. Оставляя второе слагаемое, энергию можно записать в виде:

$$E_{\mathbf{k}} = E_0 + \frac{1}{2} \sum_{ij} A_{ij} k_i k_j. \quad (4.44)$$

Здесь  $E_0$  — значение энергии в минимуме;  $A_{ij}$  — тензор второго ранга. Однако в кубическом кристалле все тензоры второго ранга сводятся к скаляру:

$$A_{ij} = A \cdot \delta_{ij}. \quad (4.45)$$

В результате для энергии получаем простое выражение

$$E_{\mathbf{k}} = E_0 + \frac{1}{2} A (k_x^2 + k_y^2 + k_z^2) = E_0 + \frac{1}{2} A k^2. \quad (4.46)$$

Чтобы привести это выражение к стандартному виду, введем эффективную массу электрона  $m$ :

$$A \equiv \hbar^2/m — \text{это есть определение эффективной массы.} \quad (4.47)$$

Тогда

$$E_{\mathbf{k}} = E_0 + \frac{\hbar^2 \mathbf{k}^2}{2m}. \quad (4.48)$$

Мы получили закон дисперсии как для свободного электрона, только вместо массы  $m_0$  стоит эффективная масса  $m$ . Как уже было отмечено выше,  $E_0$  — это значение энергии в минимуме, откуда ведется отсчет энергии ( $E_0 = \text{const}$ ).

Итак, при слабых внешних полях электрон в кристалле ведет себя как свободный электрон с эффективной массой  $m$ . В самом деле, скорость электрона с учетом закона дисперсии (4.48) равна

$$\mathbf{v} = \frac{1}{\hbar} \frac{\partial E_{\mathbf{k}}}{\partial \mathbf{k}} = \frac{\hbar \mathbf{k}}{m} = \frac{\mathbf{p}}{m}, \quad (4.49)$$

где  $\mathbf{p}$  — квазиимпульс,  $m$  — эффективная масса. Согласно закону Ньютона,

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}, \quad \text{тогда} \quad m \frac{d\mathbf{v}}{dt} = \mathbf{F}. \quad (4.50)$$

Мы получили уравнение движения такое же, как для свободной частицы с эффективной массой  $m$ . Потенциал решетки учитывается через эффективную массу  $m$ .

Рассмотрим более сложный случай, когда минимум энергии не находится в центре зоны Бриллюэна. Например, в кремнии (Si) минимумы зоны проводимости находятся на осях куба (это линии  $\Delta$ ) вблизи края зоны Бриллюэна (но не на самом краю). В этом случае следует разложить энергию в ряд Тейлора вблизи точки минимума  $\mathbf{k}_0$ .

<sup>1)</sup> Это не так вблизи точек вырождения.

Разложение в ряд Тейлора (до второго порядка) имеет вид квадратичной формы по  $\mathbf{k}$ :

$$E_{\mathbf{k}} = E_0 + \frac{1}{2} \sum_{ij} A_{ij} (k_i - k_{i0})(k_j - k_{j0}). \quad (4.51)$$

Из курса линейной алгебры известно, что всякую квадратичную форму можно привести к диагональному виду, то есть к сумме квадратов. Тогда выражение для энергии (4.51) можно записать в виде

$$E_{\mathbf{k}} = E_0 + \frac{1}{2} [A_1(k_{x'} - k_{x'0})^2 + A_2(k_{y'} - k_{y'0})^2 + A_3(k_{z'} - k_{z'0})^2]. \quad (4.52)$$

Поскольку мы рассматриваем минимум энергии, то вторые производные от энергии, то есть  $A_1$ ,  $A_2$  и  $A_3$ , положительны. Можно считать, что

$$A_1 = \frac{\hbar^2}{m_x}, \quad A_2 = \frac{\hbar^2}{m_y}, \quad A_3 = \frac{\hbar^2}{m_z}. \quad (4.53)$$

Тогда для энергии получаем окончательное выражение

$$E_{\mathbf{k}} = E_0 + \frac{\hbar^2}{2} \left\{ \frac{(k_{x'} - k_{x'0})^2}{m_x} + \frac{(k_{y'} - k_{y'0})^2}{m_y} + \frac{(k_{z'} - k_{z'0})^2}{m_z} \right\}. \quad (4.54)$$

Мы получили закон дисперсии для случая, когда экстремумы не лежат в точке Г. В этом случае эффективная масса анизотропна.

#### 4.4. Зонная структура полупроводников

Для точного расчета энергетического спектра кристаллов используются довольно сложные методы, такие как метод ортогонализированных плоских волн, метод псевдопотенциала и другие методы.

Ниже для установления общих свойств энергетического спектра  $\epsilon(\mathbf{k})$  рассматриваются два наиболее простых метода расчета зонной структуры кристаллических твердых тел:

- 1) метод сильной связи,
- 2) метод слабой связи.

Почему мы прибегаем к приближенным методам расчета зонной структуры?

В общем случае мы должны решать уравнение Шредингера

$$\left[ \frac{\hat{\mathbf{p}}^2}{2m_0} + U(\mathbf{r}) \right] \psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (4.55)$$

Здесь  $U(\mathbf{r})$  — потенциальная энергия. В общем случае потенциал  $U(\mathbf{r})$  не известен. Однако приближенные модели позволяют выяснить, чем определяется зависимость энергии от волнового вектора, от чего зависит ширина запрещенной зоны; эффективная масса, и т. д.

Есть еще одна модель для описания зонной структуры — это модель Кронига–Пени. Это одномерная модель, которая описывает потенциал в виде прямоугольных ям (рис. 4.2).

Эта модель решается точно. (Мы не будем этого делать.) Если в нулевом приближении в яме есть уровень, то в случае 2 ям уровень расщепляется на два уровня; если есть три квантовых ямы, то уровень расщепляется на три уровня, если есть бесконечное число квантовых ям, то уровни превращаются в полосы, то есть получаются зоны.

Следует отметить, что ни приближение сильно связанных, ни приближение слабо связанных электронов не описывает правильно с количественной точки зрения состояние электронов в зоне проводимости кристалла.

Как будет показано ниже, оба эти приближения не могут быть использованы для количественных расчетов энергетического спектра и волновых функций электронов в конкретных кристаллах. Однако эти два метода дают хорошую иллюстрацию к общим выводам о движении электрона в периодическом поле.

**4.4.1. Приближение сильной связи.** Рассмотрим одномерный кристалл — систему изолированных атомов (ионов). Потенциал такой системы представлен на рис. 4.3. В минимумах потенциальных ям расположены ионы. Потенциальные ямы разделены барьерами. Смысл приближения сильной связи состоит в следующем. Рассмотрим  $i$ -ю потенциальную яму (то есть  $i$ -й ион). Волновую функцию электрона, центрированную на этом  $i$ -м ионе (это отмечено пунктирной линией на рис. 4.3), мы обозначим через  $\varphi(x - x_i)$  (где  $x_i$  — положение  $i$ -го иона). Для электрона, центрированного на  $j$ -м ионе, волновая функция равна  $\varphi(x - x_j)$ .

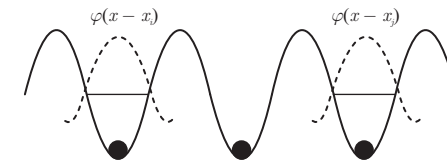


Рис. 4.3. Кривая потенциальной энергии электрона в кристалле. Пунктиром изображена волновая функция электрона, локализованного на атоме

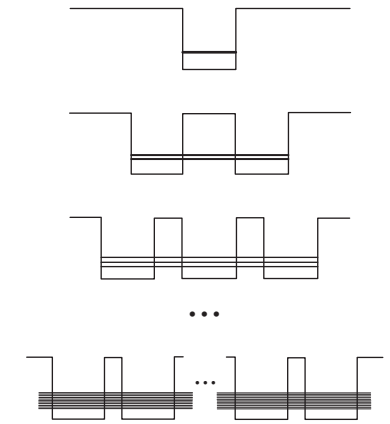


Рис. 4.2. Схема расщепления уровней энергии в нескольких связанных квантовых ямах. Нижний рисунок иллюстрирует образование энергетических зон в кристалле



Рассмотрим случай, когда в каждой яме имеется по одному энергетическому уровню и они одинаковы.

Положение уровня энергии обозначим через  $E$ . Рассмотрим один электрон, который может находиться в любой квантовой яме, поэтому существует вырождение; то есть много состояний, которые могут быть заняты электроном с одинаковой энергией (каждое состояние характеризуется своей волновой функцией, имеющей одну и ту же энергию). Потенциал, представленный на рис. 4.3, состоит из барьеров конечной высоты и конечной ширины. В этом случае нельзя пользоваться теорией возмущений для нахождения спектра. Но нужно построить правильную волновую функцию для такой задачи. Если у нас имеются два уровня, то волновая функция такой системы представляет собой линейную комбинацию двух электронных волновых функций. В случае потенциала, показанного на рис. 4.3, волновую функцию системы следует взять в виде линейной комбинации (несмотря на то, что имеется бесконечное число уровней):

$$\psi_k = \sum_j c_j \varphi(x - x_j). \quad (4.56)$$

Здесь  $c_j$  — коэффициенты разложения. Можно было бы считать  $c_j$  вариационными параметрами и находить их на основе вариационного принципа. Но можно поступить иначе. Можно выбрать коэффициенты  $c_j$  так, чтобы волновая функция имела правильный вид, то есть вид блоховской функции. Пусть  $c_j$  есть волна, то есть  $c_j = e^{ikx_j}$ .

В результате, для волновой функции системы имеем

$$\psi_k = \sum_j e^{ikx_j} \varphi(x - x_j). \quad (4.57)$$

Покажем, что  $\psi_k$  является блоховской волновой функцией; для этого перепишем (4.57) в виде

$$\psi_k = e^{ikx} \sum_j e^{ik(x_j - x)} \varphi(x - x_j) = e^{ikx} u_k(x). \quad (4.58)$$

Здесь  $u_k(x) = \sum_j e^{ik(x_j - x)} \varphi(x - x_j)$  играет роль блоховской амплитуды. Напомним, что  $x_j$  — положение  $j$ -го иона, при этом  $u_k(x)$  — периодическая функция с периодом решетки. Действительно, при смещении  $x$  на несколько узлов решетки (то есть при трансляции на вектор, соединяющий два произвольных узла) функция  $U_k(x)$  останется той же самой, так как сумма не меняется:

$$u_k(x + a) = \sum_j e^{ik(x_j - a - x)} \varphi(x + a - x_j) = \sum_{j'} e^{ik(x_{j'} - x)} \varphi(x - x_{j'}) = u_k(x), \quad (4.59)$$

где  $x_{j'} = x_j - a$ . Этот результат показывает, что функция  $\psi_k(x)$  (4.57) полностью удовлетворяет требованиям, предъявляемым к функциям

Блоха (то есть функции удовлетворяют условиям трансляционной симметрии). В этом случае можно иметь в виду, что одномерная цепочка атомов либо свернута в кольцо, либо имеет бесконечную длину; так что при этом соблюдаются условия периодичности Борна–Кармана.

Зная волновую функцию, мы можем вычислить энергию электронов. Энергию в первом приближении мы найдем, вычислив диагональные матричные элементы гамильтониана кристалла

$$E_k = \int \psi_k^*(x) \hat{H} \psi_k(x) dx. \quad (4.60)$$

Прежде чем вычислить энергию, необходимо нормировать волновые функции  $\psi_k(x)$ .

**Условие нормировки.** Пусть волновая функция  $\psi_k(x)$  нормирована на единицу (мы рассматриваем конечную цепочку длины  $L$ ):

$$\int_{-\infty}^{+\infty} |\psi_k(x)|^2 dx = 1. \quad (4.61)$$

Обозначим нормировочную константу через  $A$ . Тогда волновую функцию  $\psi_k$ , согласно (4.57), представим в виде

$$\psi_k(x) = A \sum_j e^{ikx_j} \varphi(x - x_j). \quad (4.62)$$

Подставив (4.62) в (4.61), получаем

$$\begin{aligned} \int |\psi_k(x)|^2 dx &= \\ &= A^2 \left[ \sum_j \int \varphi^2(x - x_j) dx + \sum_{i \neq j} e^{ik(x_i - x_j)} \int \varphi(x - x_i) \varphi(x - x_j) dx \right]. \end{aligned} \quad (4.63)$$

В первом интеграле (4.63) произведем замену  $x - x_j = x'$ , тогда

$$\int_{-\infty}^{+\infty} \varphi^2(x - x_j) dx = \int_{-\infty}^{+\infty} \varphi^2(x') dx' = 1, \quad (4.64)$$

так как атомные волновые функции  $\varphi(x)$  нормированы на единицу. В сумме по  $j$  в (4.63) число таких интегралов (4.64) равно  $N$ . Здесь  $N$  — это число атомов на нормировочной длине  $L$  (нормировочный объем). Окончательно, для первого интеграла в (4.63) имеем

$$A^2 \sum_j \int \varphi^2(x - x_j) dx = A^2 N. \quad (4.65)$$

Теперь рассмотрим второй интеграл в (4.63). В этот интеграл входит произведение атомных волновых функций на разных узлах  $i$  и  $j$ . Волновая функция  $\varphi(x - x_i)$  заметно отлична от нуля лишь вблизи

$i$ -го атома (иона); соответственно,  $\varphi(x - x_j)$  заметно отлична от нуля вблизи  $j$ -го атома. Если узлы решетки разнесены, то интеграл близок к нулю, поэтому во втором интеграле (4.63) достаточно учитывать только соседние узлы. Вследствие этого такой интеграл называется интегралом перекрытия (рис. 4.4).

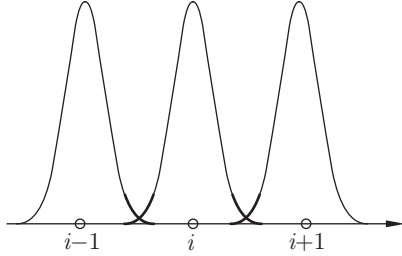


Рис. 4.4. Перекрытие волновых функций электронов локализованных на собственных атомах. Показано (жирные участки) слабое перекрытие волновых функций в рамках приближения сильно связанных электронов

Поскольку решетка периодическая и ее узлы одинаковые, все интегралы перекрытия, входящие в (4.63), равны. Обозначим интеграл перекрытия через  $I$ , тогда

$$I = \int \varphi(x - x_i) \varphi(x - x_{i+1}) dx = \int \varphi(x + a) \varphi(x) dx. \quad (4.66)$$

При выводе (4.66) мы учли, что  $x_{i+1} - x_i = a$ , где  $a$  — постоянная решетки. Кроме того, в (4.66) мы, для простоты, считаем, что атомные волновые функции являются вещественными. Тогда с учетом сказанного для второго слагаемого в (4.63) получим

$$\sum_{i \neq j} e^{ik(x_i - x_j)} \int \varphi(x - x_i) \varphi(x - x_j) dx = A^2 N I (e^{ika} + e^{-ika}). \quad (4.67)$$

Первое слагаемое соответствует соседу справа, а второе — соседу слева для каждого атома.

Подставляя выражения (4.65) и (4.67) в (4.63) получим уравнение для определения постоянной нормировки  $A$

$$\int |\psi_k(x)|^2 dx = A^2 N (1 + 2I \cos ka) = 1, \quad (4.68)$$

$$A = \frac{1}{\sqrt{N} \sqrt{1 + 2I \cos ka}}. \quad (4.69)$$

По своему смыслу интеграл перекрытия  $I$ , как правило, мал; тогда в (4.69) можно разложить квадратный корень в ряд. Таким образом,

$$A = \frac{1}{\sqrt{N}} (1 - I \cos ka). \quad (4.70)$$

**Вычисление энергии.** Переходим теперь к вычислению энергии электронов. Нормированная волновая функция электрона имеет вид (4.62). Подставим эту волновую функцию из (4.62) в (4.60). Итак, для энергии электронов имеем выражение

$$E_k = A^2 \left[ \sum_j \varphi(x - x_j) \hat{H} \varphi(x - x_j) dx + \sum_{i \neq j} e^{ik(x_j - x_i)} \int \varphi(x - x_j) \hat{H} \varphi(x - x_i) dx \right]. \quad (4.71)$$

(Здесь, как и при вычислении  $A$ , мы имеем суммы двух типов: с одинаковыми  $x_j$  и с разными  $x_i$  и  $x_j$ ). Первое слагаемое в (4.71) содержит сумму  $N$  одинаковых интегралов. Мы предполагаем, что второе слагаемое отлично от нуля только при учете ближайших соседей, при этом  $j = i + 1, i - 1$ . Причем, как и выше (при вычислении  $A$ ), все интегралы от произведений функции соседних атомов типа интеграла перекрытия одинаковы. Поэтому для энергии получаем выражение

$$E_k = A^2 N \left[ \int \varphi(x) \hat{H} \varphi(x) dx + 2 \cos ka \int \varphi(x + a) \hat{H} \varphi(x) dx \right]. \quad (4.72)$$

При выводе (4.72) мы поместили начало координат в нуль ( $x_i = 0$ ).

Гамильтониан электрона представим в виде

$$\hat{H} = \hat{T} + \sum_i U(x - x_i). \quad (4.73)$$

Здесь  $\hat{T}$  — оператор кинетической энергии электрона;  $\sum_i U(x - x_i)$  — потенциальная энергия электрона в поле всех центров (самосогласованный периодический потенциал, действующий на электрон). Удобно разбить сумму во втором члене (4.73) на две части. Выделим атом, который находится в начале координат; тогда потенциальную энергию электрона в поле данного атома мы будем обозначать через  $U(x)$ , а потенциальную энергию электрона в поле остальных атомов — через  $\Delta U(x)$ . В результате, гамильтониан (4.73) можно представить в виде

$$\hat{H} = \hat{T} + U(x) + \Delta U(x) \equiv \hat{H}_0 + \Delta U(x). \quad (4.74)$$

В этом случае атомная волновая функция  $\varphi(x)$  (волновая функция электрона в изолированном атоме) является собственной функцией гамильтониана  $\hat{H}_0$ :

$$\hat{H}_0 \varphi(x) = E_0 \varphi(x), \quad (4.75)$$

где  $E_0$  — соответствующий уровень энергии электрона в изолированном атоме. Подставляя (7.53) в (7.51) и учитывая уравнение (7.54),

получим

$$E_k = (1 - 2I \cos ka)[E_0(1 + 2I \cos ka) + \overline{\Delta U} + 2J \cos ka]. \quad (4.76)$$

Здесь мы ввели обозначения:  $\overline{\Delta U} = \int \varphi(x) \Delta U(x) \varphi(x) dx$  — среднее значение потенциальной энергии электрона выделенного атома в поле остальных атомов,

$$J = \int \varphi(x+a) \Delta U(x) \varphi(x) dx. \quad (4.77)$$

Выражение для  $J$  имеет вид интеграла перекрытия (см. (7.45)). Если бы перекрытия между волновыми функциями соседних атомов не было, интеграл в (7.56) обратился бы в нуль; кроме того, если бы  $\Delta U$  отсутствовало, то  $J$  также обратился бы в нуль. Таким образом, из (7.56) следует, что величина  $J$  играет основную роль в определении энергии электрона в рамках метода сильной связи. Выражение для энергии (7.55) можно упростить. При раскрытии скобок опустим слагаемые пропорциональные  $I^2$  и  $J^2$ , тогда для спектра электронов окончательно получаем

$$E_k = E_0 + \overline{\Delta U} + 2J \cos ka. \quad (4.78)$$

Физический смысл полученных результатов состоит в том, что при образовании кристалла из отдельных атомов энергия электрона в изолированном атоме  $E_0$  в результате взаимодействия с соседними атомами расщепляется в энергетическую зону и смещается на величину  $\overline{\Delta U}$  (рис.4.5). При этом в пределах такой зоны энергия электрона периодически зависит от волнового вектора  $k$ . Сдвиг уровня энергии  $\overline{\Delta U}$ , как правило, отрицателен. Отрицательное значение для  $\overline{\Delta U}$  можно обосновать следующим образом. На рис.4.6 представлен потенциал изолированного атома  $U(x)$ , а пунктирной линией показан самосогласованный периодический потенциал, действующий на электрон  $\sum_i U(x - x_i)$ . Если предположить, что взаимодействие

между атомами снижает потенциальные барьеры для электронов, как это показано на рис.4.6, то согласно (4.73) и (4.74),  $\Delta U(x) < 0$  и, следовательно, сам интеграл  $\overline{\Delta U} < 0$ . Такому расщеплению (как показано на рис.) подвергается каждый стационарный энергетический уровень изолированного атома. Ширина разрешенной энергетической зоны, как следует из (4.78), пропорциональна интегралу перекрытия  $J$ , то есть матричному элементу соответствующего перехода электрона к соседнему атому, и равна  $4J$ . Если интегралы перекрытия малы, то стационарный энергетический уровень энергии изолированного атома остается почти неизменным. Формулы, полученные в рамках метода сильной связи, применимы в тех случаях, когда перекрытие атомных оболочек мало.

Проиллюстрируем свойства электрона, движущегося в периодическом поле кристалла, на примере сильно связанных электронов в про-

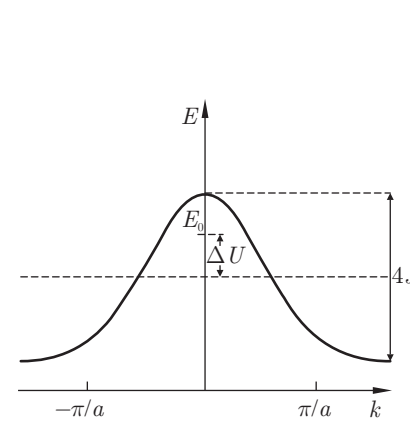


Рис. 4.5. Образование энергетических зон из энергетических уровней при сближении атомов

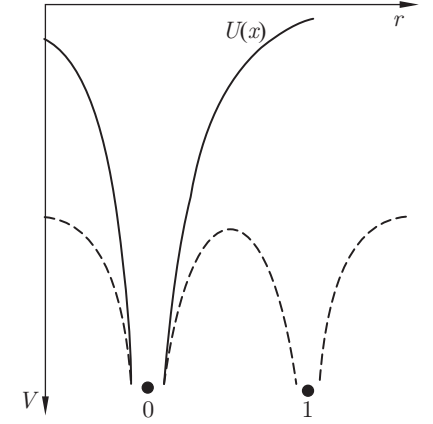


Рис. 4.6. Потенциальная энергия электрона в кристалле, как функция расстояния  $r$ . Сплошной линией представлен потенциал изолированного атома. Пунктирной линией представлен самосогласованный периодический потенциал

стой кубической решетке. Для кубического кристалла эффективная масса электрона определяется выражением

$$A = \frac{\hbar^2}{m_c}, \quad (4.79)$$

где  $A = \partial^2 E_k / \partial k^2$ . Найдем эффективную массу вблизи экстремумов ( $k = 0, k = \pm \pi/a$ ). Энергетический спектр электрона при малых  $k$  определяется, согласно (7.57), выражением

$$E_k \approx E_0 + \overline{\Delta U} + 2J \left(1 - \frac{k^2 a^2}{2}\right). \quad (4.80)$$

Таким образом, эффективная масса электрона у верхнего края зоны равна

$$m_c = -\frac{\hbar^2}{2Ja^2}. \quad (4.81)$$

Во-первых, мы получили, что эффективная масса отрицательна. Во-вторых, как следует из (7.60), чем уже разрешенная зона (то есть чем меньше величина  $J$ ), тем эффективная масса электрона больше. Можно сказать, что экстремум зоны в точке  $k = 0$  соответствует дыркам, тогда эффективная масса дырки равна

$$m_h = -m_c = \frac{\hbar^2}{2Ja^2}.$$

Αίαεϊάε÷ιύϊ ιάδὰçιϊ άυ÷εñεεϊ γόδἀεοεάϊόρ ιαññό γεάεοδρίά ό ίεα•  
 ιάαϊ εδὰγ çίίυ άάεεçε  $k = \frac{a}{a}$ . Υίαδἀάοε÷άñεεε ñίάεοδ γεάεοδρίά άάεεçε  
 γόϊάϊ ιεϊεϊόϊά εϊάάο άεά

$$E_k = E_0 + 4U \quad 2J \quad 1 \quad \frac{a^2(k)^2}{2}, \quad (4.82)$$

άάά  $k = -k$ .

Ñεάαϊάαοάεüíí, άεγ γόδἀεοεάϊέ ιαññύ ό ίεαϊάαϊ εδὰγ çίίυ  
 ιάοϊάεϊ

$$m = \frac{h^2}{2Ja^2}: \quad (4.83)$$

Όάεεϊ ιάδὰçιϊ, άεγ εόάε÷άñεϊάϊ εδενάεεά γόδἀεοεάϊάγ ιαññ ά άüδέε  
 ό άάδøεϊύ çίίυ ñίάϊάάάά (ñ άάεε÷εϊά) ñ γόδἀεοεάϊέ ιαññίε γεάε•  
 οδρίά ό ίεαϊάαϊ εδὰγ çίίυ.

Άñεε άύ ιü δαññιαοδεάαεε ίά ιάεϊ εϊεάεεçíααίíúε όδίααίü γεά ε•  
 οδρίά ίά άοίíá, ά ίαñεϊεüεϊ, όϊ ñεό÷εεε άü, ÷όϊ ÷άϊ άüøά όδία άίü  
 (ñ γίαδἀεε), όαϊ øεδά δαçδἀάίáγ çίíá; ÷άϊ ίεαά ñ γίαδἀεε çίíá,  
 όαϊ ñίá όαεά.

Ιδελίαεϊññόü ñάάεε ñεεüííε ñάγçε. Ά δαίεαδ γόίε ñάάεε  
 ιü δαññιüδάεε ιάεϊ εϊεάεεçíααίíúε όδίααίü γεάεοδρίά ίά εά αεάíí  
 άοίíá. Ñεάαϊάαοάεüíí, δαçδἀάάίíúά çίíý όçεεά, ά çáíδáüáíí üά øε•  
 δίεεά. Όεάγ ñάάεü ιδελίαεϊά άεγ ñεñáíεγ ñίáεοδρίά ñεάεο εγδρίü  
 εδενάεεεϊά, ίεñεϊά, ίí, εάε ιδάάεεϊ, ίά ιδελίαεϊά άεγ ñεñ áíεγ  
 ñίáεοδρίά ñεόíδíáíáíεεϊά. Άεγ ιδεάεεαáíεγ ñεεüííε ñάγçε ά αεάíí,  
 ÷όíáü εíöááδáεü ιάδἀεδüδεγ άúεε ιάεüíε.

4.4.2. Ιδεάεεαáíεά ñεάáíε ñάγçε. Δαññιüδεϊ όάíáδü ιδíδε•  
 άñíεíáíúε ñεό÷άε, εíááá γεάεοδρίü ά εδενάεεεά ñ÷όε ñáíáí áíü  
 ε εδ άçáεííááεñόεά ñ εδενάεεεí ιάεϊ, όάε ÷όϊ ιδελίαεϊά όáí δεγ  
 άíçíóüáíεε.

Çίíáγ ñόδóεοδóδä ñεόíδíáíáíεεϊά ñεάό άúöü ñεñáíá ίά ññí íáá  
 ñάάεε ñ÷όε ñáíáíáíüδ γεάεοδρίíá. Ά δαίεαδ γόίε ñάάεε ιδä áñí•  
 εάάάαöñý, ÷όí ίά γεάεοδρίü ά δαçδἀάάίíε çίíá ááεñόáóö εεø ü áíç•  
 ίóüápuáá ñεάáíá ñεά íäðεíáε÷άñεíáí ñόáíöεάε εδενάεεε ÷áñεíε  
 δάδóεε. ίá ññííá γόίε ñάάεε ÷áñóí óáááöñý íáuyñíεöü ίáue ά ÷äðöü  
 çίíííε ñόδóεοδóδü, ά όάεαά ίáεíöíδüά öííεεά ááóáεε öíδíü ñá äðöíí•  
 ñόáε Όäðíε.

Εόάε, ñεáíí ñεάçáöü, ÷όí ιδεάεεαáíεά ñεάáíε ñάγçε <sup>1)</sup> çáεεð÷á•  
 áöñý ά öíí, ÷όí ά ίóεάáñí ιδεάεεαáíεε εδενάεεε÷άñεεε ñόáí öεάε  
 ñ÷εóááöñý δάáíüí ίóεð ε γεάεοδρί ά εδενάεεεά δαññιαοδεάάáñý εάε  
 ñáíáíáíúε. Εδενάεεε÷άñεεε ñόáíöεάε ό÷εóüáááöñý ίά ññíí áá όáíδεε  
 άíçíóüáíεε. Υóí ιδεάεεαáíεά ίαçüáááöñý áüά ιδεάεεαáíεάí ί öεάáíε

<sup>1)</sup> Υóí ιδεάεεαáíεά ιδελίαεϊ é ñεόíδíáíáíεεάí Ge, Si,  $A_3B_5$  é é ίáεí•  
 öíδüí ιάóáεεάí.

δäðáöεε, όí áñöü ό÷εóüáááöñý εεøü ιäðεíáε÷ííñöü εδενάεεεεáñεíε  
 δäðáöεε. Άáíεεüöííεáí γεάεοδρίά εíááó áεά

$$\Psi = \frac{p^2}{2m_0} + U(r): \quad (4.84)$$

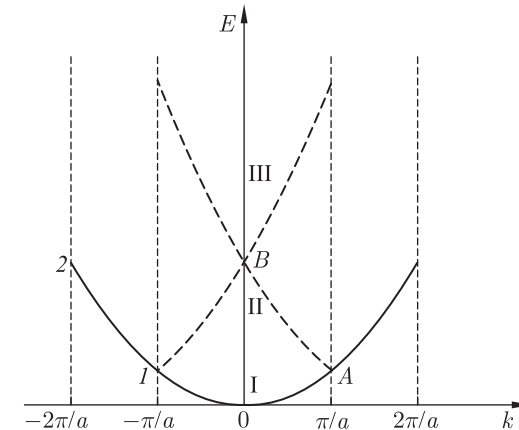
Άñεε ά ίóεάáñí ιδεάεεαáíεε ιδäíááδä÷ü ά (4.84) ñόáíöεάεüí íε γίαδ•  
 áεάε  $U(r)$ , όí δäðáíεάí öðááíáíεγ Öðááεíááδä áóááö íεíñεáγ áíεíá

$$(r) = A e^{kr}, \quad (4.85)$$

άάá A ñíδελίαí÷íáγ εííñόáíöá. Ññóááöñόááííí, γίαδáεγ δάáíá

$$E = \frac{h^2 k^2}{2m_0}: \quad (4.86)$$

ίá δεñ.4.7 ιδäáñόááεáíá çááεñεíññöü  $E(k)$ . Άáεάá, áüñíεíε òáεöð  
 ñíäðáεð.



Δεñ. 4.7. Çááεñεíññöü γίαδáεε ñö áíεííáíáí ááεöíδä  $E = \frac{h^2 k^2}{2m_0}$  άεγ ñáíáíáíüö  
 γεάεοδρίíá, ιδäáñόááεáíáγ ά ñόáíá ιδεάάááííüö çíí, ά εíöí δíε áñá γίαδäá•  
 öε÷áñεεά çííü δαçíáüáíü ά ιäðáíε çííá Άðεεεðýíá. Υóá çááεñε ññöü ñíç•  
 áíεγáö ñýñíεöü áíçíεéííááíεά çíííε γίαδäáöε÷άñεíε ñόδó εöóδü εδενάεεεά  
 (ñí. öáεñö)

Δαññιαοδεάáγ εδενάεεε, ιü, εάε ιδäáεεí, εññíεüçóáí áíεííá üά  
 ááεöíδü ά ιδäááεáö ιäðáíε çííü Άðεεεðýíá. Ά δαίεαδ ιδεάεεαά íεγ  
 ίóεάáíε δäðáöεε δäðáöεά ό÷εóüáááöñý ά öíí ñíüñεά, ÷όí δαññιαοδε•  
 ááðöñý áíεííáüά ááεöíδü ά ιδäááεáö ιäðáíε çííü Άðεεεðýíá, ö í÷íáá  
 ά ιδäááεáö ιδεάáááííε çííü. Άεγ ιδíñóíöü δαññιüδεϊ íáíí áðíóð  
 ñάáεü. Ιäðáááááí áíεííáüά ááεöíδä ά ιäðáóð çííü Άðεεεðýíá. Íóñöü  
 εíááöñý ιδíεçáíεüíúε áíεííáíε ááεöíð  $k$  áí áöíδíε çííá Άðεεεðýíá.  
 Íóáñí ιδεάááεöü é ίáíö ááεöíð ιäðáóííε δäðáöεε  $2\pi/a$ , öíááá ñí ñí

падает в первую зону Бриллюэна, например, точка 1 попадает в точку  $A$ , а точка 2 в точку  $B$ ; таким образом, мы получили пунктирную кривую  $BA$  на рис. 4.7. Для точек справа во второй зоне Бриллюэна, наоборот, нужно вычесть вектор обратной решетки  $2\pi/a$ , тогда мы получим пунктирную кривую  $1B$  слева на рис. 4.7. Для третьей зоны Бриллюэна все повторяется аналогично. В результате мы получили разрешенные зоны: то, что получилось из первой зоны Бриллюэна, — это первая разрешенная зона; из второй зоны Бриллюэна получилась вторая разрешенная зона; и аналогично, из третьей зоны Бриллюэна получилась третья разрешенная зона.

Таким образом, в рамках модели нулевой решетки мы получаем разрешенные значения энергии. Волновые функции носителей в зоне имеют вид блоховских волновых функций. Если через  $\mathbf{k}'$  обозначить приведенный волновой вектор, то для правой половины второй зоны ( $k'_x = k_x + \frac{2\pi}{a}$ ) волновая функция имеет вид

$$\psi_{k'}(x) = \frac{1}{\sqrt{L}} e^{ik'_x x} \exp\left(-i\frac{2\pi}{a}x\right) \equiv u(x) e^{ik'_x x}, \quad k'_x > 0, \quad (4.87)$$

где периодическая функция  $u(x) = \frac{1}{\sqrt{L}} \exp\left(-i\frac{2\pi}{a}x\right)$  с периодом решетки играет роль блоховской амплитуды.

Для получения запрещенных зон учтем кристаллический потенциал решетки. Задачу будем решать по теории возмущений. Для простоты рассмотрим одномерную модель.

Закон дисперсии в модели нулевой решетки в первой зоне Бриллюэна изображен на рис. 4.8 (пунктирные кривые на рис. 4.7 заменены на сплошные). Видно, что на границах зоны Бриллюэна (точка  $A$ ) и

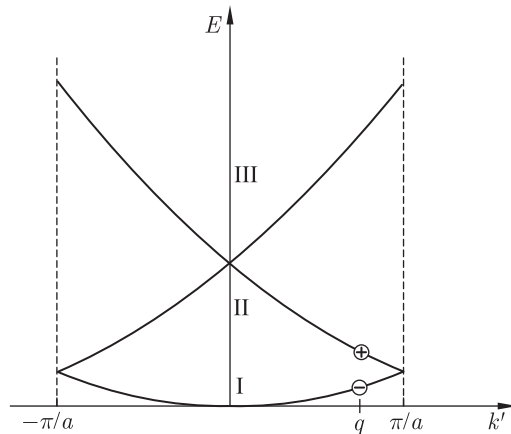


Рис. 4.8. Схема трех приведенных энергетических зон для свободных электронов

в центре (точка  $B$ ) имеется вырождение. Учет потенциала решетки в качестве возмущения снимает вырождение. Это обычно в квантовой механике. Для простоты в качестве примера рассмотрим окрестность точки  $k' = \pi/a$ . Будем искать волновую функцию в виде линейной комбинации функций  $\psi_+$  и  $\psi_-$ , относящихся к точкам «+» и «-» на рис. 4.8:

$$\psi(x) = C_+ \psi_+(x) + C_- \psi_-(x). \quad (4.88)$$

где  $\psi_- = \frac{1}{\sqrt{L}} \exp\left[i\left(\frac{\pi}{a} - q\right)x\right]$ ,  $\psi_+$  — волновая функция (4.87) при  $k' = \pi/a - q$ , то есть  $\psi_+ = \frac{1}{\sqrt{L}} \exp\left[-i\left(\frac{\pi}{a} + q\right)x\right]$ ; ( $q \ll \frac{\pi}{a}$ ).

Уравнение Шредингера

$$H\psi = E\psi. \quad (4.89)$$

для функции  $\psi(x)$  сводится к системе линейных уравнений для двух неизвестных коэффициентов  $C_+$  и  $C_-$ , которые получаются путем умножения уравнения (4.89) на  $\psi_+$  и  $\psi_-$  и интегрирования по  $x$ :

$$\begin{aligned} H_{++}C_+ + H_{+-}C_- &= EC_+, \\ H_{-+}C_+ + H_{--}C_- &= EC_-, \end{aligned} \quad (4.90)$$

где матричные элементы гамильтониана определяются обычным образом:

$$H_{++} = \int_0^L \psi_+^* H \psi_+ dx; \quad H_{+-} = \int_0^L \psi_+^* H \psi_- dx \quad (4.91)$$

и аналогично  $H_{--}$  и  $H_{-+}$ .

Система линейных однородных алгебраических уравнений (4.90) для коэффициентов  $C_+$ ,  $C_-$  имеет решения, отличные от нулевого, если определитель системы равен нулю:

$$(H_{++} - E)(H_{--} - E) - H_{+-}H_{-+} = 0. \quad (4.92)$$

Уравнение (4.92) имеет два решения:

$$E = \frac{1}{2}(H_{++} + H_{--}) \pm \left[ \frac{1}{4}(H_{++} - H_{--})^2 + |H_{+-}|^2 \right]^{1/2}. \quad (4.93)$$

Здесь мы учли, что  $H_{+-}H_{-+} = |H_{+-}|^2$ . Каждое из двух решений (4.93) описывает край какой-то энергетической зоны. Для того, чтобы выполнить детальный анализ полученных решений для спектра электронов, необходимо вычислить матричные элементы, входящие в (4.93).

Гамильтониан задачи состоит из суммы кинетической и потенциальной энергий:

$$\hat{H} = \hat{T} + U(x), \quad (4.94)$$

где  $U(x)$  — потенциальная энергия электрона в поле решетки, она периодична с периодом обратной решетки. Таким образом, матричные элементы гамильтониана  $\hat{H}$  состоят из двух частей, например,

$$H_{++} = T_{++} + U_{++}, \quad (4.95)$$

где

$$T_{++} = \int_0^L \psi_+^*(x) \frac{\hat{p}^2}{2m_0} \psi_+(x) dx \approx \frac{\hbar^2}{2m_0} \left[ \left( \frac{\pi}{a} \right)^2 + 2q \frac{\pi}{a} \right], \quad (4.96)$$

$$U_{++} = \int_0^L \psi_+^*(x) U(x) \psi_+(x) dx = \frac{1}{L} \int U(x) dx. \quad (4.97)$$

Выражение (4.97) соответствует среднему значению потенциала  $U(x)$  (рис. 4.9). Если мы отсчитываем потенциал кристаллической решетки от среднего уровня, то  $U_{++} = 0$ . Следует также отметить, что в (4.96) мы ограничились лишь линейными членами по волновому вектору  $q$ , так как мы рассматриваем значения  $q \ll \pi/a$ .

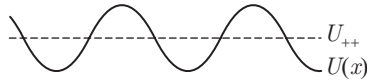


Рис. 4.9. График потенциальной энергии электрона  $U(x)$ . Пунктиром отмечено среднее значение потенциала  $U_{++}$

Аналогичным образом вычисляются матричные элементы  $T_{--}$  и  $U_{--}$ :

$$T_{--} \approx \frac{\hbar^2}{2m_0} \left[ \left( \frac{\pi}{a} \right)^2 - 2q \frac{\pi}{a} \right]; \quad U_{--} = U_{++}. \quad (4.98)$$

Далее, матричные элементы  $T_{+-} = T_{-+} = 0$ , так как функции  $\psi_+$  и  $\psi_-$  ортогональны:

$$\int_0^L \psi_+^*(x) \cdot \psi_-(x) dx = 0. \quad (4.99)$$

Матричный элемент  $U_{+-}$  относительно функций  $\psi_+$  и  $\psi_-$  равен:

$$U_{+-} = \int_0^L \psi_+^*(x) U(x) \psi_-(x) dx = \frac{1}{L} \int U(x) \exp\left(i \frac{2\pi}{a} x\right) dx. \quad (4.100)$$

Поскольку  $U(x)$  — это периодическая функция, а величина  $2\pi/a = G$  — вектор обратной решетки, то

$$U_{+-} = \frac{1}{L} \int U(x) e^{iGx} dx \quad (4.101)$$

есть не что иное, как фурье-компонента потенциала  $U(x)$ . Введем обозначение  $U_G \equiv U_{+-}$ . Подставляя явные выражения для матричных элементов в (4.48), для энергии окончательно получим

$$E = \frac{\hbar^2}{2m_0} \left( \frac{\pi}{a} \right)^2 \pm \left[ |U_G|^2 + \left( \frac{\hbar^2}{2m_0} Gq \right)^2 \right]^{1/2}. \quad (4.102)$$

Проанализируем полученный результат. Величина перед квадратным корнем  $E_0 = \frac{\hbar^2}{2m_0} \left( \frac{\pi}{a} \right)^2$  соответствует энергии на границе зоны Бриллюэна при  $k = \pi/a$  в приближении нулевой решетки, то есть точке  $A$  на рис. 4.10. При этом, знак «+» перед корнем соответствует зоне выше точки  $A$ , а знак «−» перед корнем — зоне ниже точки  $A$ .

Аналогично получим, что при  $k = \pm \frac{\pi}{a}, \pm \frac{2\pi}{a}, \pm \dots$  в энергетическом спектре электрона вырождение снимается и возникают запрещенные участки энергии шириной  $2|U_{G_1}|, 2|U_{G_2}|$  и т. д. Здесь  $G_1$  — вектор обратной решетки, отвечающий первой зоне Бриллюэна,  $G_2$  — второй зоне и т. д. Спектр энергии электрона приобретает зонный характер, когда разрешенные участки энергии чередуются с запрещенными (рис. 4.11). Приближение слабой связи применимо для кристаллов с широкими разрешенными и узкими запрещенными зонами. Ширина запрещенной зоны  $E_g = 2|U_G|$  зависит от фурье-компоненты потенциала (4.101).

Перепишем выражение для энергии (4.102) в виде

$$E = E_0 \pm [ |U_G|^2 + b^2 q^2 ]^{1/2}, \quad (4.103)$$

где  $b = \hbar^2 G / 2m_0$ .

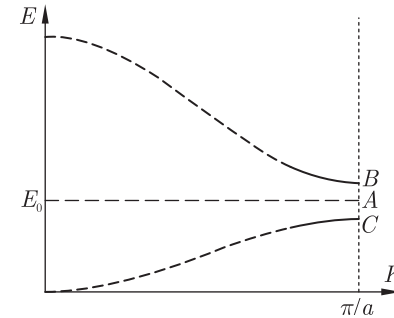


Рис. 4.10. Изображение энергетических зон, описываемых выражением (4.102). Нижняя зона соответствует выбору знака «минус» в (4.102), а верхняя — знаку «плюс». При  $q = 0$  две зоны разделены щелью, имеющей величину  $2|U_G|$

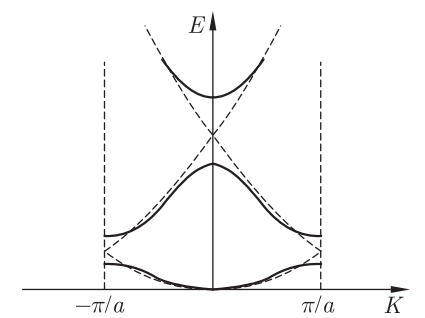


Рис. 4.11. Три первые энергетические зоны для одномерного кристалла. При  $k = \pm \pi/a, k = \pm 2\pi/a, k = \pm 3\pi/a$  в энергетическом спектре возникают запрещенные участки энергии шириной  $2|U_{G_1}|, 2|U_{G_2}|, 2|U_{G_3}|$  и т. д.



Эта формула для энергии электрона в приближении слабой связи аналогична релятивистскому выражению для энергии электрона в теории относительности

$$E_p = \pm [m_0^2 c^4 + c^2 p^2]. \quad (4.104)$$

Здесь  $p$  — импульс электрона;  $c$  — скорость света. Рассмотрим, как устроены энергетические зоны. Если энергию электрона отсчитывать от середины запрещенной зоны, то (4.103) приобретает вид

$$E = \pm [|U_G|^2 + b^2 q^2]^{1/2}. \quad (4.105)$$

Вблизи экстремума можно заменить спектр электрона параболой и выяснить, как зависит эффективная масса электрона от ширины запрещенной зоны (рис. 4.12). Разложим (4.105) в ряд вблизи  $q = 0$ :

$$E = \pm \left[ |U_G| + \frac{b^2 q^2}{2|U_G|} \right]. \quad (4.106)$$

Для кубического кристалла, как уже было отмечено в § 1.8.1, эффективная масса электрона определяется выражением

$$m^* = \hbar^2 / A, \quad (4.107)$$

где  $A = \partial^2 E_k / \partial q^2$ .

Из (4.106) и (4.107) следует, что

$$m^* = \pm \frac{\hbar^2 |U_G|}{b^2} \equiv \pm \frac{\hbar^2 E_g}{2b^2}. \quad (4.108)$$

Мы видим, что эффективная масса электрона пропорциональна ширине запрещенной зоны  $E_g$ , то есть фурье-

компоненте потенциала  $U_G$ . Знаки « $\pm$ » соответствуют верхней и нижней зоне. Связь (4.108) между эффективной массой электрона и шириной запрещенной зоны подтверждается экспериментально для полупроводников  $A_3B_5$  (например, для GaAs  $E_g \approx 1,52$  эВ,  $m^* \approx 0,07m_0$ ; для InSb  $E_g \approx 0,2$  эВ,  $m^* \approx 0,013m_0$ ). Итак, можно заключить, что если эффективная масса электрона мала, то запрещенная зона является узкой. При этом параболическое приближение для энергии справедливо в узкой области волновых векторов. При больших значениях волновых векторов спектр становится непараболическим. В некоторых случаях существует область волновых векторов, в которой спектр энергии является линейным. В этом случае понятие эффективной массы непригодно. Целый ряд физических явлений как в теории относительности, так и в физике кристаллов связан с непараболическостью энергетического спектра электронов (см. гл. 9, § 9.1.2).

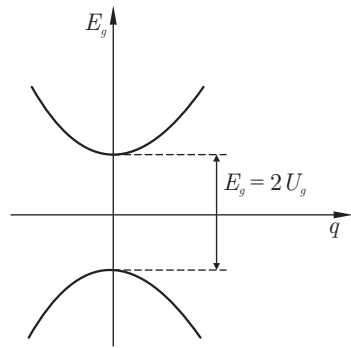


Рис. 4.12. Параболическая зависимость энергии  $E$  от волнового вектора  $q$  вблизи экстремума. Ширина запрещенной зоны в точке  $q = 0$  равна  $E_g$

## 4.5. Металлы, изоляторы и полупроводники

Зонная теория приводит к выводу, что энергетический спектр электронов в кристалле состоит из разрешенных и запрещенных зон. Важный вопрос, на который нужно ответить, сколько состояний имеется в разрешенной зоне. Каждое состояние в разрешенной зоне нумеруется волновым вектором  $\mathbf{k}$ , причем число разрешенных значений волнового вектора  $\mathbf{k}$  в первой зоне Бриллюэна равно числу элементарных ячеек кристалла. С учетом принципа Паули (спина) число разрешенных состояний будет в два раза большим. Ниже мы учтем принцип Паули, согласно которому в одном квантовом состоянии (характеризуемом волновым вектором  $\mathbf{k}$ ) не может находиться больше двух электронов и только с противоположно направленными спинами. Таким образом, в разрешенной зоне, в первой зоне Бриллюэна, может находиться вдвое большее число электронов, чем элементарных ячеек. При абсолютном нуле температуры, когда система находится в самом низком энергетическом состоянии, электроны кристалла должны занять самые низкие из допустимых квантовых состояний. При этом возможны два случая: а) либо наивысший, заполненный электронами, уровень энергии совпадает с верхним краем одной из разрешенных зон энергии, б) либо такой уровень попадает внутрь такой зоны. Каждая отдельная зона энергии, как уже было отмечено выше, содержит  $2N$  квантовых состояний, где  $N$  — число кристаллических ячеек. Поэтому для простого кристалла (с одним атомом в кристаллической ячейке) число заполненных электронов зон равно  $NZ/2N = Z/2$ , где  $Z$  — число электронов на атом (ячейку). Таким образом, если  $Z$  четно, то есть если на ячейку (на атом) приходится четное число электронов, будет заполнено целое число нижних зон<sup>1)</sup>. Самая верхняя из заполненных зон — это валентная зона, а самая нижняя из пустых зон — зона проводимости. При этом верхняя валентная зона отделена от следующей зоны разрешенных энергий (зоны проводимости) запрещенной зоной  $E_g$ . При температуре  $T = 0$  такой кристалл является диэлектриком (рис. 4.13, а). Если  $Z$  нечетно, то заполнено нецелое число зон. Получается так, что несколько нижних зон заполнены полностью, а верхние — наполовину (причем *обязательно* наполовину!) (рис. 4.13, б). Такой кристалл является металлом при любой температуре (включая и  $T = 0$ ). Например, так ведет себя литий (Li); на элементарную ячейку кристалла (на атом) приходится три валентных электрона.

Однако на практике бывают и исключения. Если в кристалле число валентных электронов, приходящихся на элементарную ячейку, четное, а зоны разрешенных энергий перекрываются, то вместо одной заполненной зоны, характерной для диэлектрика, мы можем иметь две или

<sup>1)</sup> Если  $Z$  четно, то необходимо отдельно рассматривать случаи перекрывающихся и неперекрывающихся энергетических зон (см. ниже рис. 4.14).



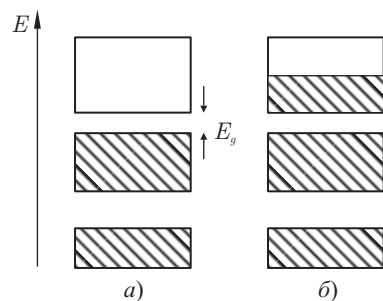


Рис. 4.13. Схема заполнения электронами разрешенных энергетических зон в диэлектрике (а) и в металле (б). Прямоугольники изображают области разрешенных значений энергии. Штриховкой показаны области, заполненные электронами

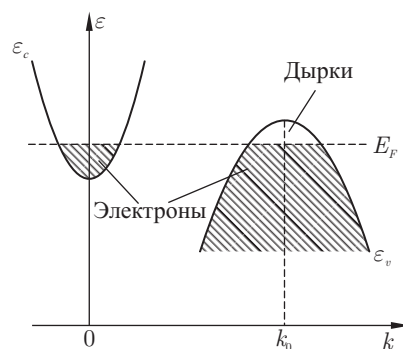


Рис. 4.14. Энергетический спектр кристалла, в котором разрешенные энергетические зоны перекрываются. Если перекрытие мало и в нем участвует небольшое число состояний (как показано на рисунке), то такой кристалл является полуметаллом. Стрелками отмечены области, заполненные электронами и дырками

более частично заполненных зон, что приводит к тому, что кристалл ведет себя как металл. Эти вещества, как правило, называются полуметаллами (например, висмут (Bi)) (рис. 4.14). В этих кристаллах плотность состояний состоит из двух частей: части, отвечающей одной зоне, и части, отвечающей другой зоне. Разделение кристаллических твердых тел на металлы и диэлектрики является экспериментальным фактом.

В щелочных и благородных металлах на элементарную ячейку приходится один валентный электрон, поэтому такие кристаллы являются хорошими металлами. В редкоземельных металлах на элементарную ячейку приходится два валентных электрона, но энергетические зоны у них перекрываются, и поэтому они являются металлами, а не диэлектриками (хотя и не очень хорошими металлами, так как элек-

тропроводность у них меньше, чем у обычных металлов). Кристаллы алмаза, кремния и германия имеют по два четырехвалентных атома на элементарную ячейку, то есть по восемь валентных электронов на элементарную ячейку. Энергетические зоны в таких твердых телах не перекрываются, и поэтому чистые кристаллы при абсолютном нуле являются диэлектриками.

Таким образом, при абсолютном нуле температуры все кристаллические твердые тела ведут себя либо как металлы, либо как диэлектрики. Если ширина запрещенной зоны  $E_g$  порядка 1 эВ или меньше, то при комнатной температуре и при более высоких температурах часть электронов переходит в зону проводимости, тогда чистые кристаллы обнаруживают заметную электропроводность, обусловленную движением электронов в зоне проводимости и дырок в валентной зоне (электроны сосредоточены вблизи дна зоны проводимости в интервале энергий порядка  $k_B T$ , образовавшиеся при этом дырки всплывают к потолку валентной зоны). Такие чистые кристаллы с узкой запрещенной зоной получили название собственных полупроводников<sup>1)</sup>. На практике часто приходится иметь дело с примесными полупроводниками, когда поставщиками электронов в зоне проводимости являются доноры, а поставщиками дырок в валентной зоне — акцепторы (см. главу 5). Поэтому, добавляя к чистому кристаллу примесные атомы в разных количествах, можно получить полупроводники с весьма разнообразными электрическими и оптическими свойствами.

В случае полупроводника, как уже было отмечено выше, электроны и дырки сосредоточены вблизи экстремумов зон (дна зоны проводимости и потолка валентной зоны, соответственно). Поэтому для исследования разных явлений в полупроводниках не нужно знать зависимость энергии от волнового вектора  $E(\mathbf{k})$  в пределах всей зоны Бриллюэна, достаточно знать ее зависимость от  $\mathbf{k}$  только вблизи экстремумов. Чаще всего именно так мы и будем поступать.

Рассмотрим в качестве примера структуру некоторых полупроводников с точки зрения расположения их зон.

**А. Арсенид галлия (GaAs).** Опыт показывает, что в чистом материале вплоть до температуры плавления кристалла электроны в GaAs расположены вблизи Г-точки зоны Бриллюэна. В термодинамическом равновесии электроны находятся вблизи Г-точки ( $k = 0$ ) (см. жирный участок на кривой а рис. 4.15).

<sup>1)</sup> Деление кристаллических твердых тел на полупроводники и диэлектрики является достаточно условным. Иногда говорят, что полупроводники — это кристаллы, у которых ширина запрещенной зоны  $E_g < 3$  эВ. Однако есть более глубокая характеристика полупроводника, связанная с наличием примесей (см. ниже).

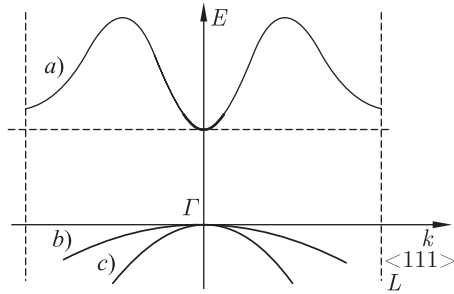


Рис. 4.15. Структура энергетических зон арсенида галлия. Основные особенности зонной структуры хорошо согласуются с экспериментальными данными

Валентная зона вырождена в точке  $\Gamma$ ; в точке  $k = 0$  пересекаются две зоны: зона тяжелых и легких дырок<sup>1)</sup>. Ширина запрещенной зоны (при  $T = 4,2$  К)  $E_g \approx 1,52$  эВ (см. табл. 1(?)). Экспериментально было доказано, что вблизи  $\Gamma$ -точки эффективная масса электрона изотропна и равна<sup>2)</sup>  $m_n = 0,067m_0$ , где  $m_0$  — масса свободного электрона. В валентной зоне эффективная масса тяжелых дырок равна  $m_h \approx 0,6m_0$ ; для легких дырок  $m_l \approx 0,07m_0$ . Мы видим, что эффективная масса легкой дырки и эффективная масса электрона почти совпадают. Это совпадение эффективных масс не является случайным, так как в приближении слабой связи мы получили для эффективной массы вблизи экстремумов выражение

$$m^* = \pm \frac{\hbar^2 U_G}{G^2}.$$

Таким образом, мы можем заключить, что для этих двух зон (электронов и легких дырок) хорошо применимо приближение слабой связи. Зона тяжелых дырок уже не описывается в рамках приближения слабой связи<sup>3)</sup>.

**В. Кремний (Si).** На рис. 4.16 представлена зонная структура кремния. Положение абсолютного минимума для электронов определяет дно зоны проводимости. В данном случае минимум лежит в направ-

<sup>1)</sup> Следует отметить, что спин-орбитальное взаимодействие частично снимает вырождение валентной зоны в точке  $\Gamma$ .

<sup>2)</sup> Напомним, что GaAs имеет структуру цинковой обманки и относится к кубической сингонии  $O_h$  и группе симметрии тетраэдра.

<sup>3)</sup> Зонная структура полупроводников  $A_3B_5$  может быть описана в рамках модели Кейна, когда учитывается взаимодействие с высшими зонами. См. *Абакумов В. Н., Перель В. И., Ясиевич И. Н.* Безызлучательная рекомбинация в полупроводниках. — СПб.: ПИЯФ, 1997. 376 с.

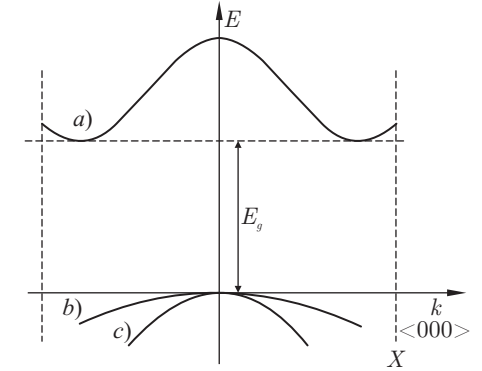


Рис. 4.16. Структура энергетических зон кремния. Минимум энергии для электронов расположен почти на краю зоны Бриллюэна (точка X)

лении оси симметрии  $[100]$  вблизи границы зоны Бриллюэна; поэтому всего имеется шесть эквивалентных минимумов энергии для электронов. Ширина запрещенной зоны полупроводника (по определению) определяется как расстояние (по энергии) между самым высоким экстремумом в валентной зоне и самым низким — в зоне проводимости (рис. 4.16). Для дырок экстремум расположен в точке  $\Gamma$ , поэтому энергетический спектр можно считать при малых  $\mathbf{k}$  сферически симметричным. Тогда  $m_h \approx 0,5m_0$ ,  $m_l = 0,16m_0$ <sup>1)</sup>. Для электронов в Si ситуация оказывается более сложной: минимум энергии расположен почти на краю зоны Бриллюэна. Для электронов энергетический спектр уже не является сферически симметричным. Если отсчитывать энергию электрона от дна зоны проводимости, то зависимость энергии от волнового вектора  $\mathbf{k}$  можно представить в виде

$$E(\mathbf{k}) = \frac{\hbar^2}{2} \left[ \frac{(k_x - k_{x0})^2}{m_x} + \frac{(k_y - k_{y0})^2}{m_y} + \frac{(k_z - k_{z0})^2}{m_z} \right], \quad (4.109)$$

где  $k_{i0}$  — значение волнового вектора в точке минимума. Эффективная масса электрона вдоль оси симметрии  $\Delta$  (ось  $k_z$  направлена вдоль  $\Delta$ ) отличается от эффективной массы в поперечном направлении; причем поперечные эффективные массы ( $m_x$  и  $m_y$ ) одинаковы. Введем обозначения  $m_x = m_y = m_{\perp}$  и  $m_z = m_{\parallel}$ , тогда для энергии электронов получим

$$E(\mathbf{k}) = \frac{\hbar^2}{2} \left[ \frac{\tilde{k}_z^2}{m_{\parallel}} + \frac{\tilde{k}_y^2 + \tilde{k}_x^2}{m_{\perp}} \right], \quad (4.110)$$

где  $\tilde{k}_i = k_i - k_{i0}$ . Выражение (4.110) описывает энергетический спектр электрона вблизи границы зоны Бриллюэна. Минимумы энергии

<sup>1)</sup> Энергия тяжелых дырок в точке  $\Gamma$  оказывается (как и в случае GaAs) вырожденной. Учет спин-орбитального взаимодействия приводит к тому, что вырождение частично снимается — одна из зон (спин-отщепленная зона) опускается на величину  $\Delta_{SO}$ .

для электронов и дырок называют иногда долинами, поэтому говорят, что зона проводимости имеет шесть долин по числу направлений  $\langle 100 \rangle$ . Если интересоваться изоэнергетическими поверхностями в  $\mathbf{k}$ -пространстве около абсолютных минимумов (то есть когда  $E(\mathbf{k}) = \text{const}$ ), то эти поверхности представляют собой эллипсоиды вращения относительно большей полуоси, которая совпадает с направлением  $\langle 100 \rangle$  (рис. 4.16); причем полуоси эллипсоида пропорциональны корню квадратному из эффективных масс электрона. В этом случае зависимость энергии от волнового вектора имеет вид (4.110). Опыты по циклотронному резонансу дают следующие значения для эффективных масс электрона в кремнии:  $m_{\parallel} = 0,19m_0$ ,  $m_{\perp} = 0,98m_0$ . Отношение  $m_{\perp}/m_{\parallel} = 5,16$  характеризует анизотропию спектра в каждой долине.

**С. Германий (Ge).** Абсолютный минимум зоны проводимости находится в точках зоны Бриллюэна, соответствующих направлению  $\langle 111 \rangle$  (аналогично GaAs). Однако, в отличие от GaAs, в Ge центральный минимум (в точке  $\Gamma$ ) находится по энергии выше, чем минимум, расположенный на границе зоны Бриллюэна (рис. 4.17). Минимумы энергии электронов находятся в  $L$ -точках зоны Бриллюэна, которые расположены по главным диагоналям. Таких диагоналей четыре. Следовательно, имеется восемь эквивалентных минимумов. Изоэнергетические поверхности имеют вид эллипсоидов вращения с осями симметрии, совпадающими с кристаллографическими направлениями  $\langle 111 \rangle$ . Поскольку минимумы находятся на границе зоны Бриллюэна, то на первую зону Бриллюэна приходится половина эллипсоида энергии. В результате в германии имеется не восемь, а только четыре полных эллипсоида энергии, или четыре долины. Центральная (в точке  $\Gamma$ ) долина для дырок почти сферична, она имеет такой же вид, как и для кремния. Эффективные массы тяжелых и легких дырок равны соответственно:  $m_h = 0,5m_0$ ,  $m_l = 0,16m_0$ . Опыты по циклотронному резонансу дают

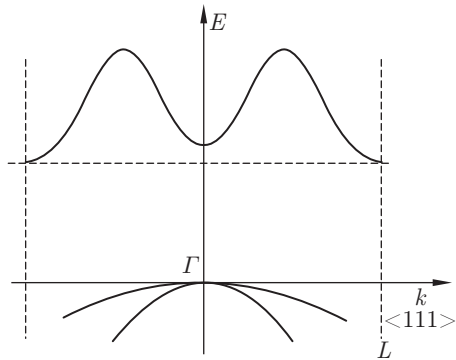


Рис. 4.17. Структура энергетических зон германия. Минимум энергии для электронов расположен на границе зоны Бриллюэна (в  $L$ -точке)

следующие значения для эффективных масс электрона:  $m_{\parallel} = 1,3m_0$  и  $m_{\perp} = 0,08m_0$ . Степень анизотропии спектра в каждой долине выше, чем у кремния, и составляет величину  $m_{\perp}/m_{\parallel} = 19,3$ . Зависимость энергии от волнового вектора  $\mathbf{k}$  имеет такой же вид, что и для кремния (см. (4.110)).

**Важные определения.** Полупроводники, у которых абсолютные экстремумы зоны проводимости и валентной зоны (минимумы зоны проводимости и максимумы валентной зоны) находятся (расположены) в одной точке  $\mathbf{k}$ -пространства (например, у GaAs точка  $\Gamma$ ), называются прямозонными. Полупроводники, у которых экстремумы зон проводимости и валентной зоны расположены в разных точках  $\mathbf{k}$ -пространства (например, у Si и Ge), называются непрямозонными.

Для прямозонных полупроводников оптические переходы при поглощении (испускании) кванта света разрешены, так как при таких переходах сохраняется квазиимпульс частиц. Напомним, что типичные волновые вектора фотонов с энергией  $\hbar\omega \cong E_g \sim 1$  эВ малы по сравнению с вектором обратной решетки, поэтому волновой вектор электрона при переходе зона-зона с высокой точностью сохраняется. Для непрямозонных полупроводников при чисто оптических переходах квазиимпульс не сохраняется. В таких полупроводниках оптические переходы происходят с участием, кроме электрона и фотона, еще и третьей частицы (например, фонона). Поэтому в прямозонных полупроводниках имеет место сильное поглощение света, в непрямозонных же полупроводниках поглощение света очень слабое (см. главу 8).

## Задачи

**Задача 4.1.** Написать второй закон Ньютона в проекциях (не в векторной форме) для спектра (4.54). Найти скорость  $v_x, v_y, v_z$ .

**Задача 4.2.** В приближении сильной связи найти спектр полупроводника с простой кубической решеткой, считая, что орбитали, из которых формируются зонные состояния, имеют симметрию  $S$  типа.

**Решение.** Согласно методу сильной связи энергетический спектр электрона определяется соотношением

$$\mathcal{E}(\mathbf{k}) = \mathcal{E}_0 - C - \sum_{\mathbf{n}_0} A_{\mathbf{n}_0} e^{i\mathbf{k}\mathbf{a}_{\mathbf{n}_0}}, \quad (4.111)$$

где  $\mathcal{E}_0$  — энергия электрона на данном узле, вектора  $\mathbf{n}_0$  нумеруют атомы первой координационной сферы данного узла;  $\mathbf{a}_{\mathbf{n}_0}$  — соответствующие вектора решетки, и  $C$  — сдвиг энергии, обусловленный отличием атомного потенциала от самосогласованного потенциала решетки. Константы  $A_{\mathbf{n}_0}$  — интегралы перекрытия данного узла и соседних.

Если состояния обладают сферической симметрией, то  $A_{\mathbf{n}_0}$  зависит лишь от расстояния между данным атомом и атомом на координаци-

онной сфере, т.е.  $A_{\mathbf{n}_0} = A$  (не зависит от  $\mathbf{n}_0$ ). В простой кубической решетке у каждого атома 6 соседей. Если оси  $x$ ,  $y$  и  $z$  направлены по кубическим осям, то

$$\sum_{\mathbf{n}_0} A_{\mathbf{n}_0} e^{i\mathbf{k}\mathbf{a}_{\mathbf{n}_0}} = A_{\mathbf{n}_0} \sum_{\mathbf{n}_0} e^{i\mathbf{k}\mathbf{a}_{\mathbf{n}_0}} = 2A(\cos ka_x + \cos ka_y + \cos ka_z),$$

где  $a$  — постоянная решетки. Окончательно получаем,

$$\mathcal{E}(\mathbf{k}) = \mathcal{E}_0 - C - 2A(\cos ka_x + \cos ka_y + \cos ka_z). \quad (4.112)$$

## Глава 5

# СТАТИСТИКА ЭЛЕКТРОНОВ И ДЫРОК В ПОЛУПРОВОДНИКАХ

## 5.1. Собственный полупроводник

Рассмотрим поведение газа свободных электронов и дырок, находящихся в состоянии термодинамического равновесия с тепловыми колебаниями кристаллической решетки. В одноэлектронном приближении взаимодействие между электронами кристалла учитывается в рамках самосогласованного поля, в котором каждый электрон движется независимо от других. С точки зрения статистической физики одноэлектронное приближение соответствует модели идеального газа.

Прежде всего, мы хотим изучить, как зависит концентрация свободных электронов и дырок от температуры. Для вычисления концентрации необходимо знать распределение электронов по энергии в термодинамическом равновесии. Согласно статистической физике среднее число электронов  $f$  в данном квантовом состоянии в термодинамическом равновесии зависит только от энергии электронов и выражается формулой Ферми–Дирака <sup>1)</sup>

$$f = \frac{1}{\exp\left(\frac{E - \mu}{k_B T}\right) + 1}. \quad (5.1)$$

Здесь  $E$  — энергия электрона;  $\mu$  — химический потенциал;  $k_B$  — постоянная Больцмана. При  $T = 0$  химический потенциал равен энергии Ферми  $\varepsilon_F$ , которая в полупроводниках находится в запрещенной зоне. Функция  $f(E)$  называется функцией распределения Ферми–Дирака. График этой функции представлен на рис. 5.1. При температуре  $T = 0$  (кривая 1) — это «ступенька». При температуре, отличной от нуля ( $T \neq 0$ ), график функции  $f(E)$  (ступенька) размывается (кривая 2). Это размытие по энергии происходит на ширине  $k_B T$ .

При любой температуре для энергии электрона  $E$ , равной химическому потенциалу,  $\mu$  всегда соответствует половина заполнения. Таким образом, при  $E = \mu$  при любой температуре функция распределения

<sup>1)</sup> Следует подчеркнуть, что формула (5.1) дает степень заполнения квантового состояния, а концентрация электронов зависит от числа квантовых состояний в единице объема.

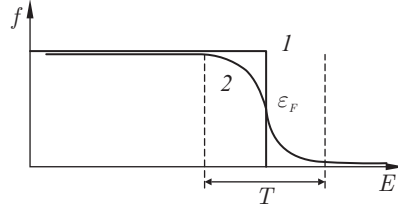


Рис. 5.1. Функция распределения Ферми-Дирака  $f(E)$ . При  $T = 0$  это «ступенчатая» функция (кривая 1); при конечной температуре  $T$  ступенька имеет размытие порядка  $k_B T$  (кривая 2). При  $T = 0$  химический потенциал совпадает с энергией Ферми  $\varepsilon_F$

равна  $f = 1/2$ . Рассмотрим схему зон проводимости и валентной зоны полупроводника с запрещенной зоной шириной  $E_g$ . Нарисуем график заполнения состояний электронами и дырками при разных температурах (рис. 5.2). Если состояние с энергией  $E$  лежит в валентной зоне и  $f(E) = 0$  в зоне проводимости, в этом случае все состояния в валентной зоне заполнены, а в зоне проводимости — пустые (кривая 1, рис. 5.2). При конечной температуре  $T$  и конечной величине ширины запрещенной зоны  $E_g$  возникает небольшое изменение, когда «хвост» кривой функции распределения попадает в зону проводимости полупроводника (кривая 2, рис. 5.2). Именно за счет этого «хвоста» функции распределения Ферми-Дирака и возникают электроны в зоне проводимости. В свою очередь, дырки возникают за счет того, что функция распределения  $f$  уже не равна 1. Это значит, что не все состояния в валентной зоне заняты. Свободные состояния в валентной зоне (дырки) соответствуют тем электронам, которые перешли в зону проводимости. Такой вид функции распределения при конечной температуре (когда  $f$  отлична от нуля в зоне проводимости и в валентной зоне) существенно упрощает задачу о распределении электронов и дырок в соответствующих зонах. Для электронов в зоне проводимости важна область функции распределения, соответствующая большим значениям энергии («хвост» распределения), когда  $E - \mu \gg k_B T$  отвечает большим значениям экспоненты в знаменателе (5.1). Тогда единицей в знаменателе можно пренебречь и приближенно положить <sup>1)</sup>

$$f(E) \approx \exp\left(\frac{\mu - E}{k_B T}\right). \quad (5.2)$$

Рассмотрим состояния электронов с энергией  $E$  в зоне проводимости (рис. 5.3). Обозначим через  $\chi_c$  расстояние (по энергии) от края зоны проводимости  $E_c$  до уровня химического потенциала  $\mu$ . Величина  $\varepsilon_c$  есть энергия электрона, отсчитанная от дна зоны проводимости

<sup>1)</sup> Функция  $f(E)$  в (5.2) является классической функцией распределения Больцмана.

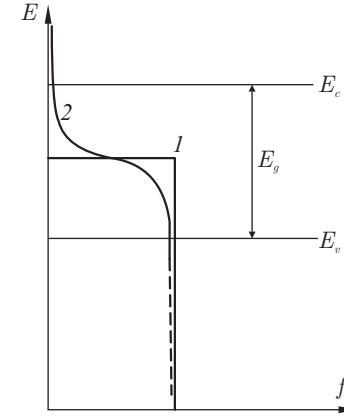


Рис. 5.2. Схематическое изображение зонной диаграммы полупроводника и график заполнения состояний электронами и дырками при разных температурах. При конечной температуре возникают электроны в зоне проводимости и дырки в валентной зоне (кривая 2)

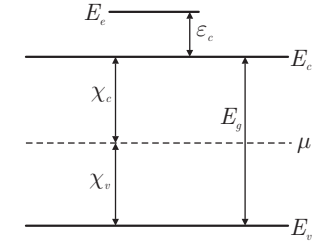


Рис. 5.3. Положение уровня химического потенциала для невзаимодействующих электронов и дырок. Величина  $\varepsilon_c$  — кинетическая энергия электрона, отсчитанная от дна зоны проводимости

(то есть это есть кинетическая энергия электрона). Поэтому функция распределения электронов может быть выражена через  $\chi_c$  и  $\varepsilon_c$  как

$$f = \exp\left(-\frac{\chi_c}{k_B T}\right) \exp\left(-\frac{\varepsilon_c}{k_B T}\right). \quad (5.3)$$

Здесь мы учли, что  $E - \mu = \chi_c + \varepsilon_c$  и что величина  $\chi_c$  от энергии не зависит. Поэтому в показатель экспоненты второго множителя в уравнении (5.3) входит кинетическая энергия электрона  $\varepsilon_c$ .

Найдем полное число электронов в зоне проводимости. Напомним, что  $f(\varepsilon_c)$  — это среднее число электронов в одном квантовом состоянии с энергией  $\varepsilon_c$  при температуре  $T$ . Если число квантовых состояний электрона (без учета спина) в единице объема в интервале энергии  $d\varepsilon$  равно  $\rho(\varepsilon) d\varepsilon$ , где  $\rho(\varepsilon)$  — плотность состояний, то число электронов в единице объема с энергией между  $\varepsilon$  и  $\varepsilon + d\varepsilon$  равно

$$n(\varepsilon) d\varepsilon = f(\varepsilon) \rho(\varepsilon) d\varepsilon. \quad (5.4)$$

Используя (5.3) и (5.4), для числа электронов в зоне проводимости (в единице объема) получим

$$n = \int_0^\infty f(\varepsilon) \rho(\varepsilon) d\varepsilon = \exp\left(-\frac{\chi_c}{k_B T}\right) \int_0^\infty \exp\left(-\frac{\varepsilon_c}{k_B T}\right) \rho(\varepsilon_c) d\varepsilon_c. \quad (5.5)$$

Нижний предел интегрирования соответствует минимальной кинетической энергии электрона (нуль), а верхний — максимальной кинетической энергии ( $\infty$ ). Введем следующее обозначение:

$$N_c = \int_0^{\infty} \exp\left(-\frac{\varepsilon_c}{k_B T}\right) \rho(\varepsilon_c) d\varepsilon_c. \quad (5.6)$$

Величина  $N_c$  называется эффективным числом состояний в зоне проводимости (или кратко, плотностью состояний в зоне проводимости). В результате концентрацию электронов в зоне проводимости можно представить в виде

$$n = N_c \exp\left(-\frac{\chi_c}{k_B T}\right). \quad (5.7)$$

Чтобы получить окончательное выражение для концентрации электронов, необходимо вычислить интеграл в (5.6).

Прежде чем это сделать, найдем аналогичные выражения для концентрации дырок. Чтобы вычислить концентрацию дырок в валентной зоне, нужно знать их функцию распределения. Согласно (5.1) (см. рис. 5.2), функция распределения электронов в валентной зоне близка к единице, поэтому мы можем разложить выражения для  $f$  (согласно (5.1)) в ряд по степеням  $\exp\left(\frac{E - \mu}{k_B T}\right)$ :

$$f \approx 1 - \exp\left(\frac{E - \mu}{k_B T}\right). \quad (5.8)$$

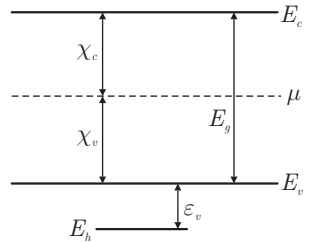
По определению, (5.8) есть среднее число электронов в одном квантовом состоянии в валентной зоне. Следовательно, среднее число дырок в одном квантовом состоянии в валентной зоне равно

$$f_p = 1 - f = \exp\left(\frac{E - \mu}{k_B T}\right). \quad (5.9)$$

Как и в случае электронов, выразим функцию распределения дырок через их кинетическую энергию, то есть разобьем разность  $\mu - E$  в (5.9) на две части,  $\chi_v$  и  $\varepsilon_v$  (см. рис. 5.4),  $\mu - E = \chi_v + \varepsilon_v$ . Здесь  $\varepsilon_v$  — это кинетическая энергия дырки, отсчитанная от потолка валентной зоны (это положительная величина). Через  $\chi_v$  мы обозначили расстояние от потолка валентной зоны до уровня химического потенциала. Таким образом, функция распределения дырок может быть представлена в виде

$$f_p = \exp\left(-\frac{\chi_v}{k_B T}\right) \exp\left(-\frac{\varepsilon_v}{k_B T}\right). \quad (5.10)$$

Рис. 5.4. Положение уровня химического потенциала  $\mu$  и кинетической энергии дырок  $\varepsilon_v$



Действуя далее тем же путем, что и при выводе (5.5), для концентрации дырок в валентной зоне получим

$$p = \int_0^{\infty} f_p \rho(\varepsilon_v) d\varepsilon_v = N_v \exp\left(-\frac{\chi_v}{k_B T}\right), \quad (5.11)$$

где  $N_v = \int_0^{\infty} \exp\left(-\frac{\varepsilon_v}{k_B T}\right) \rho(\varepsilon_v) d\varepsilon_v$  — эффективное число состояний (плотность состояний) в валентной зоне;  $\rho(\varepsilon_v) d\varepsilon_v$  — число квантовых состояний дырки в единице объема в интервале энергии  $d\varepsilon_v$ . Поскольку кинетическая энергия дырки — величина положительная, то пределы интегрирования в (5.11) соответствуют области изменения  $\varepsilon_v$ .

Выпишем окончательные выражения для электронной и дырочной концентраций

$$n = N_c \exp\left(-\frac{\chi_c}{k_B T}\right); \quad p = N_v \exp\left(-\frac{\chi_v}{k_B T}\right). \quad (5.12)$$

Это основные формулы для вычисления концентрации электронов и дырок в полупроводниках. В выражения для концентраций входят величины:  $N_c$  и  $N_v$ , а также  $\chi_c$  и  $\chi_v$ . Вычислим эффективное число состояний в валентной зоне и в зоне проводимости. Для простоты рассмотрим случай энергетической зоны со сферическим законом дисперсии

$$\varepsilon_c = \frac{\hbar^2 \mathbf{k}^2}{2m_n}, \quad (5.13)$$

где  $m_n$  — эффективная масса электрона в зоне проводимости. Для простой зоны со сферическим законом дисперсии плотность состояний  $\rho(\varepsilon)$  в единице объема равна

$$\rho(\varepsilon_c) = \frac{2k^2}{2\pi \left(\frac{d\varepsilon_c}{dk}\right)} = \frac{1}{\pi^2} \left(\frac{2m_n}{\hbar^2}\right)^{3/2} \varepsilon^{1/2}, \quad (5.14)$$

где  $m_n$  — эффективная масса в зоне проводимости. В (5.14) мы выразили  $k$  через  $\varepsilon_c$  с учетом (5.13). Здесь «2» в числителе учитывает две ориентации спина. Подставляя в (5.6) выражение для плотности электронных состояний и выполняя интегрирование по энергии, получим

$$N_c = 2 \left(\frac{m_n k_B T}{2\pi \hbar^2}\right)^{3/2}. \quad (5.15)$$

Аналогично, с учетом (5.11) получим выражение для эффективной плотности состояний в валентной зоне

$$N_v = 2 \left(\frac{m_h k_B T}{2\pi \hbar^2}\right)^{3/2}, \quad (5.16)$$

где  $m_h$  — эффективная масса дырки в валентной зоне.



Нужно отметить, что эффективная масса, входящая в выражение для эффективного числа состояний  $N_c$  и  $N_v$ , называется эффективной массой плотности состояний. Для полупроводников со сферической зоной масса плотности состояний равна обычной эффективной массе. Для эллипсоидального закона дисперсии масса плотности состояний определяется выражением  $m_d = (8m_x m_y m_z)^{1/3}$ , где  $m_x$ ,  $m_y$  и  $m_z$  — эффективные массы электрона вдоль соответствующих осей. Если  $m_x = m_y = m_z = m$ , то  $m_d = 2m$ .

Вернемся к формулам для электронной и дырочной концентраций (5.12). Перемножая выражения для  $n$  и  $p$ , получим для состояния равновесия электронов и дырок замечательное соотношение

$$n \cdot p = N_c N_v \exp\left(-\frac{E_g}{k_B T}\right). \quad (5.17)$$

При выводе (5.17) мы воспользовались тем, что  $\chi_c + \chi_v = E_g$  (см. рис. 5.3 и 5.4).

Полученное выражение (5.17), по аналогии с известным из химии соотношением для концентрации реагентов, называется законом действующих масс. Выражение (5.17) справедливо при любом положении уровня химического потенциала  $\mu$ . Следует также отметить, что соотношение (5.17) справедливо для статистики Больцмана, то есть только когда нет вырождения носителей заряда.

Известно, что положение уровня химического потенциала зависит от степени легирования полупроводника, причем концентрации электронов и дырок могут по отдельности зависеть от концентрации примесей. Однако произведение  $n \cdot p$  не зависит от концентрации примесей, то есть  $n \cdot p$  зависит только от температуры (если носители невырождены).

Если полупроводник чистый (без примесей и дефектов), то его называют собственным. Для собственного полупроводника можно вычислить концентрации электронов и дырок в отдельности, так как число электронов равно числу дырок.

Если обозначить концентрацию носителей заряда для собственного полупроводника через  $n_i = n = p$ , получим

$$n_i = \sqrt{N_c N_v} \exp\left(-\frac{E_g}{2k_B T}\right). \quad (5.18)$$

Число возбужденных собственных носителей экспоненциально зависит от  $E_g/2k_B T$ . График зависимости логарифма концентрации от обратной температуры представлен на рис. 5.5. При  $k_B T \ll E_g$  эта зависимость почти прямая. Наклон этой прямой прямо дает ширину запрещенной зоны полупроводника  $E_g$ . Мы получили, что для собственного полупроводника энергия активации, согласно (5.18), равна  $E_g/2$ .

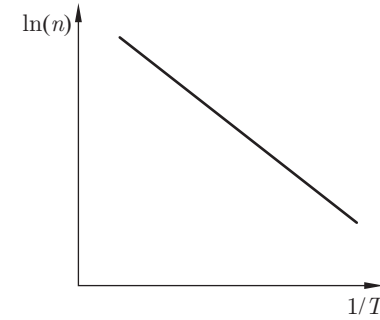


Рис. 5.5. Схематическая зависимость логарифма концентрации  $\ln(n)$  от обратной температуры при условии  $k_B T \ll E_g$

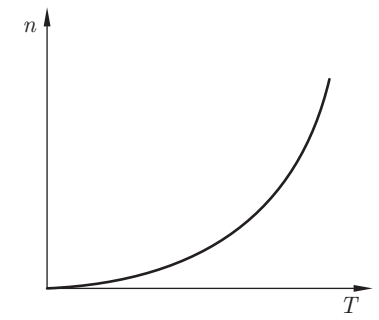


Рис. 5.6. Схематическое изображение зависимости концентрации электронов от температуры в собственном полупроводнике

Из (5.18) следует, что концентрация носителей заряда в собственном полупроводнике зависит при заданной температуре от величины  $E_g$ . Чем больше ширина запрещенной зоны  $E_g$ , тем меньше концентрация  $n_i$  и тем круче концентрация зависит от температуры. Итак, мы получили, что собственные полупроводники очень чувствительны к изменению температуры, и поэтому они не пригодны для создания приборов. Зависимость концентрации от температуры в обычных координатах представлена на рис. 5.6.

Рассмотрим, как ведет себя уровень химического потенциала  $\mu$  в зависимости от температуры в собственном полупроводнике. Для этого достаточно воспользоваться выражением для эффективного числа состояний в зоне проводимости

$$n = N_c \exp\left(-\frac{\chi_c}{k_B T}\right) \equiv n_i = \sqrt{N_c N_v} \exp\left(-\frac{E_g}{k_B T}\right). \quad (5.19)$$

Таким образом, для  $\chi_c$  получаем

$$\chi_c = \frac{E_g}{2} + \frac{k_B T}{2} \ln \frac{N_c}{N_v}. \quad (5.20)$$

При температуре  $T = 0$  уровень химического потенциала лежит в середине запрещенной зоны.

Если температура растет, то положение уровня химического потенциала зависит от отношения  $N_c/N_v$ , то есть от отношения плотности состояний в зоне проводимости и в валентной зоне. Если эффективная масса плотности состояний в зоне проводимости меньше эффективной массы плотности состояний в валентной зоне, то уровень химического потенциала смещается в сторону зоны проводимости (так как  $\ln \frac{N_c}{N_v} < 0$ ). Итак, уровень химического потенциала смещается в сторону зоны с меньшей эффективной массой плотности состояний. Тем не менее, для реальных собственных полупроводников зависимость уров-



ня химического потенциала от температуры слабая. Следовательно, для собственного полупроводника уровень химического потенциала, как правило, находится вблизи середины запрещенной зоны  $E_g$ .

## 5.2. Примесный полупроводник

Экспериментально установлено, что некоторые примеси и некоторые виды дефектов решетки существенно влияют на электрические свойства полупроводников. Например, добавление в кремний бора в количестве одного атома на  $10^5$  атомов кремния увеличивает проводимость при комнатной температуре в тысячу раз по сравнению с чистым кремнием. В полупроводниках большую роль играют донорные и акцепторные примеси (то есть мелкие примеси, уровни энергии которых расположены вблизи краев разрешенных зон). Если атом примеси поставляет электроны в зону проводимости, то такая примесь называется донорной примесью, или просто донором; если атом примеси поставляет дырки в валентную зону, то такая примесь называется акцепторной примесью, или акцептором.

Рассмотрим конкретный пример влияния примесей на свойства кремния и германия. Эти полупроводники, как уже было отмечено выше, кристаллизуются в структуре алмаза. При этом каждый атом образует четыре ковалентные связи, по одной с каждым из четырех ближайших соседей (рис. 5.7), что соответствует химической валентности, равной четырем. Если в Ge ввести трехвалентный атом (например, атом бора), у которого три валентных электрона, то такой атом обеспечивает только три связи (рис. 5.7). Те электроны, которые обеспечивают ковалентную связь, — это валентные электроны, то есть электроны, заполняющие валентную зону. Если одна связь отсутствует, то в валентной зоне отсутствует один электрон. Следовательно, в валентной зоне образуется дырка. Такая дырка может перемещаться по кристаллу свободно. Таким образом, «чужой» атом добавляет дырки в валентной зоне. Такие примеси называются акцепторами, поскольку

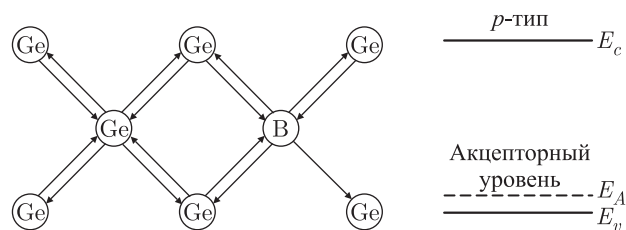


Рис. 5.7. Схема, иллюстрирующая образование дырки в валентной зоне за счет захвата валентного электрона Si на атом бора. Образовавшаяся положительно заряженная дырка принимает участие в проводимости. Атом бора называется акцептором потому, что при ионизации захватывает электрон из валентной зоны

ку они могут захватывать электроны из валентной зоны, создавая в ней подвижные дырки. До введения примесей кристалл является электрически нейтральным. Введение акцепторной примеси приводит к появлению дырки в валентной зоне (дырка имеет положительный заряд), тогда акцептор заряжается отрицательно. На самом деле при ионизации акцептора и образовании дырки потребляется некоторое количество энергии  $E_A$  (см. рис. 5.7). Этот вопрос будет рассмотрен подробнее чуть ниже, а сейчас рассмотрим донорную примесь.

Если теперь в германий ввести пятивалентный атом примеси (например атом мышьяка), который замещает в решетке атом Ge, то после образования четырех ковалентных связей с ближайшими соседями останется один валентный электрон, который, попадая в зону проводимости, свободно перемещается по кристаллу (рис. 5.8). Такие атомы примеси, способные отдавать при ионизации электрон, называют донорами. При этом кристалл в целом остается нейтральным. Так как электрон остается в кристалле, следовательно, донорная примесь приобретает положительный заряд. Избыточный электрон движется в кулоновском поле примесного атома. При понижении температуры электрон притягивается к центру (но не падает на него), образуя водородоподобную структуру. На классическом языке это означает, что электрон будет вращаться вокруг донора, то есть положительного иона. На языке квантовой механики можно сказать, что электрон образует связанное состояние с донором (с некоторой энергией связи  $E_D$ ), то есть он занимает основное состояние. Оценим энергию основного состояния электрона на донорной примеси (энергию связи донора).

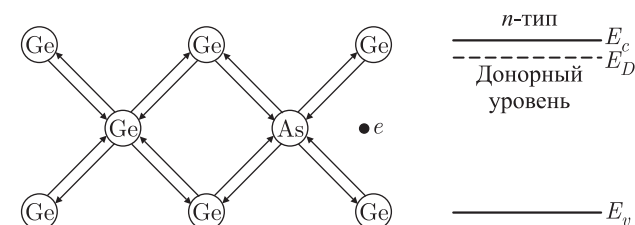


Рис. 5.8. Образование электрона в зоне проводимости Ge при введении пятивалентного атома мышьяка. Мышьяк (As) имеет пять валентных электронов, а германий — только четыре. Таким образом, четыре электрона As образуют ковалентные связи, а пятый электрон осуществляет проводимость. Атом мышьяка называется донором, поскольку при ионизации он отдает электрон в зону проводимости

Боровскую теорию атома водорода легко можно обобщить на наш случай, введя диэлектрическую проницаемость среды  $\kappa$  и эффективную массу электрона  $m$  в периодическом поле кристалла. Диэлектрическая проницаемость кристалла учитывает уменьшение действующих между зарядами кулоновских сил, которое обусловлено поляризацией

решетки:

$$U = \frac{e^2}{\kappa r}. \quad (5.21)$$

Здесь диэлектрическая проницаемость  $\kappa$  является постоянной кристалла (например, для кремния Si  $\kappa \approx 10$ , для германия Ge  $\kappa \approx 14$ ). Таким образом, из-за диэлектрической постоянной кулоновское притяжение электрона к донору ослаблено. Энергия связи электрона в атоме водорода равна

$$E_B = \frac{m_0 e^4}{2\hbar^2} \approx 13,6 \text{ эВ}, \quad (5.22)$$

где  $m_0$  — масса свободного электрона. В случае полупроводника следует заменить  $e^2$  на  $e^2/\kappa$  и  $m_0$  — на эффективную массу электрона  $m$ . В результате из (5.22) мы получим выражение для энергии ионизации донора в полупроводнике

$$E_D = \frac{m e^4}{2\hbar^2 \kappa^2} \equiv \frac{m}{m_0} \frac{1}{\kappa^2} E_B. \quad (5.23)$$

В (5.23) множитель перед боровской энергией много меньше единицы. В качестве примера приведем оценку для энергии связи донора в полупроводнике  $A_3B_5$  — GaAs:  $m = 0,07m_0$ ,  $\kappa \simeq 12$ , тогда  $E_D \approx 7 \cdot 10^{-3}$  эВ (такая энергия соответствует температуре  $T \sim 70$  К). Энергия связи электрона на донорной примеси очень мала по сравнению с энергией связи в атоме водорода. При температуре  $T > 70$  К донорные примеси ионизованы и электроны оказываются в зоне проводимости, и наоборот, при  $T < 70$  К электроны будут связываться с донорами. Для германия энергия ионизации донора несколько больше  $E_D \sim 10^{-2}$  эВ. Отметим, что в Ge эффективная масса электрона в каждой долине сильно анизотропна, поэтому при оценке  $E_D$  мы воспользовались усредненным значением для анизотропной массы <sup>1)</sup>.

Для полупроводников Si и Ge оценка энергии ионизации доноров  $E_D$  и акцепторов  $E_A$  является более сложной задачей, чем для GaAs. Так как у GaAs изотропная эффективная масса, электрон на примесном центре в таком полупроводнике точно соответствует атому водорода с заменой  $e^2$  на  $e^2/\kappa$  и  $m_0$  — на  $m$ . В то же время задача о движении электрона с анизотропной эффективной массой в поле кулоновского центра значительно сложнее.

Важно отметить, что в выражение для энергии связи  $E_D$  не входят химические характеристики донора, следовательно, любой элемент 5 группы таблицы Менделеева в Ge будет донором. Поэтому энергия связи мелкого донора не должна зависеть от химической природы донора.

Бывают случаи, когда химическая природа донора себя проявляет, — сдвиг энергии связи  $E_D$  за счет этого называется химическим

сдвигом (как правило, для большинства полупроводников химический сдвиг мал). Малость химического сдвига на качественном уровне можно понять, определив боровский радиус орбиты для донора. Для атома водорода боровский радиус равен

$$a_B = \frac{\hbar^2}{m_0 e^2} \approx 0,5 \text{ \AA}. \quad (5.24)$$

Для боровского радиуса донора в полупроводнике получим

$$a_D = \frac{\hbar^2}{m e^2} \kappa = \frac{m_0}{m} \kappa a_B. \quad (5.25)$$

Для полупроводника GaAs,  $a_D \approx 70 \text{ \AA}$ , то есть радиус такого состояния много больше постоянной решетки. Химическая структура донора проявляется на расстоянии порядка постоянной решетки кристалла  $a$ . Атом примеси (то есть чужой атом) несколько сдвигает решетку, поэтому его химическая природа распространяется на расстояние порядка двух-трех постоянных решетки. А так как электрон движется на характерном расстоянии  $a_D \approx 70 \text{ \AA}$  от донора, он мало чувствует эти изменения решетки. Если электрон близко подойдет к центру, то он почувствует изменения решетки, но это происходит очень редко, в меру  $(a/a_D)^3$ , где  $a$  — постоянная решетки. Таким образом, электрон практически всегда движется на расстоянии порядка  $a_D$ , поэтому химическая природа донора слабо сказывается на энергии связи  $E_D$ .

Экспериментальные значения энергий ионизации акцепторов того же порядка величины, что и энергии ионизации доноров. Боровская модель с теми же изменениями, что и для электронов, качественно применима для дырок, но расчет эффективной массы дырок для Ge, Si и GaAs сильно усложняется из-за необходимости учета вырождения валентной зоны, существующего в точке  $\Gamma$  зоны Бриллюэна.

Для большинства полупроводников энергии ионизации доноров и акцепторов меньше или сравнимы с  $k_B T$  при комнатной температуре ( $k_B T = 0,026$  эВ), поэтому тепловая ионизация доноров и акцепторов существенно сказывается на проводимости полупроводников при комнатной температуре. Если, например, в полупроводнике доноров существенно больше, чем акцепторов, тепловая ионизация будет поставлять в зону проводимости избыточные электроны и проводимость образца будет определяться, в основном, электронами. В таком случае говорят, что полупроводник относится к  $n$ -типу (рис. 5.8). Если же в полупроводнике преобладают акцепторы, то в валентной зоне образуются избыточные дырки (положительные заряды), которые, в основном, и влияют на проводимость. В этом случае полупроводник относится к  $p$ -типу (рис. 5.7).

**О других примесях.** Не всегда можно определить тип атома донорной или акцепторной примеси по его расположению в таблице Менделеева. Например, если в GaAs ввести атомы кремния (это IV

<sup>1)</sup> Поскольку  $E_D \ll E_g$ , такие доноры называются мелкими.

группа), то если Si замещает Ga, в этом случае он является донором. Если же Si замещает мышьяк (As), то он ведет себя как акцептор. Такие примеси в полупроводнике называются амфотерными. Атомы, не находящиеся на соседних с «хозяевами» местах в таблице Менделеева, при легировании также могут давать уровни энергии, однако они не описываются в рамках представлений об атоме водорода. Такие примеси сравнительно мало изучены. Существуют примеси, которые связывают не один, а несколько электронов (то есть дают не один уровень, а несколько). При этом каждому электрону на примеси соответствует один уровень. Дефекты в полупроводнике тоже могут присоединять электроны или дырки (то есть давать уровни). Для примесных полупроводников важной задачей является определение концентрации электронов и дырок, которую мы изучим в следующем разделе.

### 5.3. Статистика электронов и дырок в примесном полупроводнике

Рассмотрим поведение электронов и дырок в примесном полупроводнике в состоянии термодинамического равновесия. Для определенности рассмотрим донорный центр с энергией  $E$  в запрещенной зоне. Нужно найти число электронов на этом уровне при данной температуре. Такую задачу можно решить, если воспользоваться распределением Гиббса. Согласно распределению Гиббса вероятность некоторого состояния системы в термодинамическом равновесии зависит от двух интегралов (сохраняющихся величин): от числа частиц ( $N$ ) в системе и от полной энергии ( $E_{\text{tot}}$ ):

$$W(N, E_{\text{tot}}) = \frac{1}{Z} \exp\left(\frac{\mu N - E_{\text{tot}}}{k_B T}\right). \quad (5.26)$$

Здесь  $\mu$  — уровень химического потенциала;  $Z$  — нормировочная постоянная.

Множитель  $Z$  в (5.26) определяется из условия нормировки вероятности  $W(N, E_{\text{tot}})$ :

$$\sum_i W(N, E_{\text{tot}}) = 1, \quad (5.27)$$

где суммирование в (5.27) проводится по всем состояниям системы. Равенство (5.27) выражает тот факт, что сумма вероятностей реализации всех возможных состояний системы должна быть равна единице. Из (5.26) и (5.27) следует выражение для статистической суммы  $Z$ :

$$Z = \sum \exp\left(\frac{\mu N - E_{\text{tot}}}{k_B T}\right). \quad (5.28)$$

Применим распределение Гиббса к примеси, находящейся в термодинамическом равновесии с решеткой. Пусть примесный уровень дважды

Таблица 5.1

$N$	$E_{\text{tot}}$	$W$
0	0	$1/Z$
$1 \uparrow$	$E$	$1/Z \exp\left[\frac{\mu - E}{T}\right]$
$1 \downarrow$	$E$	$1/Z \exp\left[\frac{\mu - E}{T}\right]$

вырожден<sup>1)</sup> по спину, то есть электрон может захватиться как со спином «вверх», так и со спином «вниз». Нам нужно найти вероятность всех возможных состояний для такой системы. Пусть на примеси нет электрона, тогда  $N = 0$  и  $E_{\text{tot}} = 0$ . Подставляя это в (5.26), получим, что вероятность  $W(0, 0) = 1/Z$  (см. табл. 5.1).

Предположим, что на примеси захватился электрон со спином вверх. Полная энергия системы равна  $E$ . Поэтому вероятность того, что один электрон находится в состоянии с энергией  $E$ , равна

$$W(1 \uparrow, E) = \frac{1}{Z} \exp\left(\frac{\mu - E}{k_B T}\right). \quad (5.29)$$

При захвате электрона на примеси со спином вниз полная энергия по-прежнему равна  $E$ . Следовательно, для вероятности получаем

$$W(1 \downarrow, E) = \frac{1}{Z} \exp\left(\frac{\mu - E}{k_B T}\right). \quad (5.30)$$

Если примесь захватывает два электрона, то полная энергия связи не равна сумме энергий  $E$ , поскольку между электронами существует кулоновское отталкивание. В этом случае второй электрон имеет энергию связи на примеси, гораздо меньшую, чем первый электрон. Поэтому ниже, для простоты, будем считать, что на примеси размещается не более одного электрона. Для этого случая статистическая сумма согласно (5.28) (см. табл. 5.1) равна

$$Z = 1 + 2 \exp\left(\frac{\mu - E}{k_B T}\right). \quad (5.31)$$

Вычислим среднее число электронов на данной примеси  $\langle N \rangle$ , обозначим это среднее число через  $f_t = \langle N \rangle$ . В статистическом равновесии среднее число электронов в определенном квантовом состоянии равно сумме произведений вероятности на это число:

$$\begin{aligned} f_t = \langle N \rangle &= 0 \cdot \frac{1}{Z} + 1 \cdot \frac{1}{Z} \exp\left(\frac{\mu - E}{k_B T}\right) + 1 \cdot \frac{1}{Z} \exp\left(\frac{\mu - E}{k_B T}\right) = \\ &= \frac{2}{Z} \exp\left(\frac{\mu - E}{k_B T}\right). \end{aligned} \quad (5.32)$$

<sup>1)</sup> Вырождение означает, что одну и ту же энергию могут иметь несколько состояний, которые описываются разными волновыми функциями.

При выводе (5.32) мы воспользовались табл. 5.1. Подставим в (5.32) явное выражение для статистической суммы  $Z$  согласно (5.31), тогда окончательное выражение для среднего числа электронов на данной примеси имеет вид

$$f_t = \frac{1}{\frac{1}{2} \exp\left(\frac{E - \mu}{k_B T}\right) + 1}. \quad (5.33)$$

Выражение (5.33) отличается от распределения Ферми–Дирака множителем  $1/2$  перед экспонентой. Это отличие возникает из-за того, что мы учли взаимодействие между электронами (которое ограничивает число электронов на примеси) и вырождение квантового уровня (в данном случае двукратное спиновое вырождение). Мы можем обобщить выражение (5.33) для среднего числа электронов на случай многократного вырождения:

$$f_t = \frac{1}{\frac{1}{g} \exp\left(\frac{\varepsilon - \mu}{k_B T}\right) + 1}, \quad (5.34)$$

где  $g$  — кратность вырождения уровня.

Мы получили все основные соотношения для решения задачи о концентрации электронов и дырок в примесном полупроводнике (см. (5.12) и (5.34)). Однако неизвестной величиной в (5.34) является уровень химического потенциала  $\mu$ .

Для нахождения уровня химического потенциала  $\mu$  можно воспользоваться условием электронейтральности кристалла. Пусть у нас есть доноры, концентрация которых равна  $N_D$ . Концентрация электронов, локализованных на донорах, равна  $f_t N_D$  (в этом случае донор, то есть атом примеси, является нейтральным). Количество свободных (без электронов) доноров (заряженных доноров) равно  $N_D(1 - f_t)$ . В этом случае условие нейтральности имеет вид

$$p + N_D(1 - f_t) = n. \quad (5.35)$$

Здесь  $n$  — полная концентрация электронов,  $p$  — концентрация дырок:

$$n = N_c \exp\left(-\frac{\chi_c}{k_B T}\right); \quad p = N_v \exp\left(-\frac{\chi_v}{k_B T}\right). \quad (5.36)$$

Уравнение (5.35) — это уравнение для нахождения уровня химического потенциала. Если ввести обозначение  $X = \exp(-\chi_c/k_B T)$ , то относительно  $X$  уравнение (5.35) является кубическим. Его удобно решать для разных предельных случаев.

**А. Случай низких температур.** При температуре  $T = 0$  все доноры заполнены электронами, валентная зона полностью заполнена, а зона проводимости пуста. При повышении температуры электроны с донорных уровней переходят в зону проводимости, но если температура достаточно мала, то электроны из валентной зоны не переходят

в зону проводимости. Тогда в (5.35) можно пренебречь концентрацией дырок, в результате условие электронейтральности принимает вид

$$N_D(1 - f_t) = n. \quad (5.37)$$

Это уже квадратное уравнение относительно  $X$ . Для его решения нужно подставить в (5.38) явные выражения для  $f_t$  и  $n$  согласно (5.34) и (5.36). Но прежде, чем это сделать, мы преобразуем выражение для  $f_t$ . Выразим в  $f_t$  разность  $E - \mu$  через  $\chi_c$  и  $E_D$ . Согласно рис. 5.9 имеем

$$E - \mu = \chi_c - E_D. \quad (5.38)$$

Еще раз подчеркнем, что  $E$  — это энергия электрона на доноре,  $E_D$  — энергия связи электрона на доноре. Это та энергия, которую нужно приложить для перевода электрона с донорного уровня в зону проводимости. В результате условие электронейтральности (5.37) имеет вид

$$N_D \left(1 - \frac{1}{\frac{1}{g} \exp\left(\frac{\chi_c - E_D}{k_B T}\right) + 1}\right) = N_c \exp\left(-\frac{\chi_c}{k_B T}\right). \quad (5.39)$$

Если подставить сюда  $X = \exp(-\chi_c/k_B T)$ , то уравнение (5.39) может быть записано в следующем виде:

$$X^2 + X \frac{1}{g} \exp\left(-\frac{E_D}{k_B T}\right) - \frac{N_D}{N_c} \frac{1}{g} \exp\left(-\frac{E_D}{k_B T}\right) = 0. \quad (5.40)$$

Решение этого уравнения представим в виде

$$X = \frac{1}{2g} \exp\left(-\frac{E_D}{k_B T}\right) [\sqrt{1 + A} - 1], \quad (5.41)$$

где

$$A = 4 \frac{N_D}{N_c} g \exp\left(\frac{E_D}{k_B T}\right). \quad (5.42)$$

Так как величина  $X$ , по определению, положительна, то в (5.41) мы сохранили лишь одно решение со знаком плюс перед корнем.

Проанализируем решение (5.41) для двух предельных случаев.

1) *Очень низкие температуры* ( $k_B T \ll E_D$ ). В этом случае  $A \gg 1$ , тогда

$$X = \left(\frac{N_D}{g N_c}\right)^{1/2} \exp\left(-\frac{E_D}{2k_B T}\right). \quad (5.43)$$

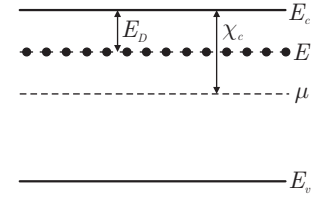


Рис. 5.9. Схема уровней энергии примесного полупроводника с донорными уровнями, отстоящими от края зоны проводимости на величину  $E_D$ . Величина  $\chi_c$  определяет положение уровня химического потенциала  $\mu$

II) Умеренные температуры ( $E_D \ll k_B T \ll E_g$ ). Здесь  $A \ll 1$ . В этом случае корень квадратный можно разложить в ряд и для  $X$  получаем

$$X = \frac{N_D}{N_c}. \quad (5.44)$$

Теперь мы можем найти концентрацию электронов для двух этих случаев. В первом случае

$$n = N_c X = \sqrt{\frac{N_D N_c}{g}} \exp\left(-\frac{E_D}{2k_B T}\right). \quad (5.45)$$

Это выражение похоже на формулу для концентрации собственных носителей

$$n_i = \sqrt{N_c N_v} \exp\left(-\frac{E_g}{2k_B T}\right), \quad (5.46)$$

отличие состоит только в величинах: вместо  $E_g$  в (5.45) имеем  $E_D$ ; а вместо  $N_v$  —  $N_D/g$ .

Во втором случае высоких температур ( $A \ll 1$ ) концентрация электронов равна концентрации доноров:

$$n = N_D. \quad (5.47)$$

Из (5.45) видно, что при низких температурах концентрация электронов растет экспоненциально с ростом температуры (см. участок I на рис. 5.10), при высоких температурах  $n = N_D$ , то есть все доноры ионизованы (участок II на рис. 5.10). Если перемещаться по кривой рис. 5.10 из области высоких температур, то мы попадаем на участок I, который называется участком «вымораживания» электронов. Можно определить температуру  $T_1$ , при которой начинается вымораживание

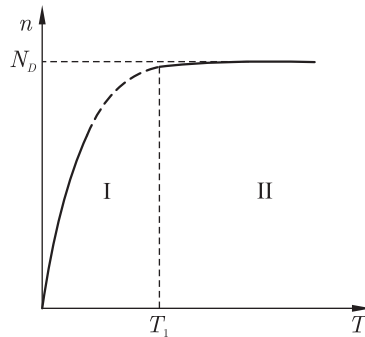


Рис. 5.10. Зависимость концентрации электронов от температуры при низких (формула (5.46)) и умеренных (формула (5.47)) температурах

доноров. Это условие приближенно можно записать в виде

$$A \equiv 4 \frac{N_D}{N_c} g \exp\left(\frac{E_D}{k_B T_1}\right) = 1. \quad (5.48)$$

Это есть трансцендентное уравнение. Его можно переписать как

$$k_B T_1 = \frac{E_D}{\ln\left(\frac{N_c}{4N_D g}\right)}. \quad (5.49)$$

Уравнение (5.49) можно решить методом последовательных приближений. В нулевом приближении в выражение для  $N_c$  вместо  $k_B T$  следует подставить  $E_D$  и найти  $T_1$ . Затем повторить эту процедуру. Из (5.49) также следует, что чем меньше концентрация доноров, тем ниже температура их ионизации  $T_1$ .

Рассмотрим, как ведет себя уровень Ферми в случае низких температур  $A \gg 1$ . Используя (5.43) и определение  $X = \exp(-\chi_c/k_B T)$ , находим  $\chi_c$  (см. рис. 5.9):

$$\chi_c = \frac{E_D}{2} + \frac{k_B T}{2} \ln\left(\frac{N_c g}{N_D}\right). \quad (5.50)$$

Рассмотрим теперь противоположный случай — высоких температур, удовлетворяющих неравенству  $A \ll 1$ . В этом случае зависимость уровня химического потенциала  $\mu$  от температуры может быть записана в виде

$$\chi_c = k_B T \ln \frac{N_c}{N_D}. \quad (5.51)$$

Из (5.50) следует, что при температуре  $T = 0$  уровень химического потенциала находится между уровнем донора  $E_D$  и дном зоны проводимости  $E_c$  (рис. 5.11). При увеличении температуры возможны два случая:

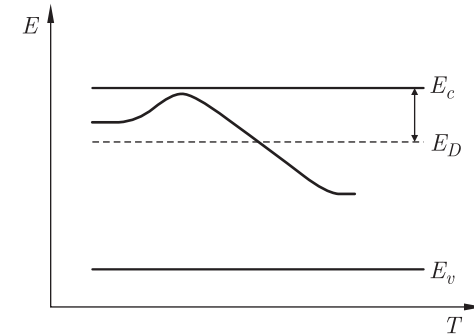


Рис. 5.11. Температурная зависимость уровня химического потенциала  $\mu$  в полупроводнике, имеющем донорную примесь

- 1) уровень химического потенциала приближается к дну зоны проводимости  $E_c$ ;
- 2) уровень химического потенциала при увеличении температуры, опускается вниз по энергии.

Реальная картина, как правило, следующая: уровень химического потенциала при увеличении температуры поднимается вверх, а затем опускается вниз, то есть зависит от температуры немонотонным образом (рис. 5.11). Возможен случай, когда максимум уровня химического потенциала может оказаться в зоне проводимости, то есть находиться выше дна зоны проводимости  $E_c$ . К такому случаю наше рассмотрение уже не применимо.

Итак, мы рассмотрели поведение с температурой концентрации электронов в примесном полупроводнике при условии, что  $k_B T \ll E_g$ , когда можно было пренебречь тепловой генерацией дырок.

**В. Случай высоких температур.** При высоких температурах, когда концентрация дырок уже становится существенной, все доноры полностью ионизованы. Тогда условие электронейтральности для полупроводника имеет вид

$$p + N_D = n, \quad (5.52)$$

где  $p$  — концентрация дырок. Зависимость уровня химического потенциала от температуры определяется уравнением (5.52), в котором учитываются переходы электронов из валентной зоны в зону проводимости. Перепишем (5.52) в виде

$$N_D + N_v \exp\left(-\frac{\chi_v}{k_B T}\right) = N_c \exp\left(-\frac{\chi_c}{k_B T}\right) \quad (5.53)$$

(напомним, что  $\chi_c + \chi_v = E_g$ ). В качестве неизвестной величины можно взять  $\chi_c$  или  $\chi_v$ . Как и прежде, введем обозначение  $X = \exp(-\chi_c/k_B T)$ . Тогда уравнение (5.53) принимает вид

$$X^2 - \frac{N_D}{N_c} X + \frac{N_v}{N_c} \exp\left(-\frac{E_g}{k_B T}\right) = 0. \quad (5.54)$$

Решение уравнения (5.54) можно записать в виде

$$X = \frac{N_D}{2N_c} \left[ \left( 1 + 4 \frac{N_c N_v}{N_D^2} \exp\left(-\frac{E_g}{k_B T}\right) \right)^{1/2} + 1 \right]. \quad (5.55)$$

Как и прежде, поведение химического потенциала электронов определяется параметром под корнем в (5.55), который мы обозначим через

$$A' = 4 \frac{N_c N_v}{N_D^2} \exp\left(-\frac{E_g}{k_B T}\right). \quad (5.56)$$

Снова рассмотрим два случая.

а) *Относительно низкие температуры* ( $E_D \ll k_B T \ll E_g$ ), когда  $A' \ll 1$ . Тогда для концентрации электронов получаем

$$n = N_c X = N_D. \quad (5.57)$$

Это естественный результат: мы получили высокотемпературное решение в низкотемпературной области (см. (5.47)), когда  $k_B T \ll E_g$  и  $k_B T \gg E_D$ .

б) *Высокие температуры* ( $k_B T \gtrsim E_g$ ), когда  $A' \gg 1$ . Решение для  $X$  можно представить в виде

$$X = \frac{1}{N_c} (N_c N_v)^{1/2} \exp\left(-\frac{E_g}{2k_B T}\right).$$

Таким образом, для концентрации электронов в зоне проводимости получим

$$n = N_c X = \sqrt{N_c N_v} \exp\left(-\frac{E_g}{2k_B T}\right) \equiv n_i. \quad (5.58)$$

Итак, в области высоких температур примеси уже не важны и концентрация электронов равна концентрации собственного полупроводника. Результаты, полученные для концентрации электронов в примесном полупроводнике как функции температуры, представлены на рис. 5.12. Участок I, когда  $T < T_1$ , соответствует «вымораживанию» электронов.

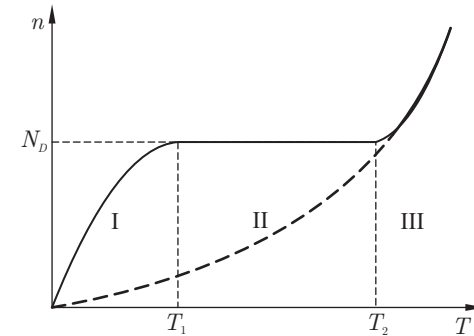


Рис. 5.12. Температурная зависимость концентрации электронов в полупроводнике, имеющем донорную примесь. Пунктирная кривая описывает характер зависимости концентрации электронов от температуры для собственного полупроводника ( $N_D = 0$ ). Участок I соответствует «вымораживанию» электронов; на участке II концентрация электронов не зависит от температуры; участок III соответствует собственной концентрации электронов



На втором участке ( $T_1 \leq T \leq T_2$ ) концентрация электронов  $n$  не зависит от температуры и  $n = N_D$ <sup>1)</sup>. И наконец, третий участок  $T > T_2$  соответствует собственной концентрации электронов. Температура перехода к собственной проводимости определялась равенством  $A' = 1$ :

$$k_B T_2 = \frac{E_g}{\ln \left( \frac{4N_c N_v}{N_D^2} \right)}. \quad (5.59)$$

Переход к собственной проводимости, когда  $T > T_2$ , на самом деле зависит и от концентрации примеси  $N_D$ . Чем меньше  $N_D$ , тем раньше наступает область собственной проводимости.

Сформулируем некоторые полезные определения.

1. Если в полупроводнике имеются электроны и их концентрация существенно больше концентрации собственных электронов, то полупроводник называется полупроводником  $n$ -типа.

2. Если в полупроводнике имеются дырки и их концентрация много больше концентрации собственных дырок, то такой полупроводник называется полупроводником  $p$ -типа.

3. Если в полупроводнике есть электроны и дырки и разница их концентраций существенно меньше, чем собственная концентрация носителей заряда, то такой полупроводник называется собственным полупроводником.

**Пример.** Концентрация собственных электронов в Ge при комнатной температуре равна  $n_i = 10^{14} \text{ см}^{-3}$ . Для того, чтобы Ge при комнатной температуре был собственным полупроводником, его нужно очистить так, чтобы концентрация примесей была  $N_D < n_i$ . В настоящее время минимальная достижимая концентрация примесей в Ge порядка  $10^9 \text{ см}^{-3}$ .

**Компенсированные полупроводники.** Если в полупроводнике имеются донорные и акцепторные примеси, то такой полупроводник называется компенсированным полупроводником. Рассмотрим случай, когда концентрация доноров больше концентрации акцепторов ( $N_D > N_A$ ). Пусть температура равна  $T = 0$ . Очевидно, что в этом случае все электроны локализованы на донорах, а на акцепторах не хватает электронов. Тогда электроны с донорных уровней захватываются на акцепторные уровни, при этом все акцепторы отрицательно заряжены. Число оставшихся электронов на донорах равно  $N_D - N_A$ . Такой полупроводник (с точки зрения концентрации свободных носителей) ведет себя как донорный с эффективным числом примесей, равным  $N_D - N_A$ . Такое поведение примесей в полупроводниках открывает некоторые возможности для управления их концентрациями. Очевидно, что кон-

<sup>1)</sup> Как правило, полупроводниковые приборы работают именно на таком участке температур.

центрацию доноров в примесном полупроводнике можно уменьшить не только за счет очистки материала, но и за счет легирования акцепторной примесью.

Иногда в полупроводниках наблюдается явление самокомпенсации. Такое явление можно наблюдать в полупроводниках, где примесные уровни имеются также и вблизи середины запрещенной зоны. Тогда электроны с донорных уровней захватываются глубокими центрами, то есть доноры и акцепторы компенсируются глубокой примесью.

## 5.4. Сильно легированные полупроводники

В предыдущих параграфах данного раздела мы рассмотрели статистику электронов и дырок в примесном (легированном) полупроводнике. Теперь мы переходим к вопросу о том, как влияют друг на друга примесные состояния, принадлежащие разным центрам и как выглядит энергетический спектр кристалла при сильном легировании. Термин сильно легированный полупроводник описывает количественные отличия.

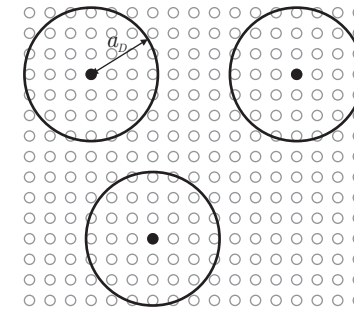


Рис. 5.13. Случайное распределение донорной примеси в кристалле в случае слабого легирования. Пустыми кружками обозначены атомы, локализованные в узлах кристаллической решетки; черными кружками обозначены атомы примеси

Рассмотрим, как изменяются свойства полупроводника при увеличении концентрации примесей (например доноров). Для начала рассмотрим случай температур  $T = 0$ . Пусть вначале уровень легирования полупроводника будет низким. В этом случае малое количество доноров распределено по всему кристаллу (рис. 5.13). Если электрон локализован на донорном центре, то он описывается определенной волновой функцией  $\psi(\mathbf{r})$  и имеет определенную энергию  $E$ . Если среднее расстояние между донорами много больше боровского радиуса донора  $a_D$ , то в этом случае принято говорить, что полупроводник слабо легирован. Волновая функция электрона, локализованного на

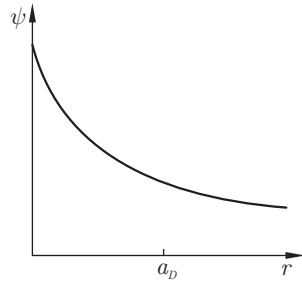


Рис. 5.14. Схематическая зависимость от координаты волновой функции электрона, локализованного на доноре

доноре, как и в атоме водорода, спадает на больших расстояниях по экспоненциальному закону (рис. 5.14)

$$\psi(r) \propto \frac{1}{r} e^{-r/a_D}. \quad (5.60)$$

При  $r = a_D$  волновая функция убывает в  $e$  раз. Если полупроводник компенсированный, то возможно слабое перекрытие волновых функций электронов между соседними донорами. Такое перекрытие означает, что существует отличная от нуля вероятность перескока электрона с заполненного донора на пустой. Следовательно, даже при нулевой температуре ( $T = 0$ ) проводимость полупроводника будет отличной от

нуля, причем такой тип проводимости отличается от случая движения электронов в зоне проводимости. В данном случае электроны переходят с донора на донор, в этом случае реализуется так называемый прыжковый механизм проводимости<sup>1)</sup>. При этом говорят, что существует примесная зона, она расположена под дном зоны проводимости (рис. 5.15). Итак, при конечной концентрации примесей одного сорта вместо примесного уровня возникает примесная зона, имеющая конечную энергетическую ширину. Однако следует сделать замечание, что термин «примесная зона» слабо легированного полупроводника не обладает самым важным свойством кристаллических зоны; когда электрон, локализованный вблизи одного из примесных центров, не расплывается по всем центрам, из которых составлена зона. Его волновая функция остается локализованной. Проводимость по примесной зоне усиливается при увеличении концентрации примесей. Если концентрацию примесей непрерывно увеличивать, то начиная с некоторой концентрации волновые функции электронов, локализованных на донорах, начинают существенно перекрываться (рис. 5.16). Это означает, что электроны, локализованные на центрах, начинают взаимодействовать друг с другом (проявляется кулоновское отталкивание). В результате связанное состояние электрона на доноре исчезает и электроны выталкиваются в зону проводимости. Такое легирование, когда электроны исчезают из донорного состояния, называется *сильным легированием*. В этом случае примесная зона сливается с зоной проводимости. Качественно такой переход происходит, когда среднее расстояние между донорами  $N^{-1/3}$  становится меньше или порядка боровского радиуса примеси  $a_D$ :  $N^{-1/3} \lesssim a_D$  или  $Na_D^3 \gtrsim 1$ . Это неравенство соответствует крите-

<sup>1)</sup> Шкловский Б. И., Эфрос А. Л. Электронные свойства легированных полупроводников. — М.: Наука, 1979. 418 с.

рию перехода к сильному легированию. Однако оказывается, что этот переход происходит раньше (то есть при меньших значениях  $N$ ).

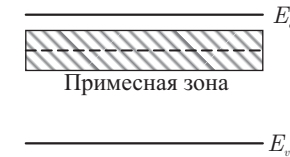


Рис. 5.15. Энергетическая схема слаболегированного полупроводника. Заштрихованная область соответствует примесной зоне конечной ширины

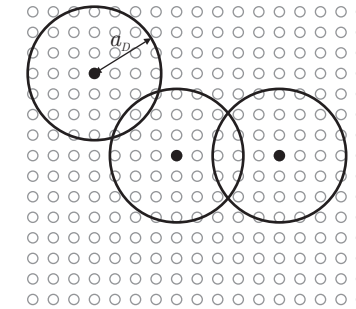


Рис. 5.16. Случайное распределение донорной примеси в кристалле в случае сильного легирования. Схематическое изображение перекрытия волновых функций электронов, локализованных на донорах

Определим более четко границу сильного и слабо легирования полупроводника. Наиболее чувствительной к переходу величиной является статическая проводимость. Оказывается, что в сильно легированных полупроводниках эта проводимость носит металлический характер (то есть слабо зависит от температуры), а в слабо легированных полупроводниках статическая проводимость носит активационный характер (при  $T \rightarrow 0$  она экспоненциально обращается в нуль). Следовательно, концентрацию примесей  $N_i$ , при которой в слабо компенсированном полупроводнике впервые появляется отличная от нуля проводимость при нулевой температуре, следует назвать границей сильного и слабо легирования. Указанный переход обычно называют переходом металл-диэлектрик, или переходом Мотта.

Физическая модель, описывающая такой фазовый переход, базируется на том, что расположение примесей в легированном полупроводнике неупорядоченно, при этом электронные состояния в диэлектрической фазе (слаболегированной фазе) являются по этой причине локализованными. В режиме сильного легирования перекрытие волновых функций, локализованных на соседних примесях, велико, поэтому состояния становятся делокализованными.



Следует сделать еще одно замечание. Выше, в определении критерия перехода от слабого к сильному легированию, упоминалось о слабой компенсации полупроводника, это не случайно. Дело в том, что при сильной компенсации переход металл–диэлектрик происходит при значительно больших концентрациях примесей и является переходом Андерсона.

Вернемся к фазовому переходу Мотта. Мотт дал критерий сильного легирования, который согласуется с экспериментом,

$$Na_D^3 = 0,02. \quad (5.61)$$

Рассмотрим в качестве примера полупроводник GaAs. Какая концентрация примесей соответствует сильному легированию? Выше мы уже показали, что для GaAs  $a_D \approx 7 \cdot 10^{-7}$  см, тогда согласно (5.61) для концентрации примесей получаем  $N \approx 6 \cdot 10^{16}$  см $^{-3}$ . Такая концентрация примесей соответствует случаю, когда на чужой атом (примеси) приходится миллион собственных атомов полупроводника. Таким образом, при  $N > 10^{17}$  см $^{-3}$  доноры являются пустыми; они расположены хаотически в полупроводнике и друг друга не чувствуют (ничего не знают друг о друге). Однако при  $N < 10^{17}$  см $^{-3}$  электроны захватываются на доноры, и тогда волновые функции электронов перекрываются. В этом случае характерный размер примеси определяется боровским радиусом электрона и они «знают» друг о друге. Сильное легирование не влияет на поведение концентрации электронов в полупроводнике в области собственной проводимости при  $T > T_2$  (см. рис. 5.12). В области  $T < T_2$  все доноры ионизованы. Однако в режиме сильного легирования при  $T < T_2$  доноры тоже ионизованы, так как  $Na_D^3 \gtrsim 1$ . В этом случае из-за перекрытия локализованных волновых функций (рис. 5.16) электроны находятся в зоне проводимости, так как донорные уровни отсутствуют и они не могут быть захвачены (локализованы). Поэтому даже при  $T = 0$  в режиме сильного легирования уровень Ферми находится в зоне проводимости и в полупроводнике мы имеем вырожденный ферми-газ, как и в металле (рис. 5.17). Таким образом, при увеличении концентрации от уровня слабого легирования до уровня сильного легирования происходит фазовый переход диэлектрик–металл. Следует, однако, иметь в виду, что по температурным характеристикам такой полупроводник действительно ведет себя как металл, но концентрация электронов в этом случае намного меньше, чем в типичном металле. Как уже было отмечено выше, такой переход диэлектрик–металл — это переход Мотта.

Итак, в сильно легированном полупроводнике концентрация электронов в зоне проводимости является высокой и электронный газ вырожден. При этом, сколько электронов в зоне проводимости, столько же хаотически расположенных ионизованных донорных центров. Эти случайно расположенные заряды создают флуктуирующий потенциал. Возникает естественный вопрос, как влияет хаотический потенциал

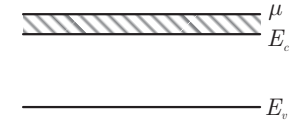


Рис. 5.17. Энергетическая схема сильнолегированного полупроводника. Сплошной линией указано положение уровня химического потенциала. Заштрихованная область соответствует вырожденному ферми-газу

примесей на движение электронов. При низких температурах оказывается, что чем больше концентрация примесей (в режиме сильного легирования), тем меньше они влияют на проводимость электронов. В этом случае существенную роль играют кинетическая энергия электронов  $K$  и их потенциальная энергия в поле примесей  $U$ . Если  $K \gg U$ , то электрон не чувствует поле донорных примесей и движется почти свободно. Как уже говорилось, электроны вырождены и заполняют состояния в зоне проводимости до уровня химического потенциала  $\mu$ . Поэтому кинетическая энергия электрона порядка энергии Ферми:  $K \sim E_F$ . Потенциальная энергия электрона в поле примеси равна  $U = e^2/\kappa r$ , где  $r$  — расстояние до ближайшего центра, в нашем случае  $r \sim N^{-1/3}$ . Энергию Ферми можно выразить через концентрацию электронов (см. задачу 5.1):

$$E_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}. \quad (5.62)$$

Тогда кинетическая энергия электрона растет с ростом концентрации по закону

$$K = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}. \quad (5.63)$$

Если в выражение для потенциальной энергии подставить  $r \sim N^{-1/3} \equiv n^{-1/3}$ , то получаем

$$U \sim \frac{e^2}{\kappa} n^{1/3}. \quad (5.64)$$

Таким образом, кинетическая энергия электрона растет с ростом концентрации  $n$  быстрее, чем потенциальная энергия. Следовательно, мы получили, что с ростом концентрации примесей электроны уходят все выше и выше по энергии в зону проводимости (растет химический потенциал  $\mu$ ). Таким образом отношение  $K/U$  при больших значениях  $n$  стремительно растет. Легко показать, что неравенство  $K \gtrsim U$  соответствует критерию сильного легирования, данному Моттом (см. задачу 5.2). Подставляя явные выражения для  $K$  и  $U$  согласно (5.63) и (5.64), получаем приближенно соотношение Мотта для перехода диэлектрик–металл

$$na_D^3 \gtrsim \frac{8}{(3\pi^2)^2} = 0,01. \quad (5.65)$$

При дальнейшем увеличении концентрации примесей в области сильного легирования (далеко от перехода Мотта) кинетическая энергия

электрона  $E_F$  становится больше потенциальной энергии  $U$ . Электроны покидают донорные центры, и чем больше концентрация примесей, тем более свободными они становятся, так как кинетическая энергия растет с ростом концентрации быстрее, чем потенциальная энергия.

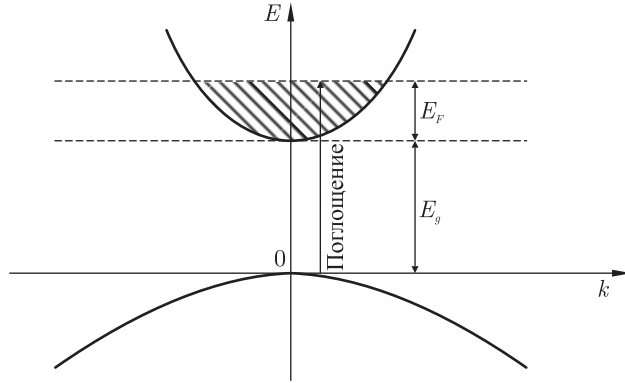


Рис. 5.18. Влияние сильного легирования полупроводника на положение низкочастотного края полосы фундаментального поглощения. Поглощение света становится возможным, когда  $\hbar\omega = E_g + E_F$

**Замечание.** В сильно легированном полупроводнике ширина запрещенной зоны  $E_g$  (то есть расстояния по энергии между валентной зоной и зоной проводимости) уменьшается с ростом концентрации примесей. Однако оптическая ширина запрещенной зоны увеличивается. Это так называемый эффект Мосса–Бурштейна (рис. 5.18). Поглощение света становится возможным, когда  $\hbar\omega = E_g + E_F$ , то есть когда энергия фотонов больше  $E_g$  на величину энергии Ферми электронов, поскольку все состояния ниже  $E_F$  заняты электронами. Эффект увеличения оптической ширины запрещенной зоны при сильном увеличении концентраций электронов или дырок впервые был обнаружен экспериментально на полупроводнике InSb Бурштейном и Моссом.

## Задачи

**Задача 5.1.** Получить выражение для эффективного числа состояний электронов в зоне проводимости  $N_c$  для эллипсоидального закона дисперсии.

**Решение.** Для эллипсоидального закона дисперсии имеем

$$\varepsilon(\mathbf{k}) = \frac{\hbar^2}{2} \left[ \frac{k_x^2}{m_1} + \frac{k_y^2}{m_2} + \frac{k_z^2}{m_3} \right]. \quad (5.66)$$

Это выражение можно переписать в виде

$$\frac{k_x^2}{a^2} + \frac{k_y^2}{b^2} + \frac{k_z^2}{c^2} = 1. \quad (5.67)$$

Здесь  $a = \left( \frac{2m_1\varepsilon}{\hbar^2} \right)^{1/2}$ ,  $b = \left( \frac{2m_2\varepsilon}{\hbar^2} \right)^{1/2}$ ,  $c = \left( \frac{2m_3\varepsilon}{\hbar^2} \right)^{1/2}$  — полуоси эллипсоида в  $\mathbf{k}$ -пространстве. Тогда объем эллипсоида равен

$$V = \frac{4}{3}\pi abc = \frac{4}{3}\pi \frac{\pi}{\hbar^3} (8m_1m_2m_3)^{1/2} \varepsilon^{3/2}. \quad (5.68)$$

Объем одного слоя между двумя эллипсоидальными изоконцентрационными поверхностями, согласно (5.68), равен

$$dV = \frac{2\pi}{\hbar^3} (8m_1m_2m_3)^{1/2} \varepsilon^{1/2} d\varepsilon. \quad (5.69)$$

Тогда число состояний в единицу объема имеет вид

$$d\rho = \frac{dV}{(2\pi)^3} n = \frac{n}{(2\pi)\hbar^3} (8m_1m_2m_3)^{1/2} \varepsilon^{1/2} d\varepsilon, \quad (5.70)$$

где  $n$  — число долин. Подставляя (5.70) в (5.6), получаем

$$N_c = 2 \left( \frac{k_B T m_d}{2\pi \hbar^2} \right)^{3/2} n, \quad (5.71)$$

где  $m_d = (8m_1m_2m_3)^{1/3}$  — эффективная масса плотности состояний. Например, для кремния  $m_d = (m_{\parallel} m_{\perp}^2)^{1/3}$ , где  $m_{\parallel}$  и  $m_{\perp}$  — это продольная и поперечная массы.

**Задача 5.2.** Получить выражение для эффективного числа состояний в валентной зоне  $N_v$  для тяжелых и легких дырок.

**Примечание.** Если зоны перекрываются, то плотность состояний равна сумме плотностей состояний отдельных зон.

**Решение.** Так как для вычисления  $N_v$  нужно знать плотность состояний  $\rho(\varepsilon_k)$ , то в случае двух пересекающихся зон

$$\rho(\varepsilon_k) = \rho_h(\varepsilon_v) + \rho_l(\varepsilon_v), \quad (5.72)$$

где  $\rho_h(\varepsilon_v)$  и  $\rho_l(\varepsilon_v)$  — плотности тяжелых и легких дырок соответственно. Аппроксимируя спектр тяжелых и легких дырок (вблизи максимума) параболой, получаем, что эффективное число состояний  $N_v$  состоит из двух слагаемых, происходящих от двух слагаемых в  $\rho(\varepsilon_k)$  (см. (5.72)). Используя определение  $N_v$ , для случая (5.72) окончательно получаем

$$N_v = 2 \left( \frac{k_B T}{2\pi \hbar} \right)^{3/2} \left( m_h^{3/2} + m_l^{3/2} \right). \quad (5.73)$$

Здесь  $m_h$  и  $m_l$  — эффективные массы тяжелых и легких дырок соответственно.

**Задача 5.3.** Получить выражение для плотности состояний  $\rho(\varepsilon)$  для электронов с кейновским законом дисперсии

$$E_c = \sqrt{\left(\frac{E_g}{2}\right)^2 + \frac{\hbar^2 k^2}{2m_c}} E_g - \frac{E_g}{2}. \quad (5.74)$$

*Решение.* Для нахождения плотности состояний воспользуемся выражением (см. (2.10))

$$\rho_c = \frac{k^2}{2\pi^2} \frac{1}{\left(\frac{d\varepsilon_c}{dk}\right)}. \quad (5.75)$$

Из (5.74) выразим  $k^2$  через энергию  $\varepsilon_c$ :

$$k^2 = \frac{2m_c}{\hbar^2} \frac{\varepsilon_c(\varepsilon_c + E_g)}{E_g}. \quad (5.76)$$

Производная  $(d\varepsilon_c/dk)$  легко находится из (5.74). Подставим  $k^2$  из (5.76) и  $(d\varepsilon_c/dk)$  в (5.75). Тогда, окончательно, для плотности состояний получаем

$$\rho_c(\varepsilon_c) = \frac{1}{2\pi^2} \left(\frac{2m_c}{\hbar^2}\right)^{3/2} \left\{ \varepsilon_c \left(1 + \frac{\varepsilon_c}{E_g}\right) \left[\frac{1}{4} + \frac{\varepsilon_c}{E_g} \left(1 + \frac{\varepsilon_c}{E_g}\right)\right] \right\}^{1/2}. \quad (5.77)$$

Для малых значений энергии, когда  $\varepsilon_c \ll E_g$ , получаем выражение (5.14)

$$\rho_c(\varepsilon_c) = \frac{1}{4\pi^2} \left(\frac{2m_c}{\hbar^2}\right)^{3/2} \varepsilon_c^{1/2}.$$

Спиновое вырождение не учитывается.

**Задача 5.4.** Выразить энергию Ферми через концентрацию электронов  $E_F$  для случая параболического закона дисперсии.

*Решение.* Воспользуемся выражением для концентрации электронов (см. (5.5)):

$$n = 2 \int_0^\infty f(\varepsilon) g(\varepsilon) d\varepsilon. \quad (5.78)$$

Плотность состояний для сферического закона дисперсии имеет вид

$$g(\varepsilon) = \frac{\sqrt{2}}{2\pi^2} \frac{m^{3/2}}{\hbar^3} \varepsilon^{1/2}. \quad (5.79)$$

Тогда для концентрации электронов найдем

$$n = \frac{2\sqrt{2}}{\pi^2} \frac{m^{3/2}}{\hbar^3} \int_0^\infty \frac{\varepsilon^{1/2} d\varepsilon}{\exp\left(\frac{\varepsilon - \mu}{k_B T}\right) + 1}. \quad (5.80)$$

Будем считать, что электроны сильно вырождены, тогда функцию распределения  $f(\varepsilon)$  можно заменить ступенькой. При этом максимальное значение энергии равно  $E_F$ . Тогда выражение (5.80) принимает вид

$$n = \frac{2\sqrt{2}}{\pi^2} \frac{m^{3/2}}{\hbar^3} \int_0^{E_F} \varepsilon^{1/2} d\varepsilon = \frac{(2m)^{3/2}}{2\pi^2 \hbar^3} E_F^{3/2}. \quad (5.81)$$

В результате

$$E_F = \frac{\hbar^2}{2m} (3\pi n)^{2/3}. \quad (5.82)$$

**Задача 5.5.** Показать, что условие  $K \gtrsim U$ , когда кинетическая энергия электрона больше потенциальной энергии взаимодействия с полем донора, соответствует критерию сильного легирования Мотта (5.65).

Кинетическая энергия электрона  $K$  порядка энергии Ферми  $E_F$ :

$$K \sim E_F = \frac{\hbar^2}{2m} (3\pi^2 n)^{2/3}. \quad (5.83)$$

Потенциальную энергию электрона можно оценить следующим образом:

$$U = \frac{e^2}{\kappa r}. \quad (5.84)$$

Здесь  $r$  — расстояние до ближайшего центра. Поэтому, зная концентрацию примесей  $N$ , можно оценить среднее расстояние как

$$r \sim N^{-1/3}. \quad (5.85)$$

В случае сильного легирования концентрация электронов равна концентрации примесей, так как все примеси ионизованы. Тогда, подставляя явные выражения для  $K$  и  $U$  через  $n$  в неравенство  $K \gtrsim U$ , получаем

$$\frac{\hbar^2}{2m} (3\pi^2)^{2/3} n^{1/3} \gtrsim \frac{e^2}{\kappa}. \quad (5.86)$$

Введем боровский радиус донора

$$a_D = \frac{\kappa \hbar^2}{m e^2}.$$

В результате выражение (5.86) принимает вид

$$n a_D^3 \gtrsim \frac{8}{(3\pi^2)^2} \sim 0,01.$$

**Задача 5.6.** Какова плотность состояний  $D(E)$  в полупроводнике с законом дисперсии

$$E(k) = \frac{\hbar^2 k^2}{2m} (1 - \gamma k^2).$$

*Указание.* Воспользуйтесь соотношением

$$\frac{V}{2\pi^2} k^2 dk = D(E) dE,$$

где  $V$  — нормировочный объем. Спиновое вырождение не учитывается.

## Глава 6

### КИНЕТИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ

#### 6.1. Проводимость полупроводников

Рассмотрим поведение электронов в полупроводнике в присутствии внешнего постоянного электрического поля  $\mathbf{E}$ . Согласно теореме об ускорении (гл. 4, § 4.3) уравнение движения электрона имеет вид

$$\frac{d\mathbf{p}}{dt} = e\mathbf{E}, \quad (6.1)$$

где  $e$  — заряд электрона, а  $\mathbf{p} = \hbar\mathbf{k}$  — квазиимпульс электрона. Из уравнения (6.1) следует, что квазиимпульс электрона в постоянном поле растет со временем неограниченно. Однако в реальных полупроводниках это не так — электрон при движении испытывает трение. Оно обусловлено взаимодействием электрона с дефектами кристаллической решетки и фононами. Таким образом, в уравнении (6.1) нужно учесть силу трения. В результате теорема об ускорении принимает вид

$$\frac{d\mathbf{p}}{dt} = e\mathbf{E} + \mathbf{F}_{\text{тр}}. \quad (6.2)$$

Теория, использующая для описания движения электронов в твердом теле уравнение (6.2), носит название теории Лоренца–Друде.

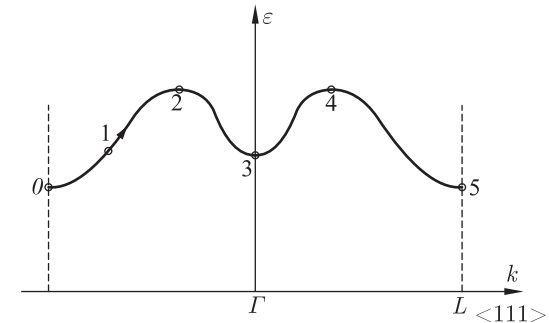


Рис. 6.1. Зависимость энергии  $\varepsilon(k)$  от волнового вектора  $k$  в одномерном случае. Точками «2», «3», «4» и «5» отмечены экстремумы  $\varepsilon(k)$ , где скорость электрона равна нулю

**Замечание.** Рассмотрим пример движения электрона в кристалле в отсутствие силы трения. Чтобы лучше понять, как происходит такое движение, рассмотрим зонную структуру Ge в направлении  $\langle 111 \rangle$  (рис. 6.1). Пусть электрон лежит на дне зоны проводимости в первой долине. Направим внешнее электрическое поле против оси  $\langle 111 \rangle$ . Тогда, согласно уравнению движения (6.1), проекция импульса электрона на ось  $\langle 111 \rangle$  линейно растет со временем (поскольку  $e < 0$ ). Скорость электрона равна  $\mathbf{v} = \partial \varepsilon / \partial \mathbf{p}$ . Следовательно, скорость электрона растет со временем (точка 1, рис. 6.1). В точке 2 (максимуме) скорость электрона равна нулю и электрон останавливается. После остановки скорость электрона становится отрицательной; электрон начинает движение в обратном направлении и в точке Г опять останавливается ( $v = 0$ ), и далее все повторяется. Схематически такое движение представлено на рис. 6.2.

Проходя первую зону Бриллюэна, электрон приходит в то же самое место  $x$ -пространства. Далее, при дальнейшем действии электрического поля квазиимпульс электрона  $\mathbf{p} = \hbar \mathbf{k}$  выходит за пределы I зоны Бриллюэна. Вычитая при этом из  $\mathbf{p}$  вектор обратной решетки, переводим электрон снова в первую зону Бриллюэна, то есть опять в точку «0» (рис. 6.1). И картина движения электрона в электрическом поле снова повторяется. Таким образом, если рассеяние отсутствует и на электрон действует внешнее электрическое поле, то электрон совершает колебательное движение в обычном  $x$ -пространстве. Переход электрона из точки «5» в точку «0» (зоны Бриллюэна) — это так называемый процесс переброса, то есть переход с изменением квазиимпульса электрона на величину вектора обратной решетки. Такое движение называется осцилляциями Блоха. Согласно квантовой механике осцилляции означают, что энергия электрона в электрическом поле квантуется (штарковская лестница). Экспериментально в однород-

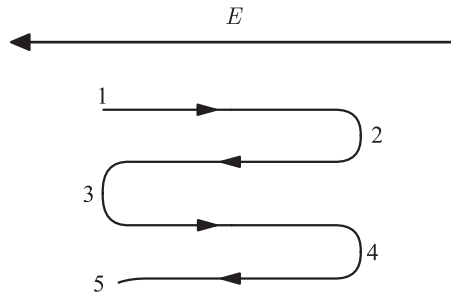


Рис. 6.2. Схематическое изображение движения электрона в электрическом поле. В области между точками «0» и «3» (рис. 6.1) скорость электрона вначале растет со временем; в точке «2» скорость электрона равна нулю (электрон останавливается). После остановки скорость электрона меняет знак — электрон движется в обратном направлении и в точке «3» опять останавливается; далее все повторяется

ных полупроводниках уровни энергии электрона в электрическом поле (штарковская лестница) не наблюдались.

В реальных полупроводниках содержится большое количество рассеивателей (дефекты, фононы и др.) Поэтому за время, необходимое для того, чтобы значение квазиимпульса  $\mathbf{p}$  заметно возросло под действием поля, электрон успевает много раз столкнуться с различными дефектами.

Частота осцилляций Блоха  $eEa/\hbar \approx 10^{10} \text{ с}^{-1}$  (где  $a$  — постоянная решетки,  $E = 300 \text{ В/см}$ ) при обычных полях меньше, чем обратное время релаксации  $\tau^{-1}$ ). Для того, чтобы имело место такое движение электрона в электрическом поле, как на рис. 6.2, электрон должен пройти всю зону Бриллюэна без столкновений. Для этого требуются гигантские электрические поля. Но при наличии сильного внешнего электрического поля раньше наступают другие процессы, такие как ударная ионизация, пробой полупроводника и т.д. Более того, при реальных значениях электрических полей электрон не выходит даже из первой долины, не доходя даже до точки «2» (рис. 6.2).

Штарковская лестница наблюдалась в искусственных многослойных структурах (сверхрешетках).

Рассмотрим теперь реальную ситуацию, когда под действием электрического поля электрон не выходит из первой долины. Вернемся к уравнению движения электрона в кристалле под действием электрического поля — к уравнению Лоренца–Друде с трением (уравнение (6.2)) — и найдем эту силу трения. Для этого рассмотрим пример.

**Математическое отступление.** Рассмотрим дифференциальное уравнение

$$\frac{dy}{dt} = -\frac{1}{\tau}y. \quad (6.3)$$

Его решение есть

$$y = y_0 e^{-t/\tau}, \quad (6.4)$$

где  $y_0$  — начальное значение величины  $y$  в момент времени  $t = 0$ . Уравнение (6.3) относится к уравнениям, описывающим распадные (релаксационные) процессы (например радиоактивный распад). Поэтому, если для какой-нибудь физической величины мы получили уравнение типа (6.3), его решение всегда имеет вид (6.4), и наоборот, если какая-нибудь физическая величина изменяется со временем по такому закону (6.4), то она, как правило, подчиняется уравнению (6.3). Как правило, распад и релаксация происходят по экспоненциальному закону (6.4).

Вернемся к уравнению движения электрона, то есть к уравнению Лоренца–Друде (6.2). Предположим, что мы выключили внешнее электрическое поле. Во внешнем поле мы сообщили электрону небольшой

<sup>1)</sup> Типичные значения  $\tau^{-1}$  в полупроводниках порядка  $10^{13} \text{ с}^{-1}$  при  $T = 300 \text{ К}$ .

квазиимпульс  $\mathbf{p}$ . При наличии силы трения этот квазиимпульс должен исчезать со временем, причем он будет затухать по закону (6.4), то есть экспоненциально. Тогда сила трения принимает вид

$$\mathbf{F}_{\text{тр}} = -\frac{\mathbf{p}}{\tau}. \quad (6.5)$$

Здесь  $\tau$  — время релаксации квазиимпульса (или просто время релаксации). Оно показывает, что если квазиимпульс электрона равен  $\mathbf{p}$ , то за время  $\tau$  его величина уменьшается в  $e$  раз за счет силы трения (то есть при рассеянии на колебаниях решетки, дефектах и др.). Таким образом, уравнение Лоренца–Друде для электрона в электрическом поле принимает вид

$$\frac{d\mathbf{p}}{dt} = e\mathbf{E} - \frac{\mathbf{p}}{\tau}. \quad (6.6)$$

Рассмотрим стационарный режим, когда квазиимпульс электрона перестает изменяться со временем, то есть  $d\mathbf{p}/dt = 0$ . В этом случае мы получаем

$$e\mathbf{E} = \frac{\mathbf{p}}{\tau}. \quad (6.7)$$

Мы получили, что в стационарном случае сила, действующая на электрон со стороны электрического поля, полностью компенсируется силой трения (например об решетку). Предположим, что стационарный квазиимпульс  $\mathbf{p} = e\mathbf{E}\tau$  настолько мал, что мы можем воспользоваться изотропным параболическим законом дисперсии для электрона  $\varepsilon = \mathbf{p}^2/2m$ . В результате квазиимпульс выражается через скорость электрона  $\mathbf{p} = m\mathbf{v}$ , где  $m$  — эффективная масса электрона. Тогда для скорости электрона получаем выражение

$$\mathbf{v} = \frac{e\mathbf{E}}{m}\tau. \quad (6.8)$$

Это та скорость, которую приобретает электрон в однородном электрическом поле, она называется дрейфовой скоростью. Коэффициент пропорциональности между дрейфовой скоростью и электрическим полем называется подвижностью и обозначается обычно буквой  $b$ :

$$\mathbf{v}_d = b\mathbf{E}; \quad b = \frac{e\tau}{m}. \quad (6.9)$$

Согласно определению, подвижность электронов равна скорости дрейфа в электрическом поле единичной напряженности.

Мы выразили подвижность  $b$  через время релаксации квазиимпульса  $\tau$ . Подвижность можно определить из эксперимента. В полупроводниках при комнатной температуре  $\tau \sim 10^{-13}$  с. Как правило, для измерения времени релаксации измеряют  $b$ . Зная эффективную массу электрона, мы находим  $\tau$  согласно (6.9).

**Замечание.** Время свободного пробега  $\tau$  можно интерпретировать несколько иначе. Пусть электроны находятся во внешнем электрическом поле  $E$ . Если в некоторый момент времени выключить внешнее

электрическое поле, то электроны будут продолжать направленное движение до тех пор, пока в результате столкновений (с колебаниями решетки или с дефектами кристалла) они не передадут решетке всю приобретенную в поле направленную скорость. Это направленное движение прекратится в среднем через время  $\tau$ . После этого электроны возвратятся в состояние хаотического теплового движения. Таким образом, столкновения приводят электроны в равновесное состояние, в то время как электрическое поле приводит к нарушению равновесного состояния. Переход системы электронов из неравновесного состояния в равновесное называют релаксационным процессом, или релаксацией, а время восстановления нарушенного равновесия называют временем релаксации. Таким образом, можно сказать, что время свободного пробега носителей заряда представляет собой время релаксации.

Если интересоваться зависимостью времени релаксации от температуры, от концентрации носителей заряда и от других параметров системы, необходимо произвести микроскопический расчет  $\tau$ , то есть учесть конкретные процессы взаимодействия электрона с дефектами кристаллической решетки и фононами (см. §§ 6.2–6.6).

**Замечание относительно уравнения Лоренца–Друде.** Уравнение движения (6.6) написано так, как если бы в полупроводнике (в системе) был только один электрон. Однако в полупроводнике электронов много и они имеют разные квазиимпульсы. Квазиимпульсы (скорости) в невырожденном газе подчиняются распределению Максвелла (см. максвелловское распределение скоростей). В нулевом электрическом поле электроны совершают хаотическое движение и средний вектор скорости равен нулю ( $\langle \mathbf{v} \rangle = 0$ ). При наличии слабого электрического поля средний вектор скорости (квазиимпульса) отличен от нуля. Таким образом, уравнение Лоренца–Друде следует понимать как уравнение для среднего вектора квазиимпульса электрона  $\langle \mathbf{p} \rangle$

$$\frac{d\langle \mathbf{p} \rangle}{dt} = e\mathbf{E} - \frac{\langle \mathbf{p} \rangle}{\tau}. \quad (6.10)$$

Пусть электрон сталкивается с какой-то примесью и рассеивается на ней (рис. 6.3). После одного столкновения у электрона остается «память» о первоначальном движении, но после многих столкновений электрон «забывает» о первоначальном движении. Если предположить, что скорость дрейфа электрона  $v_d$  меньше средней скорости беспорядочного движения  $\langle v \rangle$ , то в этом случае время релаксации  $\tau$  не зависит от средней направленной скорости (среднего квазиимпульса) и мы можем воспользоваться для решения задачи уравнением

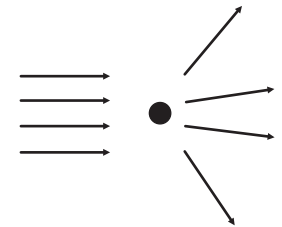


Рис. 6.3. Схема, иллюстрирующая процесс рассеяния электрона на атоме примеси



Лоренца–Друде (6.10). Условие выполнения неравенства  $\langle v \rangle \gg v_d$  рассмотрим ниже.

**Определение.** Средняя длина свободного пробега электрона  $l$  определяется равенством

$$l = \langle v \rangle \tau. \quad (6.11)$$

Следует обратить внимание на то, что  $\langle v \rangle$  — это средняя скорость хаотического (беспорядочного) движения электрона,

$$\langle v \rangle = \sqrt{\frac{8k_B T}{\pi m}}. \quad (6.12)$$

Приведем оценку  $\langle v \rangle$  для полупроводников при комнатной температуре:

$$\langle v \rangle \approx 3 \cdot 10^7 \frac{\text{см}}{\text{с}}. \quad (6.13)$$

Для оценки мы подставили в (6.12) значение  $m = 10^{-28}$  г. Однако в полупроводниках эффективная масса электрона изменяется в широких пределах от  $1,27 \cdot 10^{-29}$  г для InSb до  $1,82 \cdot 10^{-28}$  г для GaN. Поэтому средняя тепловая скорость электронов в полупроводниках имеет порядок величины  $\langle v \rangle \approx 10^7 - 10^8$  см/с. Выше мы уже отметили, что в полупроводниках при комнатной температуре время релаксации квазиимпульса<sup>1)</sup>  $\tau \approx 10^{-13}$  с. Тогда для средней длины свободного пробега при  $T = 300$  К получаем оценку:

$$l \approx 10^{-6} \text{ см} = 100 \text{ Å}. \quad (6.14)$$

Мы получили важный результат, заключающийся в том, что длина свободного пробега при  $T = 300$  К много больше постоянной решетки кристалла.

Вернемся к выражению (6.9) для дрейфовой скорости и подвижности электронов и вычислим плотность тока  $\mathbf{j}$ . Если концентрация электронов равна  $n$ , то величина, равная заряду, который проходит в единицу времени через единичное сечение, называется плотностью тока. Тогда согласно (6.9) для плотности тока  $\mathbf{j}$  мы можем записать

$$\mathbf{j} = en\mathbf{v}_d = enb\mathbf{E} = \sigma\mathbf{E}. \quad (6.15)$$

Соотношение (6.15) представляет собой закон Ома в дифференциальной форме, а величина  $\sigma$  называется удельной электрической проводимостью

$$\sigma = enb = \frac{e^2 n}{m} \tau. \quad (6.16)$$

Выражение (6.16) было получено впервые Друде. Закон Ома справедлив, пока электрическое поле не изменяет концентрацию  $n$  и по-

<sup>1)</sup> Отметим, что  $\tau$  часто называют временем свободного пробега по отношению к рассеянию импульса.

движность  $b$  электронов. Однако с ростом электрического поля  $\mathbf{E}$  концентрация электронов и их подвижность могут измениться. Вопрос о зависимости подвижности электронов от электрического поля будет рассмотрен ниже, в § 6.8 (горячие электроны).

Мы изложили теорию Лоренца–Друде для проводимости полупроводников с изотропным законом дисперсии (например, для таких, как GaAs). Для таких полупроводников, как Ge и Si, закон дисперсии в каждой долине является анизотропным (когда эффективная масса вдоль оси эллипсоида существенно отличается от поперечной массы). В Ge и Si, как уже было показано в гл. 4, имеется не одна долина, а несколько. Проводимость в отдельной долине является анизотропной, однако суммарный вклад в величину тока от всех долин дает изотропную проводимость. Такой результат является естественным, так как и Ge и Si относятся к кристаллам кубической симметрии; а в кубическом кристалле, как известно, тензор  $\Pi$  ранга сводится к скаляру и поэтому проводимость является скалярной величиной.

Рассмотрим детально проводимость таких кристаллов на примере кремния. В Si, как известно, имеется шесть эллипсоидов вдоль кристаллографических осей ( $x$ ,  $y$ ,  $z$ ) (рис. 6.4). В данном случае нам необходимо найти подвижность электронов. Будем считать, что время релаксации электронов  $\tau$  не зависит от направления электрического поля относительно кристаллографических осей, поэтому  $\tau$  является скалярной величиной. Полная плотность тока равна

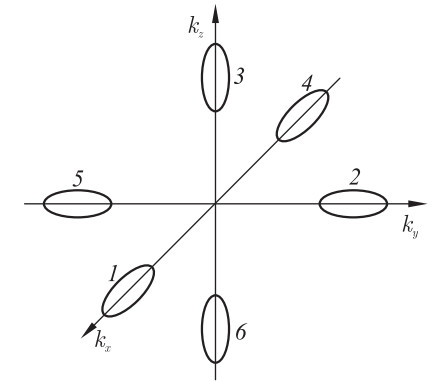


Рис. 6.4. Эллипсоиды постоянной энергии для электронов в кремнии (Si). Соответствующие параметры продольной и поперечной массы:  $m_{\parallel} = 0,98m$  и  $m_{\perp} = 0,19m$ , где  $m$  — масса свободного электрона

$$\mathbf{j} = e \sum_j n_j \mathbf{v}_{dj} = \sum_j \mathbf{j}_j. \quad (6.17)$$

Здесь  $j$  — номер долины;  $\mathbf{j}_j = en_j \mathbf{v}_{dj}$  — плотность тока в  $j$ -долине. Пусть внешнее электрическое поле направлено произвольно и имеет все три проекции,  $E_x$ ,  $E_y$  и  $E_z$ . Как видно из рис. 6.4, долины эквивалентны попарно, поэтому в (6.17) можно суммировать только три долины, а результат умножить на два. В результате выражение для плотности тока принимает вид

$$\mathbf{j} = 2e(n_1 \mathbf{v}_1 + n_2 \mathbf{v}_2 + n_3 \mathbf{v}_3). \quad (6.18)$$



Далее необходимо вычислить дрейфовую скорость  $\mathbf{v}_j$  электронов в каждой долине. Однако все долины эквивалентны и концентрация электронов в каждой долине одинакова, поэтому можно написать, что

$$n_1 = n_2 = n_3 \equiv \frac{n}{6}, \quad (6.19)$$

где  $n$  — полная концентрация электронов в кристалле. Тогда выражение для плотности тока имеет вид

$$\mathbf{j} = \frac{ne}{3}(\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3). \quad (6.20)$$

Заметим, что в каждой долине дрейфовая скорость  $\mathbf{v}_j$  анизотропна. Найдем теперь подвижность электронов в долинах. Для этого нужно воспользоваться стационарным уравнением Лоренца–Друде для среднего значения квазиимпульса (6.7). Напомним, что уравнение (6.7) было получено для случая, когда экстремумы зон (центры долин) расположены в точке  $\Gamma$  зоны Бриллюэна (в точке  $k = 0$ ). Поэтому при  $\mathbf{E} \rightarrow 0$  импульс  $\mathbf{p} \rightarrow 0$ . В Si экстремумы зон (центры долин) расположены почти на границе зоны Бриллюэна, поэтому, когда поле стремится к нулю, импульс стремится к  $\mathbf{p}_0$ , где  $\mathbf{p}_0$  — значение квазиимпульса электрона в центре долины. Поэтому для Si стационарное уравнение Лоренца–Друде имеет вид

$$e\mathbf{E} = \frac{\mathbf{p} - \mathbf{p}_0}{\tau}. \quad (6.21)$$

Далее воспользуемся теоремой о скорости

$$\mathbf{v} = \frac{1}{\hbar} \frac{\partial E_{\mathbf{k}}}{\partial \mathbf{k}}, \quad (6.22)$$

где

$$E_{\mathbf{k}} = E_0 + \frac{\hbar^2}{2} \left[ \frac{(k_x - k_{0x})^2}{m_x} + \frac{(k_y - k_{0y})^2}{m_y} + \frac{(k_z - k_{0z})^2}{m_z} \right]. \quad (6.23)$$

Тогда для скорости электрона в первой долине получаем

$$v_{1i} = \frac{p_i - p_{0i}}{m_i}, \quad (6.24)$$

где  $i = x, y, z$ . Выразим  $(p_i - p_{0i})$  из (6.21) и подставим в (6.24), тогда для компонент дрейфовой скорости электрона в первой долине имеем

$$v_{1x} = \frac{e\tau}{m_x} E_x; \quad v_{1y} = \frac{e\tau}{m_y} E_y; \quad v_{1z} = \frac{e\tau}{m_z} E_z. \quad (6.25)$$

Для первой долины большая полуось эллипсоида направлена по оси  $x$ , тогда  $m_x = m_{\parallel}$ ,  $m_y = m_z = m_{\perp}$ . В результате

$$v_{1x} = \frac{e\tau}{m_{\parallel}} E_x; \quad v_{1y} = \frac{e\tau}{m_{\perp}} E_y; \quad v_{1z} = \frac{e\tau}{m_{\perp}} E_z. \quad (6.26)$$

Во второй долине большая полуось эллипсоида направлена вдоль оси  $y$ . Тогда компоненты вектора скорости равны

$$\begin{aligned} v_{2x} &= \frac{e\tau}{m_{\perp}} E_x; & v_{2y} &= \frac{e\tau}{m_{\parallel}} E_y; \\ v_{2z} &= \frac{e\tau}{m_{\perp}} E_z. \end{aligned} \quad (6.27)$$

Аналогично для третьей долины находим

$$\begin{aligned} v_{3x} &= \frac{e\tau}{m_{\perp}} E_x; & v_{3y} &= \frac{e\tau}{m_{\perp}} E_y; \\ v_{3z} &= \frac{e\tau}{m_{\parallel}} E_z. \end{aligned} \quad (6.28)$$

Зная все компоненты вектора скорости во всех долинах, можно вычислить полный ток согласно (6.20):

$$j_i = \frac{ne^2\tau}{3} \left( \frac{1}{m_{\parallel}} + \frac{2}{m_{\perp}} \right) E_i \equiv \sigma E_i, \quad (6.29)$$

где

$$\sigma = \frac{ne^2\tau}{3} \left( \frac{1}{m_{\parallel}} + \frac{2}{m_{\perp}} \right). \quad (6.30)$$

Формула (6.30) для проводимости справедлива и для Ge. Однако вывод этой формулы для Ge является более громоздким, так как долины расположены на главных диагоналях куба.

**Отступление.** Для изотропного спектра плотность тока дается выражением (6.15), а проводимость — выражением (6.16). Обычно при анализе экспериментальных результатов принято использовать для проводимости выражение

$$\sigma = \frac{ne^2\tau}{m_c}. \quad (6.31)$$

Как в кремнии, так и в германии можно использовать выражение (6.31) для проводимости, но при этом под массой  $m_c$  нужно подразумевать эффективную массу проводимости. Если долина изотропна, то  $m_c$  совпадает с эффективной массой. Но для Si и Ge эффективная масса в каждой долине анизотропна. Тогда согласно (6.30) масса проводимости определяется соотношением

$$\frac{1}{m_c} = \frac{1}{3} \left( \frac{2}{m_{\perp}} + \frac{1}{m_{\parallel}} \right). \quad (6.32)$$

Очевидно, что масса проводимости в Si и Ge отличается от массы плотности состояний, которая равна  $m_d^3 = m_{\perp}^2 m_{\parallel}$ .

Итак, мы показали, что для кубических кристаллов (Si и Ge) электрические токи в каждой долине анизотропны, однако суммарный ток проводимости является изотропным. Проводимость при этом дается выражением (6.31), а масса проводимости — выражением (6.32).

## 6.2. Время свободного пробега (время релаксации) электронов в кристалле

Электроны в идеальной периодической решетке, согласно теореме Блоха не испытывают рассеяния, таким образом, в идеальном кристалле время свободного пробега  $\tau = \infty$ . В реальном кристалле всякое нарушение периодичности решетки приводит к рассеянию. Поэтому для определения времени релаксации  $\tau$  нужно вычислить силу трения и воспользоваться соотношением между силой трения и импульсом

$$\mathbf{F}_{\text{тр}} = -\frac{\mathbf{p}}{\tau}. \quad (6.33)$$

Нарушения периодичности кристаллической решетки могут возникать по двум причинам. Нарушения первого типа могут быть объединены под названием «статические дефекты решетки». Сюда относятся точечные дефекты (вакансии, примесные атомы) и дислокации. Как правило, рассеяние электрона на статических дефектах упругое. Возможно, однако, и неупругое рассеяние, при котором разность между энергиями начального и конечного состояний электрона равна изменению внутренней энергии рассеивателя. Например, нейтральный донор или акцептор в полупроводнике может в результате рассеяния носителя заряда перейти из основного состояния в возбужденное.

Чаще всего статические нарушения кристаллической решетки представляют собой локализованные и достаточно далеко отстоящие друг от друга дефекты. Акты рассеяния на таких дефектах можно считать независимыми, а вероятности рассеяния следует складывать. Однако такой подход возможен не всегда<sup>1)</sup>.

Общим свойством другого типа нарушений периодичности кристаллической решетки является отсутствие их локализации, когда они перемещаются по кристаллу. К такому типу нарушения, в первую очередь, относятся тепловые колебания кристаллической решетки, которые описываются при помощи квазичастиц — фононов. Кроме того, сюда относятся возбуждения в электронном спектре — экситоны, плазмоны и др. Рассеяние носителей заряда на таких нелокализованных нарушениях периодичности кристалла описывается на языке столкновений квазичастиц, при этом количество квазичастиц при столкновении не обязано сохраняться — они могут рождаться или поглощаться.

Основной величиной, описывающей рассеяние электрона, является вероятность перехода  $W(\mathbf{p}, \mathbf{p}')$  в единицу времени из состояния с квазиимпульсом  $\mathbf{p}$  в состояние с квазиимпульсом  $\mathbf{p}'$ . Вероятность  $W(\mathbf{p}, \mathbf{p}')$  удобно определять, считая, что кристалл ограничен нормировочным

<sup>1)</sup> Более детальную теорию процессов рассеяния в металлах и полупроводниках можно найти в книге: Гантмахер В. Ф., Левинсон И. Б. Рассеяние носителей тока в металлах и полупроводниках. — М.: Наука, 1985. 350 с.

объемом  $V$  и что допустимые значения импульса квазидискретны<sup>1)</sup>. Тогда вероятность  $W(\mathbf{p}, \mathbf{p}')$  — это число переходов в секунду электрона из одного квазидискретного состояния с импульсом  $\mathbf{p}$  в другое с импульсом  $\mathbf{p}'$ , размерность  $W(\mathbf{p}, \mathbf{p}')$  есть  $\text{с}^{-1}$ .

Допустим, что нам известна вероятность рассеяния  $W(\mathbf{p}, \mathbf{p}')$ . Тогда возникает вопрос, как определить силу трения. Ответ состоит в следующем. Сила трения — это тот квазиимпульс, который получает электрон от решетки в единицу времени:

$$\mathbf{F}_{\text{тр}} = \sum_{\mathbf{p}'} (\mathbf{p}' - \mathbf{p}) W(\mathbf{p}, \mathbf{p}'). \quad (6.34)$$

Здесь  $\mathbf{p} - \mathbf{p}'$  — это изменение квазиимпульса при рассеянии (см. рис. 6.5),  $W(\mathbf{p}, \mathbf{p}')$  — число переходов в секунду, суммирование в (6.34) проводится по всем квазиимпульсам.

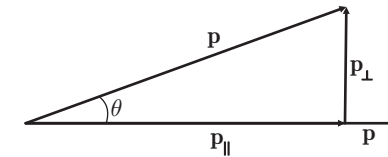


Рис. 6.5. Схема процесса рассеяния электрона. Вектор  $\mathbf{p}$  соответствует квазиимпульсу электрона до рассеяния. Вектор  $\mathbf{p}'$  — квазиимпульс электрона после рассеяния;  $\mathbf{p}'_{\perp}$  и  $\mathbf{p}'_{\parallel}$  — проекции вектора  $\mathbf{p}'$ ,  $\theta$  — угол рассеяния

Будем, для простоты, рассматривать упругое рассеяние электрона. При упругом рассеянии кинетическая энергия соударяющихся частиц практически не изменяется. Это, как правило, имеет место при рассеянии электрона на примесях. Электрон существенно легче, чем атом примеси, поэтому при рассеянии энергия электрона изменяется мало, а направление квазиимпульса изменяется существенно (рис. 6.5). Квазиимпульс электрона после рассеяния  $\mathbf{p}'$  можно представить в виде двух составляющих:  $p'_{\perp}$  и  $p'_{\parallel}$ , где  $p'_{\perp}$  — проекция вектора  $\mathbf{p}'$  на направление, перпендикулярное вектору  $\mathbf{p}$  (вектор  $\mathbf{p}$  соответствует квазиимпульсу электрона до рассеяния, это первоначальное направление квазиимпульса электрона (рис. 6.5)). Возникает вопрос о том, куда направлена сила трения. В кубических полупроводниках поперечная (по отношению к первоначальному направлению  $\mathbf{p}$ ) компонента силы трения равна нулю:

$$(\mathbf{F}_{\text{тр}})_{\perp} = \sum_{\mathbf{p}'} p'_{\perp} W(\mathbf{p}, \mathbf{p}') = 0. \quad (6.35)$$

<sup>1)</sup> Вероятность рассеяния  $W(\mathbf{p}, \mathbf{p}')$  вычисляется квантовомеханически, согласно золотому правилу квантовой механики (формула Ферми).

Действительно, в плоскости, перпендикулярной  $\mathbf{p}$ , имеется изотропия, поэтому сколько электронов будут рассеиваться в направлении  $\mathbf{p}'_{\perp}$ , столько же и в направлении  $-\mathbf{p}'_{\perp}$ , и при суммировании по  $\mathbf{p}'_{\perp}$  мы получаем нулевое значение для перпендикулярной компоненты силы трения ( $(\mathbf{F}_{\text{тр}})_{\perp} = 0$ ). Итак, остается только суммирование по  $\mathbf{p}'_{\parallel}$ . Напомним, что мы рассматриваем упругий процесс рассеяния, когда энергия электрона до и после рассеяния не меняется, и тогда  $|\mathbf{p}| = |\mathbf{p}'|$ . Согласно рис. 6.5,  $p'_{\parallel} = |\mathbf{p}'| \cos \theta \equiv p \cos \theta$ . В результате для продольной компоненты силы трения получим

$$(\mathbf{F}_{\text{тр}})_{\parallel} = -\mathbf{p} \sum_{\mathbf{p}'} W(\mathbf{p}, \mathbf{p}') (1 - \cos \theta). \quad (6.36)$$

Сравнивая это выражение для силы трения с выражением из теории Лоренца–Друде (формула (6.33)), получаем окончательное выражение для времени релаксации электронов в кристалле

$$\frac{1}{\tau} = \sum_{\mathbf{p}'} W(\mathbf{p}, \mathbf{p}') (1 - \cos \theta). \quad (6.37)$$

Видно, что время релаксации зависит через вероятность рассеяния  $W(\mathbf{p}, \mathbf{p}')$  только от модуля квазиимпульса электрона  $p$  (или от энергии  $\varepsilon$ ) и от механизма рассеяния.

Ниже мы рассмотрим несколько механизмов рассеяния электронов в кристалле.

### 6.3. Рассеяние электронов на примесях

В предыдущем параграфе было получено общее выражение для обратного времени упругого рассеяния электрона с квазиимпульсом  $\mathbf{p}$  (6.37). Из полученного результата следует, что время релаксации зависит только от квазиимпульса электрона (или его энергии  $\varepsilon$ ) и механизма его рассеяния  $W(\mathbf{p}, \mathbf{p}')$ . В этом параграфе мы вычислим время релаксации, связанное с рассеянием электронов проводимости на примесях. При описании рассеяния электронов на изолированных примесях (точечных статических дефектах) удобно воспользоваться понятием сечения рассеяния. Пусть электрон с квазиимпульсом  $\mathbf{p}$  упруго рассеивается на примеси (например он отталкивается от нее) и изменяет направление импульса на угол  $\theta$  (рис. 6.6). Свойства рассеивающей примеси будем описывать дифференциальным эффективным сечением рассеяния  $\sigma(\theta)d\Omega$ . Дифференциальное эффективное сечение рассеяния, по определению, равно числу электронов, отклоненных на угол  $\theta$  в телесный угол  $d\Omega$ , отнесенному к числу электронов, падающих на площадку площадью  $1 \text{ см}^2$ , ориентированную перпендикулярно потоку, за то же время. Другими словами, это размер площадки, на которую должен падать пучок электронов, чтобы рассеяться на телесный угол  $d\Omega$ . Дифференциальное эффективное сечение рассеяния

имеет размерность площади ( $\text{см}^2$ ) и не зависит от азимутального угла  $\varphi$ . Угол  $\varphi$  отсчитывается от некоторой плоскости, проходящей через полярную ось (ось  $z$ ), которая совпадает с направлением первоначального движения электрона. Если рассеивающий центр обладает осевой симметрией, то сечение рассеяния  $\sigma(\theta)$  не зависит от угла  $\varphi$ , но зависит от полярного угла  $\theta$  и определяет вероятность рассеяния электрона в телесный угол  $d\Omega$ , заключенный между двумя конусами  $\theta$  и  $\theta + d\theta$  (рис. 6.7). Поэтому если проинтегрировать сечение рассеяния по  $\varphi$ , то мы получим площадь кольца  $\Delta S$ , на которую должен попасть электрон, чтобы угол рассеяния лежал в интервале  $d\theta$ :  $\Delta S = 2\pi\sigma(\theta) \sin \theta d\theta$ . Это кольцо расположено перпендикулярно начальному квазиимпульсу (рис. 6.7).

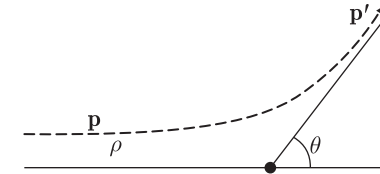


Рис. 6.6. Рассеяние электрона с начальным импульсом  $\mathbf{p}$  и прицельным параметром  $\rho$  на заряженной примеси в классической механике

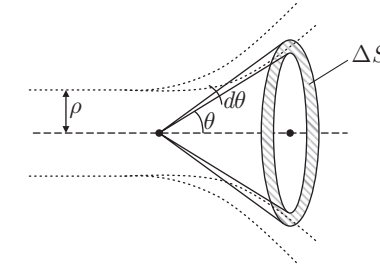


Рис. 6.7. Схема рассеяния электрона с прицельным параметром  $\rho$  в телесный угол  $d\Omega$

Найдем квазиимпульс, который все электроны отдают одной примеси, то есть силу, которая возникает при трении электронов об одну примесь, и обозначим ее через  $f_{\text{тр}}$ . Для нахождения  $f_{\text{тр}}$  нужно вычислить число рассеянных частиц. Пусть электроны падают на центр со скоростью  $v$ . Полное число электронов  $N$ , рассеянных на примеси в интервале углов между  $\theta$  и  $\theta + d\theta$  в единицу времени, равно произведению площади кольца  $\Delta S$  на концентрацию электронов  $n$  и на их скорость  $v$ :

$$N = 2\pi\sigma(\theta) \cdot n \cdot v \sin \theta d\theta. \quad (6.38)$$

Если умножить это число электронов на изменение квазиимпульса при рассеянии, равное  $\mathbf{p} - \mathbf{p}' = \mathbf{p}(1 - \cos \theta)$ , (см. рис. 6.6) и проинтегриро-

вать по  $d\theta$ , то мы получим выражение для силы трения, действующей на все электроны

$$\mathbf{f}_{\text{тр}} = 2\pi \cdot n \cdot v \mathbf{p} \int \sigma(\theta)(\cos \theta - 1) \sin \theta d\theta. \quad (6.39)$$

Разделив выражение (6.39) на число электронов в кристалле  $n \cdot V$  (где  $V$  — объем кристалла), мы получим импульс, который отдает электрон примеси в единицу времени (в 1 с)

$$\tilde{\mathbf{f}}_{\text{тр}} = \frac{2\pi v \mathbf{p}}{V} \int \sigma(\theta)(\cos \theta - 1) \sin \theta d\theta. \quad (6.40)$$

Если умножить (6.40) на число примесей в кристалле  $N_i$ , то мы найдем импульс, который отдает электрон всем примесям в кристалле в единицу времени, то есть силу трения

$$\mathbf{F}_{\text{тр}} = -2\pi v n_i \mathbf{p} \int \sigma(\theta)(1 - \cos \theta) \sin \theta d\theta \equiv -\frac{\mathbf{p}}{\tau}, \quad (6.41)$$

где  $n_i = N_i/V$  — концентрация примесей.

Таким образом, для обратного времени рассеяния электрона на примесях получили выражение

$$\frac{1}{\tau} = 2\pi n_i v \int \sigma(\theta)(1 - \cos \theta) \sin \theta d\theta. \quad (6.42)$$

Величина  $2\pi \int \sigma(\theta)(1 - \cos \theta) \sin \theta d\theta$  называется *эффективным сечением рассеяния*, или *транспортным сечением*. Часто ее называют транспортным эффективным сечением и обозначают через  $q$ . Видно, что  $q$  отличается от полного эффективного сечения рассеяния множителем  $(1 - \cos \theta)$  под интегралом. Итак, окончательное выражение для времени рассеяния электронов на примесях можно представить в виде

$$\frac{1}{\tau} = n_i v q, \quad (6.43)$$

где

$$q = 2\pi \int_0^\pi \sigma(\theta)(1 - \cos \theta) \sin \theta d\theta. \quad (6.44)$$

Отметим, что при выводе (6.43), (6.44) мы предполагали, что налетающие на примесь электроны имели заданную скорость, поэтому для расчета проводимости время рассеяния  $\tau$  необходимо усреднять по тепловому распределению носителей. Напомним, что  $v$  — это скорость беспорядочного теплового движения электрона.

Выражение (6.43) справедливо всегда, когда речь идет о рассеянии электронов на примесях. Окончательное выражение для транспортного сечения зависит от явного вида для дифференциального сечения  $\sigma(\theta)$ . Для примера рассмотрим рассеяние электронов на ионизованных цен-

трах. Такие примесные центры создают дальнедействующее кулоновское поле, и они всегда присутствуют в кристалле. В случае рассеяния на мелких ионизованных центрах дифференциальное сечение  $\sigma(\theta)$  дается формулой Резерфорда

$$\sigma(\theta) = \pi \left( \frac{2e^2}{\kappa m v^2} \right)^2 \frac{1}{\sin^4 \theta/2}, \quad (6.45)$$

где  $\kappa$  — статическая диэлектрическая проницаемость кристалла. Легко проверить, что  $\sigma(\theta)$  имеет размерность площади, величина  $2e^2/(\kappa m v^2) = r_0$  имеет размерность длины. Это то расстояние электрона от рассеивающего центра, на котором потенциальная энергия электрона в поле центра равна кинетической энергии электрона

$$\frac{e^2}{\kappa r_0} = \frac{m v^2}{2}.$$

Величина  $\pi r_0^2$  характеризует эффективную площадь рассеивающего центра. Мы видим, что дифференциальное сечение рассеяния  $\sigma(\theta)$  всегда больше площади  $\pi r_0^2$ , так как  $r_0^2$  делится на величину  $\sin^4 \theta/2$ , которая всегда меньше единицы. Когда угол рассеяния  $\theta \rightarrow 0$ , то дифференциальное сечение рассеяния стремится к бесконечности ( $\sigma(\theta) \rightarrow \infty$ ). При любом спадающем потенциале сечение рассеяния, вычисленное в рамках классической механики, будет расходиться при малых углах рассеяния<sup>1)</sup>. В случае кулоновского взаимодействия расходимость дифференциального сечения очень сильна: при  $\theta \rightarrow 0$ ,  $\sigma(\theta) \propto \theta^{-4}$ . Это является специфической особенностью кулоновского потенциала  $U(r) = e^2/\kappa r$ , который является дальнедействующим потенциалом, медленно спадающим с ростом  $r$ . Если подставить в выражение для транспортного сечения  $q$  вместо  $\sigma(\theta)$  выражение (6.45), то интеграл на нижнем пределе при  $\theta = 0$  логарифмически расходится, что обусловлено медленностью убывания кулоновского потенциала.

Проанализируем поведение транспортного сечения рассеяния на больших расстояниях от центра. Если прицельный параметр<sup>2)</sup>  $\rho$  больше среднего расстояния между примесями, то в этом случае нельзя считать, что электрон рассеивается только одной примесью; на него будут влиять и другие (расположенные в сфере объемом  $\sim \rho^3$  вблизи основной) примеси (рис. 6.8). Более важным является следующее обстоятельство: при рассеянии всегда имеются подвижные заряды, и вокруг примесного центра образуется «сфера», в пределах которой плотность электронов имеет конечную величину, и тогда возникает

<sup>1)</sup> Это верно для потенциалов, радиус действия которых бесконечен. Отметим также, что эта расходимость нам не важна, так как транспортное сечение для таких потенциалов оказывается, как правило, конечным.

<sup>2)</sup> Напомним, что прицельный параметр — это то расстояние, на котором частица прошла бы мимо центра, если бы силовое поле отсутствовало.

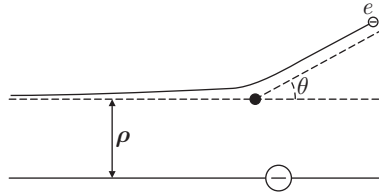


Рис. 6.8. Траектория рассеяния электрона на примесном центре. Кружок  $\ominus$  изображает расположение примеси

экранирование заряда примеси. В этом случае вместо кулоновского потенциала для электронов нужно рассматривать экранированный потенциал Дебая–Гюккеля

$$U(r) = -\frac{e^2}{\kappa r} \exp\left(-\frac{r}{r_D}\right), \quad (6.46)$$

где

$$r_D = \left(\frac{\kappa k_B T}{4\pi n e^2}\right)^{1/2} \quad (6.47)$$

— дебаевский радиус экранирования,  $n$  — концентрация электронов.

На малых расстояниях, когда  $r \ll r_D$ ,  $U(r) = -e^2/\kappa r$ , мы имеем кулоновский потенциал. На больших расстояниях, когда  $r \gg r_D$ , потенциал  $U(r)$  спадает по экспоненциальному закону, то есть быстрее, чем  $1/r$  (рис. 6.9). Часто при решении задачи о рассеянии на заряженной примеси потенциальную энергию выбирают в виде (рис. 6.10)

$$U(r) = \begin{cases} -\frac{e^2}{\kappa r}, & \text{если } r < r_D, \\ 0, & \text{если } r > r_D. \end{cases} \quad (6.48)$$

При этом ошибка за счет такого приближения невелика; более того, неважно, где «обрезать» потенциальную энергию, на величину  $r_D$  или  $2r_D$  (см. ниже).

Таким образом, на больших расстояниях при  $r \gg r_D$  взаимодействие выключается (то есть  $U(r) = 0$ ), поэтому угол рассеяния ограничен снизу величиной  $\theta_{\min}$ . Наименьшее значение угла рассеяния  $\theta_{\min}$  определяется из условия, когда прицельный параметр  $\rho$  равен дебаевскому радиусу экранирования  $r_D$ :

$$\rho = r_D. \quad (6.49)$$

В (6.49) под  $\rho$  нужно понимать максимальное значение прицельного параметра, то есть  $\rho = \rho(\theta_{\min}) = \rho_{\max}$ . Зависимость прицельного пара-

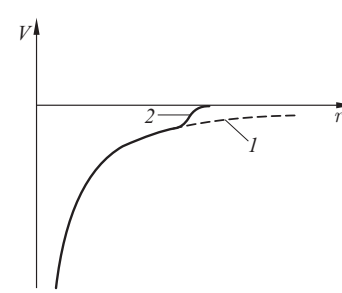


Рис. 6.9. Вид потенциальной энергии электрона для единичного заряда: «1» — неэкранированный кулоновский потенциал; «2» — экранированный дебаевский потенциал

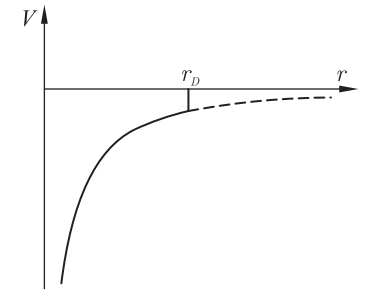


Рис. 6.10. Модельный «обрезанный» кулоновский потенциал

метра от угла рассеяния  $\theta$  дается выражением <sup>1)</sup>

$$\rho = \frac{e^2}{\kappa m v^2} \operatorname{ctg} \theta/2. \quad (6.50)$$

Теперь нетрудно найти  $\theta_{\min}$ . Для этого подставим (6.50) в (6.49) и разложим  $\operatorname{ctg} \theta/2$  в ряд, ограничиваясь первым членом разложения. В результате получим

$$\theta_{\min} = \frac{r_D}{[2e^2/(\kappa m v^2)]}. \quad (6.51)$$

Формула (6.51) справедлива, если  $r \gg r_D$ . Теперь мы можем вычислить транспортное сечение рассеяния  $q$  (см. (6.44) и (6.45))

$$q = 2\pi^2 \left(\frac{2e^2}{\kappa m v^2}\right)^2 \int_{\theta_{\min}}^{\pi} \frac{(1 - \cos \theta) \sin \theta d\theta}{\sin^4 \theta/2}. \quad (6.52)$$

Из выражения (6.52) следует, что основной вклад в интеграл дают малые значения угла  $\theta$  (см. рис. 6.11). Поэтому разложим подынтегральную функцию в ряд по  $\theta$ , кроме того, будем считать, что верхний предел интегрирования по  $\theta$  равен  $\theta_{\max}$  (причем, ввиду приближенной оценки интеграла, можно положить, что  $\theta_{\max} \sim 1$ ). Таким образом, для транспортного сечения рассеяния получаем

$$q = 16\pi^2 \left(\frac{2e^2}{\kappa m v^2}\right)^2 \ln \frac{\theta_{\max}}{\theta_{\min}}. \quad (6.53)$$

В результате транспортное сечение рассеяния слабо (логарифмически) зависит от отношения  $\theta_{\max}/\theta_{\min}$ . Поэтому до тех пор, пока

<sup>1)</sup> см. Ландау Л. Д., Лифшиц Е. М. Механика. — М.: Наука, 1988. С. 216.

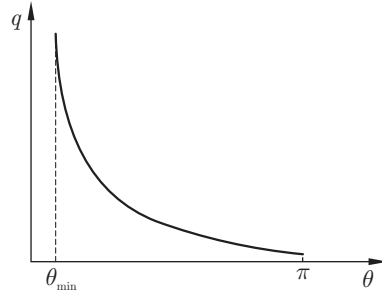


Рис. 6.11. Зависимость транспортного сечения рассеяния от угла рассеяния. Штриховой линией показано наименьшее значение угла рассеяния  $\theta_{\min}$

$\theta_{\max} \gg \theta_{\min}$ , небольшая неточность в определении  $\theta_{\min}$  не сказывается на величине  $q$ . Формулу для  $q$  можно переписать в виде

$$q = 16\pi^2 \left( \frac{e^2}{\kappa\varepsilon} \right)^2 \Lambda, \quad (6.54)$$

где  $\Lambda = \ln \left( \frac{\theta_{\max}}{\theta_{\min}} \right)$  называется кулоновским логарифмом. В (6.54) мы учли, что  $mv^2/2 = \varepsilon$ , где  $\varepsilon$  — кинетическая энергия электрона.

Видно, что транспортное сечение падает с ростом  $\varepsilon$  и пропорционально  $1/\varepsilon^2$ . Поэтому (и это очень важно), чем быстрее летит электрон, тем меньше он рассеивается.

Используя выражения (6.43) и (6.54), можно вычислить время релаксации  $\tau$

$$\tau = \frac{1}{n_i v q} = \frac{\varepsilon^{3/2}}{16\pi^2 n_i (2/m)^{1/2} (e^2/\kappa)^2 \Lambda}. \quad (6.55)$$

Мы получили, что при рассеянии электрона на заряженных примесях время релаксации пропорционально  $\varepsilon^{3/2}$  и тем больше, чем больше его энергия. Если рассматривать случай, когда  $\varepsilon \sim k_B T$ , то  $\tau \propto T^{3/2}$ . В этом случае подвижность при рассеянии электронов на ионизованных центрах растет с ростом температуры:

$$\mu = \frac{e\tau}{m} \propto T^{3/2}. \quad (6.56)$$

Из этого анализа следует, что при увеличении температуры из-за увеличения подвижности рассеяние на ионизованных примесях ослабляется.

### Некоторые замечания

**Замечание 1.** Относительно кулоновского логарифма  $\Lambda$ . Мы получили, что транспортное сечение рассеяния выражается через кулоновский логарифм

$$\Lambda = \ln \frac{\theta_{\max}}{\theta_{\min}}, \quad (6.57)$$

где

$$\theta_{\min} = \frac{r_D}{[2e^2/(\kappa m v^2)]}. \quad (6.58)$$

Если решить задачу о рассеянии электрона на кулоновском потенциале классически, то минимальный прицельный параметр — это тот радиус, при котором кинетическая энергия электрона равна потенциальной энергии:

$$\rho_{\min}^{\text{кл}} = \frac{2e^2}{\kappa m v^2}. \quad (6.59)$$

Кулоновский логарифм равен

$$\Lambda \approx \ln \frac{1}{\theta_{\min}} = \ln \frac{r_D}{[2e^2/(\kappa m v^2)]}, \quad (6.60)$$

где мы учли, что  $\rho_{\max} = r_D = \left( \frac{\kappa k_B T}{4\pi n e^2} \right)^{1/2}$ . В квантовой механике, если вычислить  $\rho_{\max}$  и  $\rho_{\min}$  в рамках теории возмущений, то для  $\rho_{\max}$  получается практически тот же результат. Однако для  $\rho_{\min}$  получается несколько другое выражение. Это связано с тем, что, когда электрон приближается к центру рассеяния, то есть когда минимальный прицельный параметр становится порядка  $\hbar/mv$  (порядка длины волны де-Бройля), классическое рассмотрение неприменимо. В результате получается, что

$$\rho_{\min}^{\text{кв}} = \frac{\hbar}{mv}. \quad (6.61)$$

Из сказанного ясно, что классическое рассмотрение справедливо при условии, что

$$\rho_{\min} \gg \frac{\hbar}{mv}. \quad (6.62)$$

Это означает, что  $\rho_{\min}$  больше, чем предел, продиктованный соотношением неопределенностей.

**Замечание 2.** Мы рассмотрели случай упругого рассеяния электрона на примеси, когда энергия электрона практически не меняется, а изменяется только направление импульса. В рамках такой постановки задачи захват электрона на примеси невозможен. Чтобы электрон захватился на примеси, он должен отдать энергию третьей частице.

**Замечание 3.** Условие применимости теории:  $\rho_{\max} \gg \rho_{\min}(\theta_{\max} \gg \theta_{\min})$ . Если эти неравенства нарушаются, то изложенная выше теория несправедлива.

Условие  $\rho_{\max} \gg \rho_{\min}$  можно переписать в виде

$$n \cdot r_D^3 \gg 1. \quad (6.63)$$

Это неравенство означает, что в сфере дебаевского радиуса должно быть много электронов. Дебаевский радиус также зависит от концентрации электронов  $r_D \propto n^{-1/2}$  (см. (6.47)). Поэтому соотношение (6.63) выполняется при не слишком высоких концентрациях электронов. Таким образом, изложенная выше теория о рассеянии электронов на примесях справедлива при низких концентрациях и высоких температурах, в случае больших концентраций изложенная теория о рассеянии не очень хороша. В теории многих частиц параметр  $n \cdot r_D^3$  называется параметром газовой, его можно переписать с учетом выражения для  $r_D$  в виде

$$\frac{e^2}{\kappa} n^{1/3} \ll k_B T. \quad (6.64)$$

Из формулы (6.64) следует, что, когда средняя потенциальная энергия электрона меньше кинетической, мы имеем газ электронов; а когда выполняется обратное неравенство, можно считать, что мы имеем жидкость.

Весь предыдущий анализ справедлив для невырожденных электронов, когда они подчиняются статистике Больцмана.

Если электронный газ вырожден, то кинетическая энергия электрона уже определяется не температурой, а энергией Ферми  $E_F$ . Тогда в (6.63) и (6.64) необходимо заменить  $k_B T$  на энергию Ферми  $E_F$ . Тогда условие газовой (6.64) принимает вид

$$\frac{e^2}{\kappa} n^{1/3} \ll E_F. \quad (6.65)$$

Мы уже показали (см. гл. 5), энергия Ферми  $E_F \propto n^{2/3}$ . Тогда условие (6.65) совпадает с условием сильного легирования, то есть, чем больше концентрация электронов, тем лучше выполняется условие.

Итак, мы можем заключить, что для вырожденного ферми-газа, чем выше концентрация, тем лучше выполняется условие газовой; а для невырожденного газа, наоборот, чем меньше концентрация, тем хуже выполняется условие газовой.

**Замечание 4.** Рассеяние на нейтральных примесях. Рассеяние на нейтральных примесях в полупроводниках исследовано сравнительно мало. Основной теоретический результат состоит в применении к полупроводнику расчетов сечения рассеяния свободного электрона атомом водорода путем введения в формулы для сечения рассеяния диэлектрической проницаемости  $\kappa$  и эффективной массы  $m$  вместо массы свободного электрона  $m_0$ . Очевидно, что сечение рассеяния на нейтральной примеси существенно меньше, чем на ионизованной примеси. Однако рассеянием на нейтральных примесях пренебрегать нельзя, особенно при очень низких температурах, когда число ней-

тральных атомов примеси намного превышает число ионизованных примесных атомов. Можно показать, что подвижность при рассеянии на нейтральных центрах слабо зависит от энергии <sup>1)</sup>.

## 6.4. Рассеяние на тепловых колебаниях решетки

Фундаментальным процессом рассеяния электронов в кристалле является их взаимодействие с тепловыми колебаниями атомов кристалла. Эти колебания имеются даже в совершенном (идеальном) кристалле. Очевидно, что другие виды рассеяния (например на примесях) связаны с наличием различного рода несовершенств кристалла.

Когда мы говорим, что атомы совершенного кристалла занимают в нем положения, образующие идеальную периодическую решетку, то это вовсе не означает, что в каждый данный момент времени атомы в кристалле образуют идеальную решетку. При отличных от нуля температурах атомы совершают небольшие колебания относительно фиксированных положений равновесия, причем средняя энергия колебаний определяется законом распределения энергии по степеням свободы. Согласно классической физике при абсолютном нуле температуры атомы должны покоиться в положениях равновесия, и в этом случае их можно рассматривать как узлы идеальной решетки. Однако согласно квантовой теории, положение атомов не может быть фиксировано в силу соотношения неопределенностей (имеются нулевые колебания). Следовательно, слова о том, что атомы кристалла «занимают определенное место», следует понимать в том смысле, что среднее положение каждого атома в идеальном кристалле (при усреднении за достаточно большой промежуток времени по сравнению с периодом колебаний) совпадает с соответствующим узлом идеальной периодической решетки.

Таким образом, можно считать, что при абсолютном нуле температуры атомы покоятся в узлах решетки, а их беспорядочные смещения при более высоких температурах можно рассматривать как случайные колебания около соответствующих положений равновесия. Такие колебания обычно и носят название тепловых колебаний решетки.

При тепловых колебаниях решетки (поскольку решетка уже не идеальна) электрон при движении испытывает рассеяние на решетке.

Приведем качественный вывод выражения для сечения рассеяния электрона на тепловых колебаниях решетки (Я. И. Френкель). Если атом не колеблется, то он не влияет на движение электрона, атом влияет на движение электрона (рассеивает электрон) в меру своего отклонения от положения равновесия. Поэтому можно предположить, что транспортное сечение рассеяния пропорционально квадрату ампли-

<sup>1)</sup> Ансельм А. И. Введение в теорию полупроводников. — М.: Наука, 1978. 615 с.



туды колебания атома  $X^2$  ( $q \sim X^2$ ). Амплитуду тепловых колебаний можно найти из условия равенства средней потенциальной энергии осциллятора средней кинетической энергии

$$\gamma X^2 \sim k_B T. \quad (6.66)$$

Таким образом, мы получаем, что транспортное сечение зависит от температуры:

$$q \propto k_B T. \quad (6.67)$$

В результате для частоты столкновений, согласно (6.43) и (6.67), получим

$$\frac{1}{\tau} \propto q \cdot v \propto T^{3/2}. \quad (6.68)$$

Здесь мы учли, что  $v$  — это средняя тепловая скорость электрона и  $v \propto T^{1/2}$ . Таким образом, при рассеянии на тепловых колебаниях решетки подвижность электронов убывает с ростом температуры, как

$$b \propto T^{-3/2}. \quad (6.69)$$

Полученный результат понятен, поскольку, чем больше температура, тем больше колебания кристаллической решетки. Это означает, что электрон будет сильнее рассеиваться на колебаниях решетки. В результате подвижность падает. Следует еще раз отметить, что весь этот вывод — качественный.

Возбуждение колебаний кристаллической решетки принято описывать как рождение фононов, квазичастиц, имеющих, как и электроны, определенный закон дисперсии  $\omega = \omega_s(\mathbf{q})$ , где  $\mathbf{q}$  — квазиимпульс фонона,  $\omega$  — его частота, а индекс  $s$  нумерует различные ветви фононного спектра (акустические и оптические, продольные и поперечные). Поэтому строгая теория о рассеянии электронов на колебаниях решетки рассматривает взаимодействие электронов с фононами. В этом случае процесс рассеяния движущихся электронов описывается в терминах поглощения и испускания фононов. В настоящей главе изучаются только однофононные процессы, простейшие и наиболее вероятные, в которых излучается или поглощается только один фонон.

На рис. 6.12 представлены фейнмановские диаграммы<sup>1)</sup> для двух указанных процессов: а) излучения фонона и б) поглощения фонона. Такие процессы рассеяния подчиняются законам сохранения энергии и импульса. Например, рассмотрим случай излучения (рис. 6.12, а). Закон сохранения импульса определяется равенством

$$\hbar \mathbf{k} = \hbar \mathbf{k}' + \hbar \mathbf{q}. \quad (6.70)$$

<sup>1)</sup> Диаграммная техника Фейнмана позволяет по рисунку восстановить аналитическое выражение для вероятности перехода. Мы не будем здесь формулировать соответствующие правила. Отметим только, что в точках контакта линий (вершинах) должен выполняться закон сохранения квазиимпульса (6.70) и энергии (6.71).

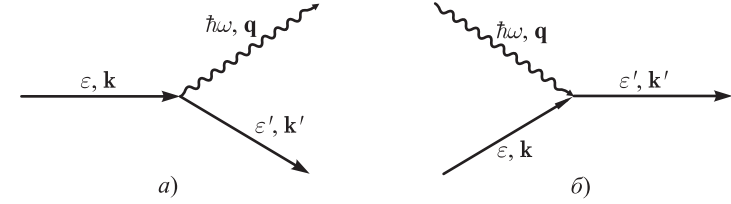


Рис. 6.12. Фейнмановские диаграммы для процессов: а) излучения фонона; б) поглощения фонона. На этих диаграммах прямые линии изображают электрон, волнистые — фонон

Аналогично, закон сохранения энергии можно записать в виде

$$\epsilon = \epsilon' + \hbar\omega_{\mathbf{q}}. \quad (6.71)$$

Здесь  $\omega_{\mathbf{q}}$  — частота фонона с волновым вектором  $\mathbf{q}$ . Частота фонона  $\omega_{\mathbf{q}}$  по-разному зависит от волнового вектора для акустических и оптических фононов (продольных и поперечных). Ниже мы отдельно рассмотрим рассеяние на акустических и оптических фононах.

## 6.5. Рассеяние на акустических фононах

**а) Излучение фонона.** Закон дисперсии для акустических фононов представлен на рис. 6.13. Значение  $q = \pi/a \equiv k_{\text{пред}}$  — это гигантская величина по сравнению с тепловым волновым вектором фонона, поэтому интересующие нас фононы имеют  $q$  и  $\omega$  вблизи 0. Тогда закон дисперсии для тепловых акустических фононов может быть представлен в виде (см. гл. 3)

$$\omega_{\mathbf{q}} = s q, \quad (6.72)$$

где  $s$  — скорость звука;  $q$  — волновой вектор фонона. Для процесса излучения фонона законы сохранения энергии и импульса имеют вид

$$\frac{\hbar^2 \mathbf{k}^2}{2m} = \frac{\hbar^2 \mathbf{k}'^2}{2m} + \hbar s q, \quad (6.73)$$

$$\mathbf{k} = \mathbf{k}' + \mathbf{q}, \quad (6.74)$$

где  $\mathbf{k}$  — начальный волновой вектор электрона, а  $\mathbf{k}'$  — волновой вектор электрона после излучения фонона.

Если проанализировать уравнения (6.73) и (6.74), то оказывается, что электрон при обычных условиях излучения акустического фонона теряет очень небольшую долю энергии. Поэтому в нулевом приближении мы считаем, что энергия электрона при рассеянии на акустических фононах не изменяется. Тогда  $k^2 \approx (k')^2$ . Направление волнового вектора электрона при рассеянии изменяется, но его величина (в нулевом приближении) практически остается прежней. Энергия фонона в нулевом приближении, согласно (6.72) и (6.74), равна

$$\hbar\omega_{\mathbf{q}} = \hbar s q = \hbar s |\mathbf{k} - \mathbf{k}'|. \quad (6.75)$$

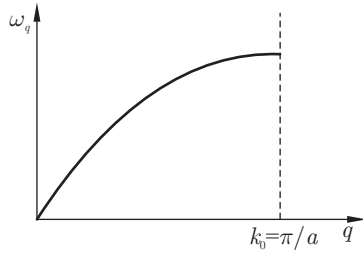


Рис. 6.13. Схематическое изображение закона дисперсии для акустических фононов в одномерной цепочке атомов

Если возвести (6.75) в квадрат и воспользоваться соотношением  $k^2 = (k')^2$ , получаем

$$(\hbar\omega_q)^2 = (\hbar s)^2 2k^2(1 - \cos\theta) = 4(\hbar s)^2 k^2 \sin^2 \theta/2, \quad (6.76)$$

где  $\theta$  — угол рассеяния электрона (угол между  $\mathbf{k}$  и  $\mathbf{k}'$ ). Выразим  $k^2$  через энергию электрона  $\varepsilon$ :

$$k^2 = \frac{2m\varepsilon}{\hbar^2}. \quad (6.77)$$

Тогда соотношение (6.76) для энергии излучаемого фонона принимает вид

$$\hbar\omega_q = (8ms^2\varepsilon)^{1/2} |\sin \theta/2|. \quad (6.78)$$

Таким образом, максимальную энергию фонона, которую может излучить (поглотить) электрон при рассеянии на акустических фононах, можно выразить соотношением

$$\hbar\omega_q^{\max} = (8ms^2\varepsilon)^{1/2}. \quad (6.79)$$

Эта энергия соответствует случаю, когда электрон рассеивается назад ( $\theta = \pi$ ).

Итак, для ответа на вопрос, верны ли наши предположения, сделанные в начале данного параграфа, необходимо сравнить максимальную энергию фонона (6.79) со средней энергией электрона  $\varepsilon \sim k_B T$ . Для этого достаточно показать, что энергия  $8ms^2$  меньше энергии электрона  $\varepsilon$ , или, что то же самое, показать, что температура  $T_s$ , отвечающая энергии  $8ms^2$ , меньше  $T$ . Дадим оценку для  $T_s$ :

$$T_s = \frac{8ms^2}{k_B} \approx 5 \text{ K}. \quad (6.80)$$

В (6.80) при оценке  $T_s$  мы учли, что масса электрона равна  $m \sim 10^{-27}$  г; скорость звука равна  $s \approx 3 \cdot 10^5$  см/с. Очевидно, что, если электроны находятся при комнатной температуре, то  $T_s \ll T$ , или, что то же самое, средняя (тепловая) энергия электрона  $\varepsilon \gg 8ms^2$ .

Таким образом, при высоких температурах (порядка комнатной) электрон излучает фононы с импульсом порядка импульса электрона, но с энергией, гораздо меньшей энергии электрона. Итак, при излучении акустического фонона электрон может потерять импульс (то есть направление импульса может изменяться как угодно), но энергию он теряет очень мало. На языке скоростей это утверждение означает, что скорость акустического фонона много меньше тепловой скорости электрона ( $s \ll v$ ). На классическом языке акустические фононы — это волна деформации, причем эта волна распространяется со скоростью звука  $s$ . Если скорость электрона  $v \gg s$ , то электрон воспринимает эту волну как стоячую деформацию.

При низких температурах порядка гелиевой ( $T \sim 4,2$  К) изложенная выше теория неверна, так как при рассеянии электрон отдает фонону большую долю энергии. В этом случае тепловая скорость электрона сравнивается со скоростью фонона  $s$ .

**б) Поглощение фонона.** В случае поглощения фонона (рис. 6.12, б) (по сравнению с излучением) изменяются только знаки в законах сохранения (6.73) и (6.74)

$$\frac{\hbar^2 \mathbf{k}^2}{2m} = \frac{\hbar^2 \mathbf{k}'^2}{2m} - \hbar s q, \quad \mathbf{k} = \mathbf{k}' - \mathbf{q}. \quad (6.81)$$

Тогда энергия фонона  $\hbar\omega_q$ , которую может поглотить электрон, дается тем же выражением (6.78)

$$\hbar\omega_q = (8ms^2\varepsilon)^{1/2} |\sin \theta/2|. \quad (6.82)$$

Максимальную энергию электрон поглощает, когда он рассеивается назад ( $\theta = \pi$ ).

Итак, при поглощении электроном фонона изменяются только знаки в законе сохранения. Однако максимальная энергия, которую может поглотить электрон при рассеянии на акустических фононах, дается выражением (6.79).

Мы получили, что при рассеянии электронов на акустических фононах излучаются или поглощаются фононы с энергией  $\hbar\omega_q \ll \varepsilon$ , если  $T \gg 8ms^2$ , где энергия электрона  $\varepsilon \sim k_B T$ . Энергия тепловых фононов в кристалле — это величина порядка  $k_B T$ . Поэтому энергия фононов, излучаемых или поглощаемых электроном, много меньше энергии тепловых фононов:  $\hbar\omega_q \ll k_B T$ .

Таким образом, можно заключить, что законы сохранения энергии и импульса (6.73) и (6.74) разрешают электронам взаимодействовать не с тепловыми фононами, а с дотепловыми фононами, энергия которых много меньше,  $\hbar\omega_q \ll k_B T$ .

Найдем число дотепловых фононов в кристалле. В состоянии термодинамического равновесия распределение фононов по энергии опре-

деляется формулой Планка. Таким образом, число дотепловых фононов в кристалле с энергией  $\hbar\omega_q \ll k_B T$  можно найти по формуле Планка:

$$N_q = \frac{1}{e^{\hbar\omega_q/k_B T} - 1} \approx \frac{k_B T}{\hbar\omega_q} \gg 1. \quad (6.83)$$

Итак, дотепловых фононов в кристалле очень много и их число пропорционально температуре  $T$ . В кристалле имеются три акустические ветви: две поперечные ветви и одна продольная. Электроны, как правило, сильнее взаимодействуют с продольными акустическими фононами; взаимодействие с поперечными акустическими фононами значительно слабее.

Рассмотрим вероятность излучения или поглощения фонона  $W_{\mathbf{p}\mathbf{p}'}$ , которая входит в выражение для времени рассеяния  $\tau$ . Отметим, что  $W_{\mathbf{p}\mathbf{p}'}$  — это есть вероятность (в единицу времени) того, что электрон, имеющий квазиимпульс  $\mathbf{p}$ , рассеиваясь, приобретает квазиимпульс  $\mathbf{p}'$ . Для вычисления вероятности  $W_{\mathbf{p}\mathbf{p}'}$  нужно воспользоваться золотым правилом квантовой механики при условии, что известен закон взаимодействия электронов с фононами. Золотое правило Ферми дает вероятность  $W_{\mathbf{p}\mathbf{p}'}$  — вероятность элементарного акта, когда электрон переходит из состояния  $\mathbf{p}$  в состояние  $\mathbf{p}'$  с излучением (поглощением) фонона. Если электрон поглощает несколько фононов, то вероятность поглощения  $W_{\mathbf{p}\mathbf{p}'}$  нужно умножить на число фононов  $N_q$ :  $W_{\mathbf{p}\mathbf{p}'}^0 \cdot N_q$ . Вероятность испускания фононов, казалось бы, не зависит от того, имеются ли в кристалле фононы или нет. При поглощении электроном фононов ситуация принципиально другая — в кристалле должны быть фононы. Однако наряду со спонтанным излучением, как показал Эйнштейн, существует еще эффект индуцированного излучения. Фонон, взаимодействуя с электроном, «вынуждает» его излучать свою реплику (то есть точно такой же фонон). На аналогичном эффекте излучения фотонов основан принцип работы лазера (см. гл. 8). Вероятность индуцированного излучения равна вероятности спонтанного излучения  $W_{\mathbf{p}\mathbf{p}'}^0$ , умноженной на число фононов  $N_q$ .

Таким образом, в выражении для вероятности взаимодействия электронов с фононами присутствуют три вклада: поглощение  $W_{\mathbf{p}\mathbf{p}'}^0 \cdot N_q$ ; спонтанное излучение  $W_{\mathbf{p}\mathbf{p}'}^0$  и индуцированное излучение  $W_{\mathbf{p}\mathbf{p}'}^0 \cdot N_q$ .

Согласно золотому правилу квантовой механики вероятность перехода электрона в единицу времени из состояния с квазиимпульсом  $\mathbf{p}$  в состояние с квазиимпульсом  $\mathbf{p}'$ , сопровождаемая испусканием фонона, равна

$$W_{\mathbf{p}\mathbf{p}'}^0 = \frac{2\pi}{\hbar} |V_{\mathbf{p}\mathbf{p}'}|^2 \delta(\varepsilon_{\mathbf{p}} - \varepsilon_{\mathbf{p}'} - \hbar\omega_q). \quad (6.84)$$

Здесь  $V_{\mathbf{p}\mathbf{p}'}$  — матричный элемент электрон-фононного взаимодействия,  $\delta$ -функция изображает закон сохранения энергии (в данном случае

для процесса излучения). Процедура вычисления вероятности согласно (6.84) громоздка, но все же выполнима.

Рассмотрим подробнее матричный элемент взаимодействия электрона с фононом. Наличие фононов означает, что в кристалле происходит деформация решетки. Рассматривая ниже тензорезистивный эффект (см. задача 6.1), мы покажем, что при деформации дно зоны проводимости смещается по энергии вверх или вниз. Для случая же взаимодействия электронов с фононами задача о нахождении изменения энергии электрона при взаимодействии с фононами была решена феноменологически Бардиным и Шокли. Энергия взаимодействия электрона с фононом (в случае квадратичного закона дисперсии для электронов и кубического кристалла)  $V(\mathbf{r})$  должна быть пропорциональной деформации, то есть тензору деформации:

$$\varepsilon_{ik} = \frac{1}{2} \left( \frac{\partial U_k}{\partial x_i} + \frac{\partial U_i}{\partial x_k} \right). \quad (6.85)$$

Тензор деформации  $\varepsilon_{ik}$  зависит не от самих смещений  $\mathbf{U}$ , а от их производных. Если смещения  $\mathbf{U}$  во всех точках кристалла одинаковы, то это не деформация, а просто сдвиг кристалла как целого. В нашем случае потенциальная энергия взаимодействия  $V(\mathbf{r})$  — это инвариант, но если  $V(\mathbf{r})$  пропорционально  $\varepsilon_{ik}$ , тогда

$$V(\mathbf{r}) = \sum_{ik} A_{ik} \varepsilon_{ik}, \quad (6.86)$$

где  $A_{ik}$  — тензор второго ранга.

Как уже было показано выше, в кубическом кристалле тензор второго ранга сводится к скаляру, следовательно,  $A_{ik} = a\delta_{ik}$ . Тогда для потенциальной энергии взаимодействия электрона с фононами получаем

$$V(\mathbf{r}) = a(\varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}), \quad (6.87)$$

где  $\varepsilon_{ii}$  — диагональные компоненты тензора деформации.

Получим выражение (6.87) другим, более простым способом. Нужно из тензора II ранга построить инвариант. Таким инвариантом является шпур тензора деформации второго ранга  $\text{Sp } \varepsilon = \varepsilon_{xx} + \varepsilon_{yy} + \varepsilon_{zz}$ . Итак, мы снова получили, что энергия взаимодействия электрона с фононами равна

$$V(\mathbf{r}) = a \left( \frac{\partial U_x}{\partial x} + \frac{\partial U_y}{\partial y} + \frac{\partial U_z}{\partial z} \right) \equiv a \text{div } \mathbf{U}, \quad (6.88)$$

где  $a$  — это константа деформационного потенциала, которая обозначается через  $E_0$ . Потенциал  $V(\mathbf{r})$  — деформационный потенциал;  $\mathbf{U}(\mathbf{r})$  — это смещение решетки в той точке  $\mathbf{r}$ , в которой находится электрон. Итак, для потенциальной энергии взаимодействия электрона с фоно-

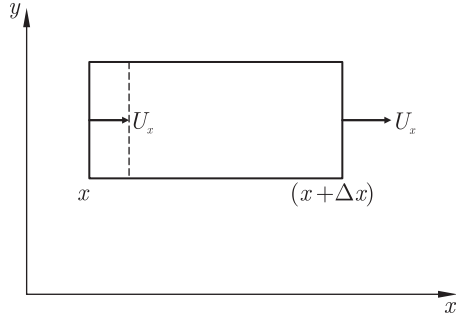


Рис. 6.14. Геометрический вывод формулы (6.91). Построение проведено для случая деформации вдоль оси  $x$

нами мы получили выражение

$$V(\mathbf{r}) = E_0 \operatorname{div} \mathbf{U}(\mathbf{r}). \quad (6.89)$$

Константа деформационного потенциала  $E_0$ , как правило, определяется из эксперимента.

#### Физический смысл $\operatorname{div} \mathbf{U}$

Для простоты рассмотрим первый член в выражении для  $\operatorname{div} \mathbf{U}$ . Выделим прямоугольник в кристалле (рис. 6.14). Левая точка имеет координату  $(x)$ , правая  $(x + \Delta x)$ . Представим, что кристалл деформирован, тогда точка  $(x)$  изменит свое положение; при этом смещение вдоль оси  $x$  обозначим через  $U_x$ ;  $U_x$  — это есть проекция вектора смещения  $\mathbf{U}$  на ось  $x$ . Точка  $(x + \Delta x)$  тоже сместится при этом, но так как деформация неоднородная, то она сместится иначе. Найдем, как при такой деформации изменится  $\Delta x$  (то есть найдем основание выделенного параллелепипеда):

$$\delta \Delta x = U_x(x + \Delta x) - U_x(x) = \frac{\partial U_x}{\partial x} \Delta x. \quad (6.90)$$

Поэтому относительное удлинение основания параллелепипеда равно

$$\frac{\delta \Delta x}{\Delta x} = \frac{\partial U_x}{\partial x}. \quad (6.91)$$

Итак, мы получили, что  $\partial U_x / \partial x$  — это есть относительное удлинение «кусочка» материала (в данном случае, параллелепипеда) вдоль оси  $x$  при деформации. Точно так же  $\partial U_y / \partial y$  и  $\partial U_z / \partial z$  есть относительное удлинение параллелепипеда вдоль соответствующих осей при деформации. Тогда для  $\operatorname{div} \mathbf{U}$  имеем

$$\begin{aligned} \operatorname{div} \mathbf{U} &= \frac{\partial U_x}{\partial x} + \frac{\partial U_y}{\partial y} + \frac{\partial U_z}{\partial z} \equiv \frac{\delta \Delta x}{\Delta x} + \frac{\delta \Delta y}{\Delta y} + \frac{\delta \Delta z}{\Delta z} = \\ &= \frac{\delta(\Delta x \cdot \Delta y \cdot \Delta z)}{\Delta x \cdot \Delta y \cdot \Delta z} = \frac{\delta V}{V}. \end{aligned} \quad (6.92)$$

Соотношение (6.92) представляет собой относительное изменение объема кристалла при деформации:

$$\operatorname{div} \mathbf{U} = \frac{\delta V}{V}, \quad (6.93)$$

где  $V = \Delta x \cdot \Delta y \cdot \Delta z$  — объем выделенного в кристалле параллелепипеда. Итак,  $\operatorname{div} \mathbf{U}$  есть не что иное, как относительное изменение объема, или, что то же самое, плотности кристалла  $\Delta \rho / \rho$  при деформации. Это значит, что электроны в кристалле при взаимодействии с фононами (при деформации) реагируют только на изменение объема (или плотности) в данной точке кристалла.

**Расчет матричного элемента электрон-фононного взаимодействия.** Вектор смещения  $\mathbf{U}(\mathbf{r})$  точки  $\mathbf{r}$  кристалла в приближении непрерывного континуума можно разложить по плоским волнам

$$\mathbf{U}(\mathbf{r}) = \sum_{\mathbf{q}, j} \mathbf{U}_{\mathbf{q}, j} e^{i\mathbf{q}\mathbf{r}}. \quad (6.94)$$

Здесь  $\mathbf{U}_{\mathbf{q}, j}$  — амплитуда,  $\mathbf{q}$  — волновой вектор волны, а индекс поляризации звуковой волны  $j$  пробегает три значения, соответствующие одной продольной и двум поперечным ветвям акустических колебаний. Тогда

$$\operatorname{div} \mathbf{U}(\mathbf{r}) = i \sum_{\mathbf{q}, j} (\mathbf{q} \mathbf{U}_{\mathbf{q}, j}) e^{i\mathbf{q}\mathbf{r}}. \quad (6.95)$$

На основании (6.95) можно сделать важный вывод. Так как  $\operatorname{div} \mathbf{U}$  пропорциональна скалярному произведению амплитуды смещения на волновой вектор фонона, то для поперечных фононов (для которых  $\mathbf{q} \perp \mathbf{U}_{\mathbf{q}, j}$ ) получаем  $(\mathbf{q} \mathbf{U}_{\mathbf{q}, j}) = 0$ . Таким образом, если закон дисперсии для электронов изотропен, то отсутствует взаимодействие электронов с поперечными фононами (их две ветви), и следовательно, остается взаимодействие только с одной продольной акустической волной. Тогда в (6.95) сумму по  $j$  можно опустить. Для определения вероятности перехода электрона  $W_{\mathbf{p}\mathbf{p}'}$  по формуле (6.84) вычислим матричный элемент энергии взаимодействия (6.89) на волновых функциях электрона в приближении плоских волн:

$$V_{\mathbf{p}\mathbf{p}'} = \int_{\Omega} \psi_{\mathbf{p}'}^*(\mathbf{r}) V(\mathbf{r}) \psi_{\mathbf{p}}(\mathbf{r}) d^3r, \quad (6.96)$$

где  $\psi_{\mathbf{p}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \exp\left(\frac{i}{\hbar} \mathbf{p} \mathbf{r}\right)$ ,  $\Omega$  — нормировочный объем кристалла.

В результате, для матричного элемента электрон-фононного взаимодействия, с учетом (6.89), (6.95) и (6.96) получаем

$$V_{\mathbf{p}\mathbf{p}'} = i E_0 \sum_{\mathbf{q}} q U_q \frac{1}{\Omega} \int_{\Omega} e^{i(\mathbf{k} + \mathbf{q} - \mathbf{k}') \mathbf{r}} d^3r. \quad (6.97)$$

Интеграл в (6.97) отличен от нуля и равен  $\Omega$  только, если  $\mathbf{k} + \mathbf{q} = \mathbf{k}'$ :

$$\frac{1}{\Omega} \int_{\Omega} e^{i(\mathbf{k}+\mathbf{q}-\mathbf{k}')\mathbf{r}} d^3r = \delta_{\mathbf{k}+\mathbf{q},\mathbf{k}'}. \quad (6.98)$$

В этом случае равенство (6.98) дает закон сохранения квазиимпульса электрона при поглощении фонона. Аналогично получается выражение для закона сохранения при излучении фонона; в этом случае

$$\mathbf{k} - \mathbf{q} = \mathbf{k}'. \quad (6.99)$$

При квантовомеханическом описании следует рассматривать не акустическую волну, а фононы с энергией  $\hbar\omega_q$  и импульсом  $\hbar\mathbf{q}$ . Поэтому случай (6.98) соответствует процессу, при котором электрон поглощает фонон с волновым вектором  $\mathbf{q}$  и энергией  $\hbar\omega_q$ , а случай (6.99) соответствует процессу испускания фонона. Итак, с учетом сказанного квадрат модуля матричного элемента для процессов испускания или поглощения равен:

$$|V_{\mathbf{p}\mathbf{p}'}|^2 = E_0^2 q^2 |U_q|^2. \quad (6.100)$$

При квантовомеханическом рассмотрении амплитуды колебаний  $U_q$  и  $U_q^*$  следует считать операторами (см. гл.3, приложение к §3.4, формулы (3.115) и (3.121)). Поскольку мы рассматриваем гармонические колебания, то матричные элементы этих операторов связывают колебательные состояния  $N_q$  и  $N_q + 1$ , отличающимися на один фонон (они соответствуют процессам рождения и уничтожения фонона в зависимости от того, происходит ли поглощение или испускание фонона). Тогда квадрат амплитуды колебаний  $|U_q|^2$  в (6.100) выражается через матричные элементы этих операторов (см. гл.3, приложение к §3.4, формулы (3.115) и (3.121)) и равен

$$|U_q|^2 = \begin{cases} \frac{\hbar N_q}{2M\omega_q}, & \text{для перехода } N_q \rightarrow N_q - 1, \\ \frac{\hbar(N_q + 1)}{2M\omega_q}, & \text{для перехода } N_q \rightarrow N_q + 1. \end{cases} \quad (6.101)$$

Здесь  $M = \rho V$ , где  $\rho$  — плотность вещества,  $V$  — нормировочный объем. Энергия осциллятора в  $N_q$ -м состоянии равна  $(N_q + 1/2)\hbar\omega_q$  и складывается из энергии  $N_q$  фононов. Поскольку таких осцилляторов в кристалле множество,  $N_q$  можно заменить средним числом фононов, которое, согласно формуле Планка, при температуре решетки  $T$  равно

$$N_q = \frac{1}{\exp\left(\frac{\hbar\omega_q}{k_B T}\right) - 1}. \quad (6.102)$$

В результате окончательное выражение для квадрата модуля матричного элемента можно представить в виде

$$|\mathbf{V}_{\mathbf{k},\mathbf{k}\pm\mathbf{q}}|^2 = E_0^2 q^2 \left( N_q + \frac{1}{2} \mp \frac{1}{2} \right) \frac{\hbar}{2\rho V \omega_q}. \quad (6.103)$$

Здесь верхний знак соответствует поглощению фонона, а нижний — испусканию. Входящая в (6.103) энергия акустического фонона  $\hbar\omega_q$  мала по сравнению с тепловой энергией  $k_B T$  (см. §6.5, пункт а). Поэтому выражение для  $N_q$  можно упростить, разлагая экспоненту в ряд по параметру  $\hbar\omega_q/k_B T \ll 1$ ,

$$N_q \approx \frac{k_B T}{\hbar\omega_q} \gg 1. \quad (6.104)$$

Таким образом, в этом приближении  $N_q + 1 \approx N_q$ . Тогда матричные элементы, описывающие поглощение и испускание фонона, совпадают и равны

$$|\mathbf{V}_{\mathbf{k},\mathbf{k}\pm\mathbf{q}}|^2 = E_0^2 q^2 \frac{k_B T}{2\rho V \omega_q^2}. \quad (6.105)$$

Для продольных акустических фононов при малых  $q$  частота  $\omega_q = S_l \cdot q$ , где  $S_l$  — продольная скорость звука, то есть скорость продольных акустических фононов. Продольная скорость звука связана с продольной упругой постоянной  $C_l$  соотношением

$$C_l = \rho S_l^2. \quad (6.106)$$

Окончательно, квадрат модуля матричного элемента равен

$$|\mathbf{V}_{\mathbf{k},\mathbf{k}\pm\mathbf{q}}|^2 = E_0^2 \frac{k_B T}{2V C_l}. \quad (6.107)$$

Здесь следует отметить, что в кубическом кристалле при распространении волны в направлении  $\langle 100 \rangle C_l = C_{11}$ , в направлении  $\langle 110 \rangle C_l = (C_{11} + C_{12} + C_{44})/2$  и в направлении  $\langle 111 \rangle C_l = (C_{11} + 2C_{12} + 4C_{44})/3$ . В других направлениях волны не являются строго продольными и имеют скорости, промежуточные по сравнению со случаями, соответствующими  $\langle 100 \rangle$  и  $\langle 111 \rangle$ ;  $C_{11}$ ,  $C_{12}$  и  $C_{44}$  — компоненты тензора упругости (см. главу 1).

Из (6.107) следует, что матричный элемент электрон-фононного взаимодействия (при рассеянии на акустических фононах) не зависит от энергии электрона и угла рассеяния. Поскольку его значения при испускании и поглощении фонона приблизительно одинаковы, мы можем учесть оба процесса, просто умножив на 2 вероятности рассеяния (6.84):

$$W_{\mathbf{p}\mathbf{p}'} = \frac{2\pi}{\hbar} |V_{\mathbf{p}\mathbf{p}'}|^2 \{ \delta(\varepsilon_{\mathbf{p}'} - \varepsilon_{\mathbf{p}} + \hbar\omega_{\mathbf{q}}) + \delta[\varepsilon_{\mathbf{p}'} - \varepsilon_{\mathbf{p}} - \hbar\omega_{\mathbf{q}}] \} \approx \\ \approx 2 \frac{2\pi}{\hbar} |V_{\mathbf{p}\mathbf{p}'}|^2 \delta(\varepsilon_{\mathbf{p}'} - \varepsilon_{\mathbf{p}}). \quad (6.108)$$

Зная вероятность, мы можем вычислить время релаксации. Согласно (6.37)

$$\frac{1}{\tau} = \sum_{\mathbf{p}} W_{\mathbf{p}\mathbf{p}'} (1 - \cos \theta). \quad (6.109)$$

Заменим суммирование по  $\mathbf{p}'$  интегралом по  $\mathbf{p}'$ -пространству в сферических координатах с полярной осью, направленной по  $\mathbf{p}$ :

$$\sum_{\mathbf{p}'} \rightarrow \frac{V}{(2\pi\hbar)^3} \int p'^2 dp' \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\varphi. \quad (6.110)$$

В результате для времени релаксации с учетом (6.108), (6.109) и (6.110) получаем

$$\frac{1}{\tau} = \frac{V}{\pi\hbar} \int k'^2 dk' |V_{\mathbf{k}\mathbf{k}'}|^2 \delta(\varepsilon_{\mathbf{k}'} - \varepsilon_{\mathbf{k}}) \int_0^\pi (1 - \cos \theta) \sin \theta d\theta. \quad (6.111)$$

В (6.111) мы сразу выполнили интегрирование по  $\varphi$  (оно дает  $2\pi$ ). Интеграл по  $\theta$  вычисляется легко и равен 2. Удобно перейти в (6.111) от интегрирования по  $k'$  к интегрированию по  $\varepsilon_{\mathbf{k}'}$ . Выполняя замену переменных

$$k'^2 dk' = \frac{m^2 v'}{\hbar^3} d\varepsilon_{\mathbf{k}'}, \quad (6.112)$$

где  $v' = \left(\frac{2\varepsilon_{\mathbf{k}'}}{m}\right)^{1/2}$ , получаем

$$\frac{1}{\tau} = \frac{2V}{\pi\hbar} \frac{m^2}{\hbar^3} \langle v \rangle \frac{E_0^2 k_B T}{2VC_l} = \frac{E_0^2 m^2 \langle v \rangle k_B T}{\pi \hbar^4 C_l}. \quad (6.113)$$

Здесь  $\langle v \rangle$  — средняя тепловая скорость электрона.

Так как средняя тепловая скорость электрона равна  $\langle v \rangle \propto T^{1/2}$ , то

$$1/\tau \propto T^{3/2}. \quad (6.114)$$

Итак, мы подтвердили результат Я. И. Френкеля для зависимости времени релаксации от температуры при рассеянии электронов на акустических фононах.

## 6.6. Рассеяние на оптических фононах

Рассмотрим пока качественно, что нужно понимать под рассеянием электронов на оптических фононах. На рис. 6.15 представлен спектр оптических фононов. Для оптических фононов так же, как и для акустических фононов, справедливо утверждение, что импульс оптического фонона (взаимодействующего с электроном) порядка импульса электрона и этот импульс очень мал (много меньше  $k_0$  (см. рис. 6.15)).

Поэтому нас интересуют оптические фононы с очень малыми волновыми векторами  $q \ll k_0$ . В отличие от акустических фононов, для которых их энергия при данном значении  $q \ll k_0$  мала по сравнению с энергией электрона  $\varepsilon$ , для оптических фононов их энергия  $\hbar\omega_0$  велика (см. рис. 6.15). Энергия оптического фонона (взаимодействующего с электроном) не зависит от энергии электрона и слабо зависит от  $q$ . Поэтому при взаимодействии электронов с оптическими фононами электрон излучает или поглощает оптический фонон с энергией  $\hbar\omega_0$ . Как правило, электрон сильнее взаимодействует с оптическими колебаниями решетки, чем с акустическими. Причем различают взаимодействие с оптическими фононами: полярное и деформационное взаимодействие.

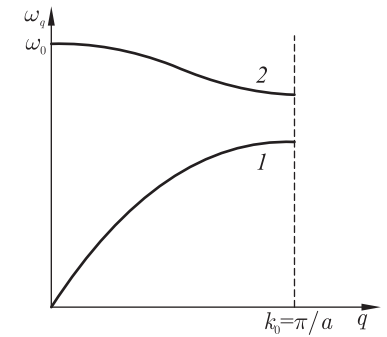


Рис. 6.15. Закон дисперсии для акустических (кривая 1) и оптических (кривая 2) фононов для двухатомной линейной решетки. Показано предельное значение частоты оптического фонона при  $q = 0$

Рассмотрим в качестве примера полупроводник GaAs с ионной связью. Когда происходят оптические колебания, то при этом в колебаниях участвуют два иона (рис. 6.16). Это соответствует колебанию дипольного момента. Такие колебания эффективно взаимодействуют с электронами (дырками), так как они — заряженные частицы.



Рис. 6.16. Образование электрического дипольного момента при колебании ионов As и Ga

В Ge и Si ситуация принципиально иная. Атомы кристаллической решетки не заряжены, и полярный механизм электрон-фононного взаимодействия не работает.

При температуре ниже азотной можно не учитывать взаимодействие электронов с оптическими фононами при вычислении подвижности. Если температура низкая, так что  $k_B T \ll \hbar\omega_0$ , то оптических фононов в кристалле нет. При таких температурах излучение электроном оптических фононов запрещено законом сохранения энергии. Если же средняя энергия электрона  $k_B T < \hbar\omega_0$ , то оптических фононов мало, но взаимодействие электронов с оптическими фононами велико. В силу того, что при этих условиях функция распределения электронов по энергии подчиняется распределению Максвелла ( $f \propto \exp(-\varepsilon/k_B T)$ ),



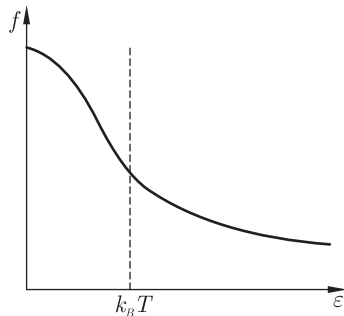


Рис. 6.17. Схематический вид функции распределения Максвелла-Больцмана при заданной температуре

рис. 6.17), всегда имеются электроны (на хвосте максвелловского распределения), которые существенно взаимодействуют с оптическими фононами (область  $\varepsilon > k_B T$  на рис. 6.17), но если  $\hbar\omega_0 \gg k_B T$ , то это взаимодействие почти отсутствует. Как правило, при комнатной температуре взаимодействие электронов с оптическими фононами в полупроводниках уже существенно.

**Замечание.** Другой режим рассеяния, имеющий особое значение для полупроводников Ge и Si, зона проводимости которых содержит более одного минимума энергии, — это так называемое междолинное рассеяние, связанное

с перебросами электрона из одного минимума в другой (рис. 6.18). В долинах «1» и «2» имеются электроны, причем из долины «2» электрон может перейти в долину «1», то есть электрон переходит из одного квантового состояния с энергией  $E_{n\mathbf{k}}$  и волновым вектором  $\mathbf{k}$  порядка  $\mathbf{k}_0$  в состояние с энергией  $E_{n'\mathbf{k}'}$  и волновым вектором  $\mathbf{k}'$ . При таком переходе электрон излучает (поглощает) фонон с импульсом, близким к предельному, то есть с импульсом порядка вектора обратной решетки  $2\pi/a$  (это связано с тем, что дно зоны расположено вблизи границы зоны Бриллюэна). И акустические, и оптические фононы с предельным импульсом имеют большую энергию (рис. 6.15), которую иногда называют энергией междолинного рассеяния. Такое рассеяние должно быть весьма неупругим, а потому и маловероятным при низких температурах из-за малости как числа фононов с большой энергией, которые могли бы участвовать в рассеянии, так и числа электронов,

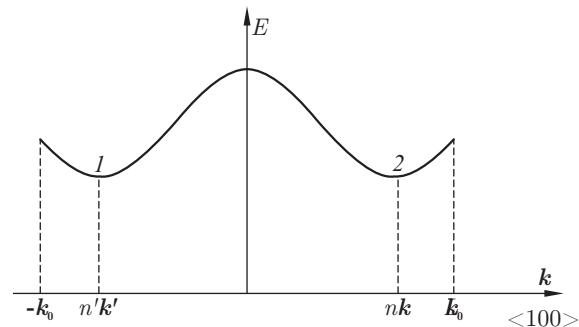


Рис. 6.18. Зонная структура полупроводника с двумя минимумами для электронов. Показана схема междолинного рассеяния электронов в таких полупроводниках

способных испускать такие быстрые фононы. Здесь ситуация такая же, как и в случае с оптическими фононами. С ростом температуры эффективность междолинного рассеяния возрастает, также как в случае рассеяния на оптических фононах. Но энергия фонона, требуемая для междолинного рассеяния, как правило меньше, чем энергия оптического фонона. Поэтому при увеличении температуры междолинное рассеяние наступает раньше, чем рассеяние на оптических фононах.

## 6.7. Температурная зависимость подвижности

Результаты предыдущих параграфов позволяют установить температурную зависимость подвижности носителей заряда  $b(T)$ . Температурная зависимость подвижности  $b$  такая же, как и температурная зависимость времени релаксации  $\tau$ . Для любого конкретного полупроводника определение вкладов различных механизмов рассеяния в разных областях температур в величину подвижности является очень сложной задачей. Относительные вклады различных механизмов рассеяния в сильной степени изменяются от одного полупроводника к другому, так что невозможно дать сколько-нибудь общий результат для всех полупроводников. Однако можно сказать, что колебания решетки различных типов дают основной вклад в рассеяние в области высоких температур. При низких температурах  $k_B T \ll \hbar\omega_0$  оптических фононов практически нет, имеются только акустические фононы и заряженные центры. Тогда, как уже было показано выше, при рассеянии на заряженных центрах  $b \propto T^{3/2}$ . С ростом температуры рассеяние на заряженных центрах становится все менее эффективным, а рассеяние на колебаниях решетки, наоборот, становится все большим. При рассеянии только на акустических колебаниях решетки  $b \propto T^{-3/2}$ . Но при температурах порядка комнатной становится существенным рассеяние на оптических колебаниях решетки, и закон  $T^{-3/2}$  для подвижности уже не выполняется. Тогда подвижность будет падать с ростом температуры быстрее, чем по закону  $T^{-3/2}$ . Экспериментально, закон  $b \propto T^{-3/2}$  выполняется в Ge. Для дырок в Si и дырок в Ge подвижность убывает с ростом температуры быстрее,  $b \propto T^{-2 \div 2.5}$ . Если при какой-то температуре действуют несколько механизмов рассеяния, тогда нужно складывать частоты столкновений, соответствующие этим механизмам. Как правило, доминирует тот механизм рассеяния, у которого время релаксации меньше. Эффективная частота столкновений тогда равна

$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_2}. \quad (6.115)$$

Если вычисленная подвижность больше экспериментально измеренной, то это означает, что не учтен еще какой-то механизм рассеяния.

На рис. 6.19 представлен общий характер поведения подвижности  $b$  как функции температуры  $T$ . Это иллюстрация экспериментальной



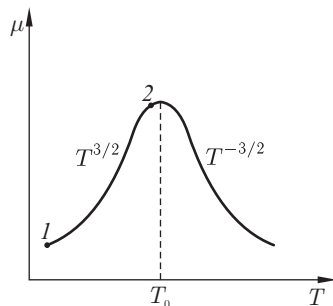


Рис. 6.19. Температурная зависимость подвижности электронов в полупроводнике при действии различных механизмов рассеяния. Слева от максимума (от точки 1 до точки 2) доминирует рассеяние на заряженных примесях; справа от максимума подвижность контролируется рассеянием на фононах

зависимости. От точки 1 до точки 2 подвижность контролируется рассеянием на заряженных примесях, от точки 2 и далее подвижность определяется рассеянием на фононах. Для полупроводника GaAs  $T_0 \approx 76$  К при концентрации примесей порядка  $10^{16}$ – $10^{17}$  см<sup>3</sup>. Важно отметить, что в области температур от точки 1 до точки 2 подвижность является характеристикой не кристалла, а его чистоты, то есть указывает на количество примесей в кристалле. В области температур от точки 2 и далее подвижность является характеристикой кристалла, так как она определяется рассеянием на фононах. Обычно подвижность при высоких температурах, когда рассеяние определяется фононами, заносят в справочник, так как она является характеристикой данного материала. Если у нас изначально был чистый кристалл и мы измерили подвижность и получили обычную зависимость от температуры (рис. 6.20, кривая 1), то при легировании кристалла время релаксации падает и, следовательно, падает подвижность (рис. 6.20, кривая 2). Но при высоких температурах кривая для  $b(T)$  легированного полупроводника (кривая 2, рис. 6.20) совпадает с подвижностью чистого материала, то есть с зависимостью  $b(T)$  для чистого кристалла. Подвижность — это очень важная характеристика кристалла, поэтому для каждого кристалла ее приводят в таблицах (табл. 6.1)<sup>1)</sup>. Обычно на практике судят о чистоте материала полупроводника по зависимости подвижности от температуры. Для чистого полупроводника подвижность убывает с ростом температуры (рис. 6.21, кривая 1). Однако с ростом количества примесей кривая 1 на рис. 6.21 резко изменяется (кривые 2, 3, 4, 5).

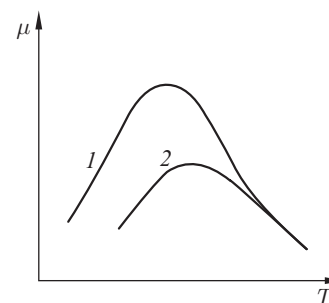


Рис. 6.20. Подвижность электронов в полупроводнике с различным количеством примесей. Кривая 1 соответствует нелегированному полупроводнику; кривая 2 — легированному

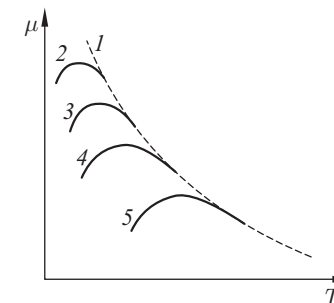


Рис. 6.21. Зависимость подвижности электронов от температуры в полупроводнике, содержащем примеси различной концентрации. Кривые пронумерованы в порядке возрастания концентрации примесей. При низких температурах подвижность ограничивается степенью чистоты кристалла (количеством примеси), а при высоких — рассеянием на оптических фононах. Пунктирная кривая 1 соответствует температурной зависимости подвижности для чистого (без примесей) полупроводника

Таблица 6.1. Подвижность носителей заряда в некоторых полупроводниках

Полупроводник	$b_n, \text{см}^2/(\text{В} \cdot \text{с}) \cdot 10^3$		$b_p, \text{см}^2/(\text{В} \cdot \text{с}) \cdot 10^3$		$b = b_n/b_p$	
	300 К	77 К	300 К	77 К	300 К	77 К
Ge	3,8	37,1	1,82	43,7	2,1	0,7
Si	1,3	45,5	0,5	11,6	2,6	3,9
InSb	78	1200	0,75	10,0	100	120
InAs	33	82	0,46	0,69	70	120
InP	4,6	24	0,15	1,2	30	20
GaSb	4,0	6	1,4	3,6	3	1,7
GaAs	8,5	21	0,42	4,2	20	5
GaP	0,11	0,5	0,075	0,42	1,5	1
AlSb	0,2	—	0,42	3,7	—	—

**Замечание.** Если два механизма рассеяния влияют друг на друга, то соотношение<sup>1)</sup>

$$\frac{1}{\tau} = \frac{1}{\tau_1} + \frac{1}{\tau_2} \quad (6.116)$$

выполняется лишь приближенно.

<sup>1)</sup> База данных: <http://www.ioffe.ru/SVA/NSM>

<sup>1)</sup> Соотношение (6.116) иногда называют правилом Матиссена.

### 6.8. Горячие электроны

При исследовании подвижности электронов мы фактически интересовались тем, как изменяется импульс электрона при столкновениях, то есть мы интересовались проекцией вектора импульса  $\mathbf{p}'$  после столкновения на первоначальное направление. Мы видели, что при рассеянии на примесях электрон практически не теряет энергию, а только изменяет свой импульс. Если кристалл помещен во внешнее электрическое поле, то электрон в таком поле набирает энергию и, рассеиваясь на примеси, не теряет ее. Получается, что энергия электрона непрерывно увеличивается, и в этом случае нет стационарного состояния. Другая ситуация возникает, например, при рассеянии на акустических фоновых. В этом случае в каждом акте рассеяния электрон теряет энергию, но эта потеря (и при поглощении и при испускании) мала, порядка

$$\Delta\varepsilon \sim \sqrt{ms^2\varepsilon} \ll \varepsilon, \quad (6.117)$$

где  $m$  — масса электрона,  $s$  — скорость звука в кристалле,  $\varepsilon$  — энергия электрона (см. § 6.5). Малая потеря энергии при рассеянии компенсирует прирост энергии, набираемой электроном на длине свободного пробега. Это означает, что при этих условиях существует некоторая стационарная энергия электрона. Эта энергия, очевидно, зависит от напряженности электрического поля. Оказывается, что уже при умеренных электрических полях стационарная энергия электрона начинает зависеть от напряженности электрического поля и происходит разогрев электронов. В этом случае энергия, приобретенная от поля, не полностью компенсируется рассеянием электронов на фоновых. В этих условиях температура электронов  $T_e$  больше, чем температура решетки  $T$ . Если средняя энергия электронов существенно возрастает под действием электрического поля, то это оказывает влияние на подвижность электронов и поэтому плотность тока уже не будет пропорциональна величине поля  $E$ . Подвижность тогда следует рассматривать как функцию электрического поля. Если плотность тока не пропорциональна электрическому полю  $E$ , то закон Ома нарушается. В слабых электрических полях плотность тока, согласно закону Ома, имеет вид

$$\mathbf{j} = enb\mathbf{E} = en\mathbf{v}_{др}, \quad (6.118)$$

где  $n$  — концентрация электронов;  $b$  — подвижность;  $\mathbf{v}_{др}$  — дрейфовая скорость электрона. Однако в умеренных сильных электрических полях (с учетом эффекта разогрева электронов) имеет место отклонение от закона Ома (сублинейная характеристика (рис. 6.22)).

Рассмотрим качественно механизм нелинейной зависимости тока от поля. В очень высоких электрических полях, когда энергия электрона  $\varepsilon$  оказывается равной ширине запрещенной зоны полупроводника  $E_g$ , становится важной ударная ионизация, и концентрация носителей резко возрастает. Это может приводить к нелинейной зависимости  $j(E)$ .

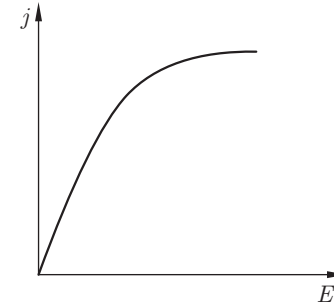


Рис. 6.22. Сублинейная зависимость плотности тока носителей заряда от напряженности электрического поля

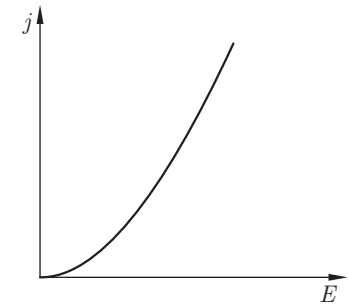


Рис. 6.23. Суперлинейная зависимость плотности тока от напряженности электрического поля

В меньших полях нелинейность связана с полевой зависимостью времени рассеяния  $\tau$ . При этом концентрацию носителей можно считать постоянной.

Если зависимость тока  $j$  от поля  $E$  сублинейная, то время релаксации убывает с ростом поля. Если же зависимость тока  $j$  от поля  $E$  суперлинейная (рис. 6.23), то время релаксации растет с ростом поля  $E$  (то есть в этом случае процесс рассеяния носителей заряда усиливается). При суперлинейной характеристике  $j(E)$  электроны разогреваются, время релаксации удлиняется и рассеяние на заряженных примесях замедляется, так как  $\tau \propto \varepsilon^{3/2}$ , где характерная энергия  $\varepsilon$  пропорциональна температуре электронного газа  $T_e$ , и при больших полях  $\varepsilon$  начинает зависеть от поля  $E$ . Итак, с ростом средней энергии электрона время релаксации  $\tau$  растет и сечение рассеяния убывает. Суперлинейная характеристика наблюдается экспериментально при низких температурах, когда электроны рассеиваются на заряженных примесях и время релаксации  $\tau \propto \varepsilon^{3/2}$ .

Сублинейная характеристика  $j(E)$  наблюдается при высоких температурах, когда электроны рассеиваются на фоновых. Время релаксации при рассеянии на фоновых убывает с ростом средней энергии электрона (то есть с ростом температуры электронов  $T_e$ ):  $\tau \propto T^{-1}v^{-1}$ . Здесь  $T$  — температура решетки (она входит через число фононов),  $v$  — скорость беспорядочного движения электрона. В случае, когда электрическое поле является сильным, скорость  $v$  пропорциональна корню квадратному из температуры электронов  $T_e$  (то есть из средней энергии электронов):  $v \propto (k_B T_e)^{1/2} = (\varepsilon)^{1/2}$ . Таким образом, при разогреве электронного газа средняя энергия электрона увеличивается и время релаксации  $\tau$  убывает.

Рассмотрим более подробно случай взаимодействия электрона с акустическими фоновыми в присутствии сильного электрического поля. Фактически, нам нужно найти закон зависимости электрического

тока  $j$  от поля  $E$  в сильных полях. Для этого полезно ввести такую характеристику, как время релаксации энергии  $\tau_\varepsilon$ . Время энергетической релаксации  $\tau_\varepsilon$  определяется из закона экспоненциального убывания превышения температуры (энергии) электронного газа над температурой решетки:  $\frac{d(T_e - T)}{dt} = -\frac{1}{\tau_\varepsilon}(T_e - T)$ . (Напомним, что под величиной  $\tau$  мы понимаем время релаксации импульса.) Дадим качественную оценку времени энергетической релаксации. Если единственным механизмом рассеяния носителей является рассеяние на заряженных примесях, то  $\tau_\varepsilon$  бесконечно велико, так как такие столкновения упругие. Однако при рассеянии на фонах время релаксации энергии  $\tau_\varepsilon$  конечно, но  $\tau_\varepsilon$  все же больше, чем время релаксации импульса  $\tau$  (поскольку импульс электрона, как мы видели в § 6.5, релаксирует быстрее, чем его энергия).

Итак, как уже было показано, возможны три процесса рассеяния на фонах: а) индуцированное излучение; его вероятность пропорциональна числу фононов  $N_q W_{\mathbf{p}\mathbf{p}'}^0$ ; б) поглощение фононов; вероятность поглощения также пропорциональна числу фононов  $N_q W_{\mathbf{p}\mathbf{p}'}^0$ ; в) спонтанное излучение фононов; его вероятность есть просто  $W_{\mathbf{p}\mathbf{p}'}^0$ . Напомним, что число фононов, эффективно взаимодействующих с электронами,  $N_q \approx \frac{k_B T}{\hbar \omega_q} \gg 1$ . Поэтому процесс поглощения и индуцированного излучения фононов гораздо сильнее, чем процесс спонтанного излучения (в силу того, что фононов много). Однако индуцированное излучение и поглощение фононов компенсируют друг друга, поскольку их вероятности одинаковы, а передача энергии излучению остается. Такое соотношение между тремя процессами рассеяния электронов на фонах сохранится, пока мы считаем фононы равновесными и их число не меняется.

Рассмотрим скорость изменения энергии электрона в единицу времени (в с) за счет трех процессов рассеяния:

$$\frac{d\varepsilon}{dt} = -N_q W_{\mathbf{p}\mathbf{p}'}^0 \hbar \omega_q + N_q W_{\mathbf{p}\mathbf{p}'}^0 \hbar \omega_q - W_{\mathbf{p}\mathbf{p}'}^0 \hbar \omega_q. \quad (6.119)$$

Первое слагаемое описывает уменьшение энергии электрона за счет индуцированного излучения фонона; второе слагаемое описывает поглощение энергии, которая приобретается электроном; третье слагаемое описывает изменение энергии за счет спонтанного излучения фонона. В общем виде в (6.119) следовало бы суммировать по всем волновым векторам фононов  $q$ . Мы же для анализа скорости  $d\varepsilon/dt$  воспользуемся упрощенным вариантом. Первые два члена в (6.119) при не слишком сильном разогреве почти компенсируют друг друга и главную роль

в (6.119) играет спонтанное излучение; тогда

$$\frac{d\varepsilon}{dt} = -W_{\mathbf{p}\mathbf{p}'}^0 \hbar \omega_q. \quad (6.120)$$

Вероятность элементарного акта  $W_{\mathbf{p}\mathbf{p}'}^0$  удобно выразить через время импульсной релаксации  $\tau$ . Скорость изменения дрейфового импульса электрона в единицу времени за счет трех процессов рассеяния:

$$\frac{dP_{\text{др},x}}{dt} = -N_q W_{\mathbf{p}\mathbf{p}'} q_x - N_q W_{\mathbf{p}\mathbf{p}'} q_x - W_{\mathbf{p}\mathbf{p}'} q_x \approx \frac{P_{\text{др},x}}{\tau}, \quad (6.121)$$

где ось  $x$  направлена вдоль импульса  $p$ . Здесь первое слагаемое описывает уменьшение импульса электрона за счет индуцированного излучения фонона; второе слагаемое описывает уменьшение импульса за счет поглощения фонона; третье слагаемое — это уменьшение импульса электрона за счет спонтанного излучения. Отметим, что все три процесса рассеяния изменяют импульс электрона. Третьим слагаемым в (6.121) можно пренебречь, так как  $N_q \gg 1$ . В (6.121) мы также учли, что импульс фонона  $q \approx P_{\text{др}}$ . С учетом сказанного (6.121) можно переписать в виде

$$\frac{dP_{\text{др},x}}{dt} \approx -N_{q,x} W_{\mathbf{p}\mathbf{p}'} q_x - N_q W_{\mathbf{p}\mathbf{p}'} q_x = -2N_q W_{\mathbf{p}\mathbf{p}'} q_x P_{\text{др},x} \equiv -\frac{P_{\text{др},x}}{\tau}. \quad (6.122)$$

Таким образом, мы выразим вероятность  $W_{\mathbf{p}\mathbf{p}'}$  через время релаксации импульса  $\tau$ :

$$W_{\mathbf{p}\mathbf{p}'} = \frac{1}{\tau N_q}. \quad (6.123)$$

В (6.123) мы пренебрегли множителем  $1/2$ , так как его учет означает превышение точности.

Подставим выражение (6.123) для  $W_{\mathbf{p}\mathbf{p}'}$  в уравнение (6.120) и учтем, что  $N_q = k_B T / \hbar \omega_q$ , тогда получим

$$\frac{d\varepsilon}{dt} = -\frac{1}{\tau} \frac{(\hbar \omega_q)^2}{k_B T}. \quad (6.124)$$

В § 8.5 мы показали, что энергия акустического фонона (взаимодействующего с электроном) зависит от энергии электрона и порядка:

$$\hbar \omega_q \sim (\varepsilon m s^2)^{1/2}. \quad (6.125)$$

Тогда для скорости изменения энергии электрона в единицу времени получаем окончательное выражение

$$\frac{d\varepsilon}{dt} = -\frac{1}{\tau} \frac{(m s^2)}{k_B T} \varepsilon = -\frac{\varepsilon}{\tau_\varepsilon}. \quad (6.126)$$

Время энергетической релаксации  $\tau_\varepsilon$  определим как

$$\tau_\varepsilon = \frac{k_B T}{2 m s^2} \tau. \quad (6.127)$$

Выражение (6.127) для времени релаксации энергии точное, поэтому мы включили в него множитель  $1/2$ . Из (6.127) следует важный вывод, что время релаксации по энергии  $\tau_\epsilon$  гораздо больше времени релаксации импульса  $\tau$ , так как  $kT \gg m s^2$ .

Отметим, что при рассеянии электронов на оптических фононах время релаксации по энергии  $\tau_\epsilon$  почти такое, как и время релаксации импульса  $\tau$ . Выражение (6.127) для  $\tau_\epsilon$  можно переписать в виде

$$\tau_\epsilon = \alpha \tau, \quad (6.128)$$

где  $\alpha$  — параметр неупругости (как правило  $\alpha \gg 1$ ). Физический смысл  $\alpha$ : эта величина показывает, сколько нужно электрону испытать столкновений, чтобы растратить всю свою энергию.

Найдем теперь энергию электрона во внешнем электрическом поле  $E$  при разогреве. Как и прежде, мы не будем проводить точный расчет этой энергии, а дадим качественную оценку ее величины. Рассмотрим стационарное состояние электрона. В стационарном состоянии энергия, приобретенная электроном на длине свободного пробега, должна в точности компенсироваться потерей энергии при рассеянии:

$$eE v_{др} = \frac{\epsilon - k_B T}{\tau_\epsilon}. \quad (6.129)$$

Здесь  $eE v_{др}$  — это энергия, приобретаемая от поля в единицу времени, а  $(\epsilon - k_B T)/\tau_\epsilon$  — скорость потери энергии:

$$\frac{d\epsilon}{dt} = -\frac{\epsilon - k_B T}{\tau_\epsilon}. \quad (6.130)$$

Выразим в (6.129) дрейфовую скорость через подвижность  $\mu$  и напряженность электрического поля  $E$ :  $v_{др} = bE = \frac{e\tau}{m}E$ . Тогда для энергии электрона получаем

$$\epsilon = \frac{e^2}{m} E^2 \tau \cdot \tau_\epsilon + k_B T. \quad (6.131)$$

Введем электронную температуру  $T_e$ :

$$\epsilon = k_B T_e. \quad (6.132)$$

Тогда для  $T_e$  получаем

$$k_B T_e = k_B T + m(bE)^2 \alpha. \quad (6.133)$$

Напомним, что  $\alpha = \tau_\epsilon/\tau$  — параметр неупругости. Из (6.133) следует, что если внешнее электрическое поле слабое, то электронная температура  $T_e$  близка к решеточной  $T$ . Если же поле сильное, то  $T_e$  больше  $T$ . Итак, мы получили (согласно (6.133)), что задолго до того, как произойдет уравнивание дрейфовой скорости  $v_{др}$  с тепловой  $v$ , произойдет разогрев электронного газа, так как дрейфовая энергия в (6.133) умножается на большое число  $\alpha$ . Рассмотрим условия существенного

разогрева, когда второе слагаемое в (6.133) становится много больше первого:

$$(bE)^2 \alpha \gg \frac{k_B T}{m}. \quad (6.134)$$

В этом случае температура электронного газа равна

$$k_B T_e = m(bE)^2 \alpha. \quad (6.135)$$

Однако в условиях сильного разогрева подвижность электронов  $b$  зависит от  $T_e$ . При рассеянии электронов на акустических фононах  $b \propto v^{-1} \propto T_e^{-1/2}$ . Тогда из (6.135) следует, что электронная температура пропорциональна напряженности электрического поля:  $T_e \propto E$ . Таким образом, подвижность электронов убывает с ростом поля по закону

$$b \propto \frac{1}{T_e^{1/2}} \propto \frac{1}{E^{1/2}}. \quad (6.136)$$

Окончательно мы получили, что при рассеянии электронов на акустических фононах плотность тока (согласно (6.118)) зависит от напряженности электрического поля нелинейно:

$$j \propto (E)^{1/2}. \quad (6.137)$$

На рис. 6.24 представлена теоретическая зависимость плотности тока от напряженности электрического поля согласно (6.137). При малых полях выполняется закон Ома; при больших полях плотность тока линейно зависит от  $E$  (закон Ома); при больших полях имеем участок нелинейной зависимости плотности тока  $j$  от поля  $E$ . Следует отметить, что такая зависимость тока  $j$  от  $E$  часто наблюдается экспериментально.

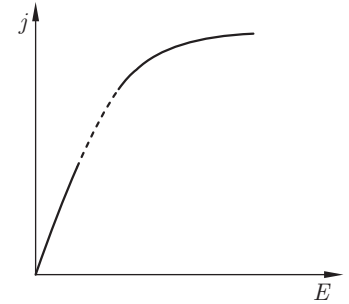


Рис. 6.24. Нелинейная зависимость плотности тока от напряженности электрического поля согласно (6.137). При малых полях выполняется закон Ома; при больших полях  $j$  зависит от  $E$  нелинейно

### Некоторые замечания

**Замечание 1. О возможных формах нелинейных ВАХ.** Экспериментально наблюдаемые зависимости тока от напряжения (вольт-амперные характеристики, или ВАХ) могут приобретать достаточно сложную нелинейную форму. Часто в основе этого — нелинейность локальных законов переноса заряда в твердых телах. Далее будем рассматривать нелинейные зависимости плотности тока  $j$  от локального электрического поля  $E$ . По аналогии с «внешними» зависимостями  $I(U)$  (ток–напряжение) их называют дифференциальными (или локальными) вольт-амперными характеристиками. Важным понятием, характеризующим локальную ВАХ, является дифференциальная проводимость

$$\sigma_d = \frac{dj}{dE}. \quad (6.138)$$

В простейшем случае плотность тока  $j$  растет с ростом поля  $E$ ; при этом  $\sigma_d > 0$ . Как уже было отмечено выше, зависимость  $j(E)$  может быть сублинейной (рис. 6.24). Для такой формы ВАХ  $\sigma_d$  уменьшается с ростом поля  $E$ . Возможен также случай, когда  $\sigma_d$  увеличивается с ростом поля; это соответствует суперлинейной зависимости  $j(E)$ .

Более сложными являются немонотонные зависимости  $j(E)$   $N$ - и  $S$ -типов. Рассмотрим сначала характеристику  $N$ -типа (рис. 6.25). Она содержит два нарастающих участка с  $\sigma_d > 0$ : один при малых значениях поля  $E$ , другой — при больших полях. Имеется также и падающий участок с отрицательной дифференциальной проводимостью  $\sigma_d < 0$  (ОДП) для промежуточных полей  $E$ .

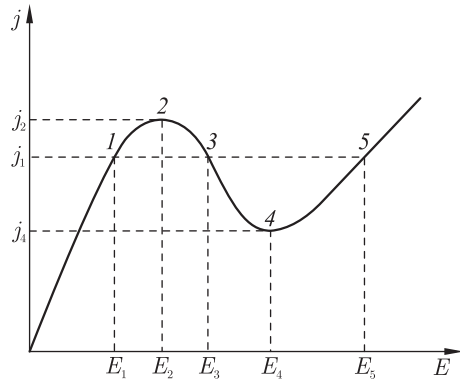


Рис. 6.25. Вольтамперная характеристика  $N$ -типа

Существенная особенность  $N$ -образной характеристики на рис. 6.25 состоит в том, что в определенной области значений токов  $j_4 < j < j_2$  напряженность поля есть многозначная функция плотности тока, а плотность тока является однозначной функцией поля. Чтобы пройти по этой кривой, всегда нужно поддерживать (то есть задавать) напряжение, а ток будет такой, какой установится. В интервале  $j_4 < j < j_2$  при заданном значении плотности тока  $j = j_1$  система может находиться в одном из трех пространственно однородных состояний. В двух из них (при  $E = E_1$  и при  $E = E_5$ ) дифференциальная проводимость положительна; в третьем (при  $E = E_3$ ) — дифференциальная проводимость отрицательна. Очевидно, что дифференциальная проводимость меняет знак, проходя через нуль в точках  $(E = E_2, j = j_2)$  и  $(E = E_4, j = j_4)$  (рис. 6.25).

Другим видом немонотонной зависимости тока от поля  $j(E)$  является  $S$ -образная характеристика, представленная на рис. 6.26.  $S$ -образная характеристика также имеет два участка, на которых  $\sigma_d > 0$ , и один участок, где  $\sigma_d < 0$  (участок 1–2 на рис. 6.26). Однако, для  $S$ -образной характеристики в некотором интервале значений поля  $E$  плотность

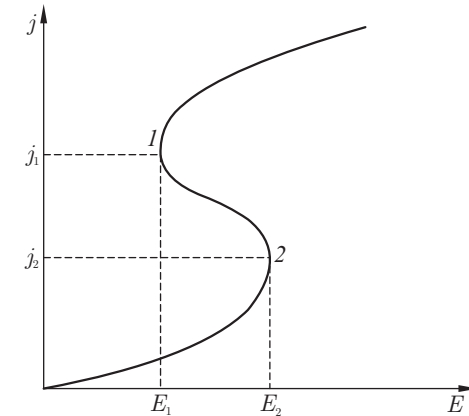


Рис. 6.26. Вольтамперная характеристика  $S$ -типа

тока  $j$  есть неоднозначная функция поля  $E$ , а дифференциальная проводимость  $\partial j / \partial E$  меняет знак, обращаясь в бесконечность в точках  $(E = E_2, j = j_2)$  и  $(E = E_1, j = j_1)$  (рис. 6.26). Чтобы пройти по  $S$ -образной характеристике, нужно задавать ток, а поле получится такое, какое установится.

Реально возможны также комбинации ВАХ  $S$ - и  $N$ -типа; на статической характеристике они могут сменять друг друга при возрастании электрического поля<sup>1)</sup>. Характеристика может с течением времени превращаться из  $S$ -образной в  $N$ -образную. На практике статическая характеристика может иметь и более сложную форму, с многозначностью тока и напряжения.

Весьма интересные последствия, связанные с неоднозначностью локальных ВАХ вещества, проявляются в реальных образцах и приборах на их основе с конечными геометрическими размерами.

Так, в образце с локальной ВАХ  $N$ -типа, имеющего большой размер по направлению протекающего тока, в объеме могут сосуществовать области (домены) с большим и малым полем (см. главу 7). Такие неоднородные состояния с доменами могут быть реализованы как за счет внешнего воздействия, так и из-за фундаментальных свойств неустойчивости распределенных систем с ОДП относительно флуктуаций.

Аналогично этому в образце с локальной ВАХ  $S$ -типа, имеющем большие размеры по направлению, перпендикулярному  $\mathbf{E}$ , возможно сосуществование целых областей с большой и малой плотностью тока (глава 7). Области с большим током  $j$  называют шнурами тока. Подобно доменам поля шнур тока может возникнуть как за счет управляемого внешнего воздействия, так и из-за неустойчивости однородных систем с ОДП к малым флуктуациям.

<sup>1)</sup> Шель Э. Самоорганизация в полупроводниках. — М.: Мир, 1991. 460 с.

**Замечание 2. О практическом применении.** Возникновение или уничтожение доменов поля и шнуров тока сопровождается большими скачками напряжения и тока во внешних электрических цепях. Эти скачки вредны для выполняемых приборами функций, а понимание соответствующих механизмов помогает с ними бороться на практике. Однако особо интересны полезные применения эффектов доменизации и шнурования. Так, периодически повторяющийся пролет доменов в пространстве между катодом и анодом в устройстве с  $N$ -образной характеристикой широко используется для генерации СВЧ-колебаний.

В системах с  $S$ -образной характеристикой можно возбуждать движение фронтов переключения тока, напоминающее распространение нервных импульсов. В последние годы ведутся активные исследования по созданию на этой основе аналоговых процессоров для искусственных нейронных сетей.

### Задачи

**Задача 6.1.** Есть две долины, одна ориентирована вдоль оси  $x$ , другая — вдоль оси  $y$ . К кристаллу приложена одноосная деформация вдоль оси  $x$ . Найти отношение концентраций в долинах в зависимости от величины относительного сдвига долин  $\Delta E$ . Оценить значение относительного сдвига долин, при котором относительная концентрация в долинах изменится в 2 раза.

**Решение.** Уровень Ферми для обеих зон одинаков. Концентрация электронов в первой долине равна:  $n_1 = N_c \exp\left(-\frac{\chi_{c1}}{k_B T}\right)$ , где  $\chi_{c1}$  — расстояние по энергии от уровня Ферми до дна зоны проводимости первой долины. Во второй долине:  $n_2 = N_c \exp\left(-\frac{\chi_{c2}}{k_B T}\right)$ , где  $\chi_{c2}$  — расстояние от уровня Ферми до дна зоны проводимости второй долины. Поскольку при деформации первая долина сместится вниз по энергии на величину  $\Delta E$ , то

$$\chi_{c2} = \chi_{c1} + \Delta E.$$

Следовательно,

$$\frac{n_1}{n_2} = \left(\frac{m_{c1}}{m_{c2}}\right) \exp\left(\frac{\Delta E}{k_B T}\right).$$

Если относительная концентрация в долинах изменилась в 2 раза, то сдвиг энергии при этом равен:  $\Delta \varepsilon = k_B T \ln 2$ .

**Задача 6.2.** Предполагая, что относительный сдвиг долин при одноосной деформации  $\Delta E$  пропорционален величине деформации  $\varepsilon$ ,  $\Delta E = \alpha \varepsilon$ , где  $\alpha$  — коэффициент, имеющий размерность энергии, и  $\alpha \sim 1\text{--}10$  эВ. Найти зависимость относительного изменения удельного сопротивления  $\Delta \rho / \rho$  от  $\varepsilon$  для Si. Нарисовать схематически график

зависимости  $\Delta \rho / \rho_0$  от  $\varepsilon$ . Давление и электрическое поле приложены по оси  $\langle 100 \rangle$ .

**Решение.** Чтобы найти  $\Delta \rho / \rho_0$  за счет деформации, нужно найти проводимость  $\sigma$ , так как  $\rho = 1/\sigma$ . Поскольку и давление, и электрическое поле приложены по оси  $\langle 100 \rangle$ , то для Si согласно (6.20), ток равен:  $j_x = \frac{e}{3}(n_1 v_{1x} + n_2 v_{2x} + n_3 v_{3x})$ . Так как вторая и третья долины расположены по осям  $y$  и  $z$ , то в них  $n_2 = n_3$ . Согласно предыдущей задаче

$$n_1 = n_2 \exp\left(\frac{\Delta E}{k_B T}\right) = n_2 \exp\left(\frac{\alpha \varepsilon}{k_B T}\right).$$

В результате

$$\sigma = \frac{ne^2 \tau}{3} \left[ \frac{2}{m_{\perp}} + \frac{\exp(\alpha \varepsilon / k_B T)}{m_{\parallel}} \right].$$

Окончательно:

$$\frac{\Delta \rho}{\rho_0} = \frac{\rho_0 - \rho}{\rho_0} = \frac{m_{\perp} [\exp(\alpha \varepsilon / k_B T) - 1]}{2m_{\parallel} + m_{\perp} \exp(\alpha \varepsilon / k_B T)}.$$

Здесь мы учли, что  $\rho_0 = 1/\sigma_0 = \frac{3m_{\perp} m_{\parallel}}{ne^2 \tau (m_{\perp} + 2m_{\parallel})}$  (см. (6.30)). Схематическая зависимость  $\Delta \rho / \rho_0$  от величины деформации  $\varepsilon$  представлена на рис. 6.27.

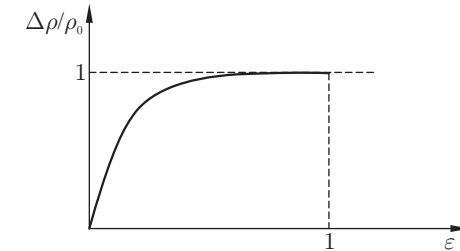


Рис. 6.27. Зависимость относительного изменения удельного сопротивления  $\Delta \rho / \rho_0$  от величины деформации  $\varepsilon$

**Задача 6.3.** Показать исходя из законов сохранения энергии и импульса, что если энергия электрона  $E \ll ms^2$ , то в этом случае электрон не может испускать фононы.

**Решение.** Запишем законы сохранения энергии и импульса для процесса испускания фонона:  $\frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 (k')^2}{2m} + \hbar s q$ ,  $\mathbf{k} = \mathbf{k}' + \mathbf{q}$ . Решая эту систему уравнений, получаем  $ms^2 = 2E \cos^2 \alpha$ , где  $E = \frac{\hbar^2 k^2}{2m}$ ,  $\alpha$  —



угол между  $\mathbf{k}'$  и  $\mathbf{q}$ . Поскольку  $\cos^2 \alpha \leq 1$ , то  $ms^2 \leq 2E$ . Таким образом, для излучения фононов энергия электрона  $E$  должна быть больше  $ms^2$ .

**Задача 6.4.** Выяснить, какой порядок величины имеет дрейфовая скорость электрона, когда начинается существенный разогрев. Считать, что рассеяние электронов происходит на акустических фононах.

**Задача 6.5.** Найти транспортное время релаксации электронов при рассеянии на примесях, описываемых потенциалом  $U(\mathbf{r}) = U_0 \delta(\mathbf{r})$ . Отличается ли транспортное сечение рассеяния от полного? Почему?

*Указание.* Воспользуйтесь уравнением (6.37). Для расчета  $W(\mathbf{p}, \mathbf{p}')$  используйте борновское приближение  $W(\mathbf{p}, \mathbf{p}') = \frac{2\pi}{\hbar} |V_{\mathbf{p}\mathbf{p}'}|^2 \delta(E_p - E_{p'})$ .

## Глава 7

# ПОВЕДЕНИЕ ИЗБЫТОЧНЫХ (НЕРАВНОВЕСНЫХ) НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКАХ

## 7.1. Введение

В предыдущих разделах мы рассмотрели свойства полупроводников, обусловленные свободными электронами и дырками, которые находятся в состоянии термодинамического равновесия. Наряду с такими равновесными состояниями, большой теоретический и практический интерес представляет изучение поведения неравновесных электронов и дырок. Избыточные электроны и дырки в полупроводнике можно создавать, например, светом. Пусть на некоторый участок полупроводника падает свет с энергией  $\hbar\omega > E_g$ , где  $E_g$  — ширина запрещенной зоны полупроводника, причем длительность импульса света  $\tau_\omega$  много меньше времени релаксации носителей заряда. В полупроводнике возникают неравновесные электроны и дырки (рис. 7.1). Возникающий пакет после выключения света расплывается и электронно-дырочные пары рекомбинируют со временем. Если теперь к полупроводнику приложено внешнее электрическое поле, то на электроны и дырки действует сила и они движутся в разные стороны (рис. 7.2). Однако далеко друг от друга они не уходят, так как они притягиваются за счет кулоновского взаимодействия. При этом по-прежнему происходит рекомбинация электронов и дырок и расплывание пакета. Можно показать (см. ниже), что пакет движется в ту сторону, в какую движутся неосновные носители полупроводника до освещения. Если неосновными носителями являются дырки, то пакет движется по направлению движения дырок, и наоборот. Таким образом, в полупроводнике, в котором созданы неравновесные носители, ситуация зависит от поведения неосновных носителей.

## 7.2. Нейтрализация пространственного заряда

Пусть в нейтральной среде случайно в каком-то месте создан пространственный заряд. Возникает вопрос, за какое время расплывается такой заряд. Чтобы ответить на этот вопрос, необходимо рассматривать следующие уравнения:



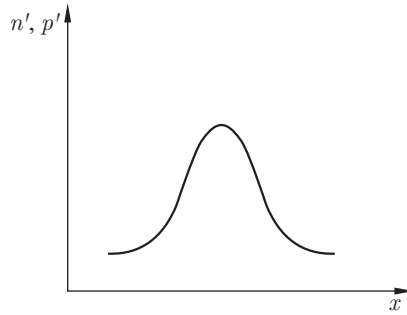
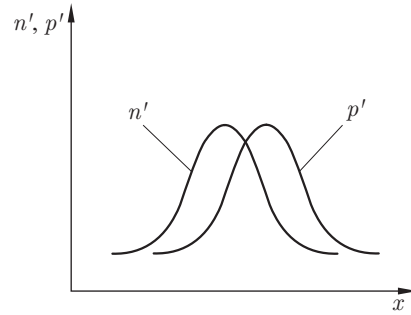


Рис. 7.1. Движение пакета неравновесных носителей в пространстве

Рис. 7.2. Движение неравновесных электронов и дырок в электрическом поле  $E_0$ . Направления скоростей движения электронов и дырок противоположны

а) уравнение непрерывности

$$\frac{\partial \rho}{\partial t} + \frac{\partial j}{\partial x} = 0, \quad (7.1)$$

б) уравнение Пуассона

$$\frac{\partial E}{\partial x} = \frac{4\pi}{\kappa} \rho, \quad (7.2)$$

в) закон Ома

$$j = \sigma E. \quad (7.3)$$

В этих уравнениях  $\rho$  — плотность заряда,  $j$  — плотность тока,  $E$  — напряженность электрического поля (для простоты мы рассматриваем одномерный случай). Тогда из (7.1)–(7.3) получаем

$$\frac{\partial \rho}{\partial t} = -\frac{4\pi\sigma}{\kappa} \rho \equiv -\frac{\rho}{\tau_m}. \quad (7.4)$$

Уравнение (7.4) и определяет поведение плотности заряда со временем

$$\rho = \rho_0 e^{-t/\tau_m}, \quad (7.5)$$

где  $\rho_0$  — плотность заряда в момент времени  $t = 0$ . Оно показывает, что плотность заряда будет распадаться со временем. Характерное время распада — это максвелловское время

$$\tau_m = \frac{\kappa}{4\pi\sigma}. \quad (7.6)$$

Как правило,  $\tau_m$  — это очень короткое время, поскольку проводимость  $\sigma$  — величина довольно большая. Все характерные времена, характеризующие другие процессы в системе, обычно гораздо больше  $\tau_m$ . Будем в дальнейшем считать, что  $\tau_m$  — самое короткое время в системе. Если так, то электронам и дыркам разойтись не удастся, они будут

двигаться вместе. Поведение электронов и дырок всегда подчиняются уравнениям Пуассона и неразрывности. Однако закон Ома (7.3) иногда нарушается. Рассмотрим общий случай, когда мы не знаем явной зависимости между плотностью тока  $j$  и напряженностью электрического поля  $E$ . Известно, что  $j$  есть функция  $E$ :  $j(E)$ . Перепишем уравнение непрерывности (7.1) в виде

$$\frac{\partial \rho}{\partial t} + \frac{\partial j}{\partial E} \frac{\partial E}{\partial x} = 0. \quad (7.7)$$

С учетом уравнения Пуассона (7.2) уравнение (7.7) можно записать так:

$$\frac{\partial \rho}{\partial t} = -\frac{4\pi}{\kappa} \frac{\partial j}{\partial E} \rho. \quad (7.8)$$

Величину  $\partial j / \partial E$  мы назовем дифференциальной проводимостью

$$\sigma_{\text{дифф}} = \frac{\partial j}{\partial E}. \quad (7.9)$$

Поэтому обобщение выражения для максвелловского времени, когда закон Ома несправедлив, нужно представить в виде

$$\tau_m = \frac{\kappa}{4\pi\sigma_{\text{дифф}}}. \quad (7.10)$$

Интересен случай, когда дифференциальная проводимость отрицательна,  $\sigma_{\text{дифф}} < 0$ . Тогда получается, что максвелловское время релаксации  $\tau_m$  также отрицательно. Рассмотрим, что физически означает, что  $\tau_m < 0$ . Поскольку поведение плотности заряда со временем определяется выражением  $\rho = \rho_0 \exp(-t/\tau_m)$ , то отрицательное значение  $\tau_m$  означает, что возникшая флуктуация заряда не распадается, а наоборот, нарастает со временем. Это означает, что в системе возникает неустойчивость. Это неустойчивость по отношению к малым возмущениям плотности заряда. Таким образом, в среде с отрицательной дифференциальной проводимостью (ОДП) состояние относительно разделения заряда неустойчиво (пример такого типа неустойчивости — маятник в верхнем стационарном положении). Подробнее о неустойчивых состояниях см. в конце этой главы. Мы пока рассмотрим электрические поля, для которых выполняется закон Ома, и флуктуации заряда («горбики») для электронов и дырок, движущихся вместе, когда нет пространственного заряда.

**Замечание.** Выше мы рассматривали плотность тока  $j$  в данной точке как функцию электрического поля  $E$  в данной точке. Но может оказаться, что ток в данной точке пространства может зависеть от поля  $E$  в других точках (это так называемая пространственная дисперсия). И кроме того, ток  $j$  в данный момент времени может зависеть от поля  $E$  в предшествующие моменты времени (это временная дисперсия). Когда необходимо учитывать временную и пространственную диспер-

сии, тогда уравнение для тока как функции поля ( $j = f(E)$ ) является интегральным уравнением.

### 7.3. Дрейф неравновесных носителей заряда

Рассмотрим поведение неравновесных электронов и дырок во внешнем электрическом поле  $E$ . Будем теперь интересоваться не вопросом расщепления горбика, а направлением его распространения. Для этого рассмотрим выражения для потоков электронов  $q_n$  и дырок  $q_p$ . По определению поток электронов равен

$$q_n = -nb_n E. \quad (7.11)$$

Здесь  $n$  — концентрация электронов,  $b_n$  — подвижность электронов,  $b_n E$  — их скорость. Знак минус в (7.11) означает, что электроны движутся против поля  $E$ . Аналогично для дырок имеем

$$q_p = pb_p E, \quad (7.12)$$

где  $p$  — концентрация дырок;  $b_p$  — подвижность дырок. Концентрация электронов  $n$  состоит из двух частей:

$$n = n_0 + n', \quad (7.13)$$

где  $n_0$  — концентрация равновесных электронов;  $n'$  — концентрация избыточных (неравновесных) электронов. Для дырок имеет место аналогичное выражение

$$p = p_0 + p' \equiv p_0 + n'. \quad (7.14)$$

Здесь для простоты считаем, что концентрация избыточных электронов  $n'$  равна концентрации избыточных дырок  $p'$ . Кроме того, напряженность электрического поля равна

$$E = E_0 + E', \quad (7.15)$$

где  $E_0$  — приложенное внешнее электрическое поле, то есть поле, создаваемое внешним источником;  $E'$  — это электрическое поле, создаваемое все-таки малым расхождением (в пространстве) электронов и дырок. В выражении для концентрации мы это расхождение не учитывали. В дальнейшем мы будем считать, что величины  $E'$  и  $n'$  малы по сравнению с  $E_0$  и  $n_0$  ( $E' \ll E_0$ ,  $n' \ll n_0$ ). Тогда выражение для потоков электронов и дырок можно представить в виде

$$\begin{aligned} q_n &= q_{n0} + q'_n, \\ q_p &= q_{p0} + q'_p, \end{aligned} \quad (7.16)$$

а избыточные потоки для электронов и дырок равны

$$\begin{aligned} q'_n &= -b_n(n'E_0 + n_0E'), \\ q'_p &= +b_p(n'E_0 + p_0E'). \end{aligned} \quad (7.17)$$

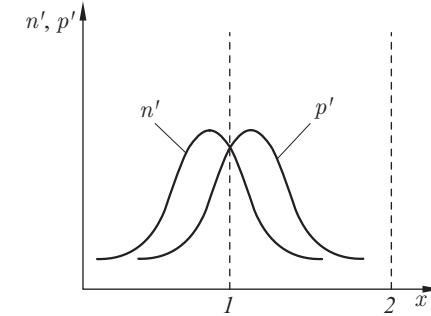


Рис. 7.3. Движение электронов и дырок в электрическом поле  $E_0$ . Создаваемый ими полный электрический ток одинаков во всех сечениях полупроводника

Видно, что избыточные потоки состоят из двух частей: одна часть создана избыточными носителями в том поле, которое было в системе, это  $E_0$ ; другая часть — это поток равновесных электронов, на которые действует добавочное поле  $E'$ . Поле  $E'$  можно найти из условия, что избыточные электроны и дырки не удаляются далеко друг от друга (за счет притяжения) и движутся одинаково. Это означает, что потоки  $q'_n$  и  $q'_p$  одинаковы:

$$q'_n = q'_p. \quad (7.18)$$

Соотношение (7.18) есть следствие того, что полный ток должен быть одинаковым во всех сечениях полупроводника: и в том месте, где есть избыточные носители  $n'$  (см. рис. 7.3, точка «1»), и там, где их нет (рис. 7.3, точка «2»). Это есть фундаментальное свойство

$$j(\text{«1»}) = e(q_{n0} - q_{p0}) + e(q'_n - q'_p) = j(\text{«2»}) = e(q_{n0} - q_{p0}). \quad (7.19)$$

Из последнего равенства (7.19) следует соотношение (7.18). При выводе (7.18) мы не учли ток смещения. Если в системе существует также и ток смещения, то соотношение (7.19) остается в силе, но при этом полный ток в точках «1» и «2» складывается из двух частей: из тока  $j$  и из тока смещения. Поскольку мы пренебрегали током смещения, то получили соотношение (7.18). Из условия равенства потоков (7.18) с учетом (7.17) найдем поле

$$E' = -\frac{b_n + b_p}{b_n n_0 + b_p p_0} E_0 \cdot n'. \quad (7.20)$$

Поле  $E'$ , возникающее из-за разделения зарядов, направлено против внешнего поля  $E_0$ , причем поле  $E'$  существует в том месте, где имеются избыточные носители  $n'$ .

Подставим поле  $E'$  в выражение для потоков и найдем, с какой скоростью будет происходить движение (дрейф) неравновесных носи-

телей:

$$q'_n = q'_p = n' \frac{(n_0 - p_0) b_n b_p}{n_0 b_n + p_0 b_p} E_0. \quad (7.21)$$

Введем обозначение

$$b_{\text{бип}} = \frac{(n_0 - p_0) b_n b_p}{n_0 b_n + p_0 b_p}. \quad (7.22)$$

Величина  $b_{\text{бип}}$  называется биполярной подвижностью. Тогда выражение для потоков (7.21) можно переписать в виде

$$q'_n = q'_p = n' E_0 b_{\text{бип}}. \quad (7.23)$$

Проанализируем полученный результат. Пусть полупроводник, который мы изучаем, является полупроводником  $n$ -типа, то есть  $n_0 > p_0$ . Тогда, согласно (7.22), биполярная подвижность положительна ( $b_{\text{бип}} > 0$ ). Следовательно, дрейфовая скорость направлена в ту же сторону, куда движутся дырки. Если же полупроводник  $p$ -типа, то есть  $n_0 < p_0$ , то дрейф направлен в сторону движения электронов.

Итак, дрейф неравновесных носителей направлен в ту сторону, куда движутся неосновные носители.

Если  $n_0 \gg p_0$ , то из (7.22) следует, что

$$b_{\text{бип}} = b_p. \quad (7.24)$$

В этом случае неравновесные носители движутся не только в ту сторону, куда движутся неосновные носители (дырки), но и с такой же подвижностью, что и неосновные носители. В результате в примесном полупроводнике ситуация следующая: электрическое поле действует на носители, которых меньше, а те носители, которых много, подстраиваются под тех носителей, которых меньше.

## 7.4. Расплывание пакета

Итак, мы нашли электрическое поле  $E'$ , возникающее при разделении волновых пакетов для электронов и дырок (рис. 7.2), и показали, что это поле направлено против внешнего поля  $E_0$ , и определили дрейф неравновесных носителей в таком поле.

Займемся теперь вопросом распыливания пакета избыточных носителей в пространстве. Такое распыливание происходит за счет процессов диффузии. Очевидно, что в то место, где частиц меньше, диффундируют частицы из области больших концентраций избыточных носителей. При этом поток направлен в сторону убывания частиц:

$$\begin{aligned} q_n &= -D_n \frac{dn}{dx}, \\ q_p &= -D_p \frac{dp}{dx}, \end{aligned} \quad (7.25)$$

где  $D_n$  и  $D_p$  — коэффициенты диффузии для электронов и дырок соответственно. Возникает вопрос, могут ли расплываться потоки  $q_n$  и  $q_p$ , согласно уравнениям (7.25). Как правило, из-за отличия в эффективных массах электроны диффундируют быстрее дырок, тогда электроны движутся быстрее, а дырки немного отстают. Может ли быть так, чтобы концентрация электронов успела выровняться, а концентрация дырок не успела? Как уже было показано выше, этого быть не может. Если есть разделение зарядов, то возникает электрическое поле, которое препятствует уходу электронов и ускоряет движение дырок. Следовательно, такого сильного нарушения электронейтральности (то есть ухода электронов и дырок на большие расстояния друг относительно друга) на самом деле не существует.

Для простоты пренебрежем внешним электрическим полем  $E_0$ . Так как даже при небольшом пространственном разделении электронов и дырок возникает электрическое поле  $E''$ , то выражения для потоков носителей  $q_{n,p}$  (7.25) нужно скорректировать с учетом этого поля:

$$\begin{aligned} q_n &= -D_n \frac{dn}{dx} - b_n n E'', \\ q_p &= -D_p \frac{dp}{dx} + b_p p E''. \end{aligned} \quad (7.26)$$

В выражении для  $q_n$  знак «минус» перед вторым слагаемым означает, что электроны движутся против поля  $E''$ . Как и прежде, представим концентрацию электронов и дырок в виде

$$\begin{aligned} n &= n_0 + n', \\ p &= p_0 + p'. \end{aligned} \quad (7.27)$$

По-прежнему, для простоты считаем, что концентрации избыточных электронов и дырок равны:  $n' = p'$ . Равновесные концентрации электронов и дырок  $n_0$  и  $p_0$  не зависят от координат. В итоге, выражения для потоков  $q_{n,p}$  принимают вид

$$\begin{aligned} q_n &\equiv q'_n = -D_n \frac{dn'}{dx} - b_n n_0 E'', \\ q_p &= -D_p \frac{dp'}{dx} + b_p p_0 E''. \end{aligned} \quad (7.28)$$

В (7.28) мы пренебрегли членами типа  $n' E''$ . Поле  $E''$  мы найдем, как и прежде, из условия того, чтобы электроны и дырки двигались вместе. Это означает, что

$$q'_n = q'_p. \quad (7.29)$$

Подставляем (7.28) в (7.29), получаем окончательное выражение

$$E'' = \frac{(D_p - D_n)}{b_p p_0 + b_n n_0} \frac{dn'}{dx}. \quad (7.30)$$

Если выполняется условие  $D_p < D_n$ , то поле  $E''$  ускоряет дырки и тормозит электроны. Подставляя поле  $E''$  в выражение для потоков  $q'_{n,p}$ , получаем

$$q'_n = q'_p = -\frac{(D_p b_n n_0 + D_n b_p p_0)}{b_p p_0 + b_n n_0} \frac{dn'}{dx} \equiv -D_a \frac{dn'}{dx}. \quad (7.31)$$

Выражение

$$D_a = \frac{(D_p b_n n_0 + D_n b_p p_0)}{b_p p_0 + b_n n_0} \quad (7.32)$$

называется амбиполярным коэффициентом диффузии.

Если концентрация электронов существенно больше концентрации дырок  $n_0 \gg p_0$ , то  $D_a = D_p$ , и наоборот, если  $p_0 \gg n_0$ , то  $D_a = D_n$ . Таким образом, распывание пакета описывается коэффициентом диффузии неосновных носителей. Здесь ситуация аналогична задаче о дрейфе импульса неравновесных носителей. Для собственного полупроводника  $n_0 = p_0$ , тогда коэффициент амбиполярной диффузии равен

$$D_a = \frac{D_p b_n + D_n b_p}{b_n + b_p}. \quad (7.33)$$

Это выражение для  $D_a$ , как правило, используется при расчетах, связанных с обычной плазмой, в которой концентрация электронов равна концентрации положительных ионов.

### 7.5. Эффект Дембера

В качестве примера рассмотрим эффект, который возникает в полупроводниках при освещении. На пластинку полупроводника пада-

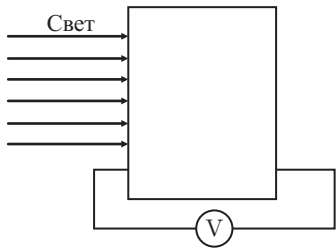


Рис. 7.4. Схема наблюдения эффекта Дембера

ется свет с энергией кванта  $\hbar\omega$ , большей ширины запрещенной зоны  $E_g$  (рис. 7.4). Электромагнитное излучение (свет) сильно поглощается пластинкой и проникает вглубь полупроводника на толщину порядка 1 мкм. На поверхности пластинки рождаются электроны и дырки. Оказывается, что между освещенной и противоположной поверхностями полупроводника возникает разность потенциалов, причем довольно существенная. Это и есть эффект Дембера. Возникающее при этом поле называется полем Дембера. Это поле возникает из-за разности коэффициентов диффузии электронов и дырок (см. (7.30)):

$$E'' = \frac{(D_p - D_n)}{b_n n_0 + b_p p_0} \frac{dn'}{dx}. \quad (7.34)$$

Выражение (7.34) определяет величину поля Дембера. Картина движения носителей следующая: электроны стремятся двигаться быстрее, чем дырки, а поле  $E''$  стремится их удержать, при этом поле  $E''$  ускоряет дырки.

### 7.6. Рекомбинация избыточных носителей в полупроводниках

Данный раздел посвящен рекомбинации избыточных (неравновесных) электронов и дырок в полупроводниках. Так как избыточные носители живут в полупроводнике достаточно долго, то рекомбинация идет медленно, так что успевают произойти процессы диффузии, дрейфа и т.д. Вначале для исследования процессов рекомбинации воспользуемся феноменологическим подходом.

**Линейная рекомбинация. Время жизни.** По-прежнему обозначим концентрацию избыточных носителей через  $n'$  и  $p'$ . Рассмотрим для определенности полупроводник  $p$ -типа (рис. 7.5). В полупроводнике имеется много равновесных дырок  $p_0$ . С помощью электромагнитного излучения (света) создаем избыточные носители заряда. При рекомбинации электроны могут рекомбинировать с любой дыркой, как с равновесной, так и с неравновесной. Так как концентрация равновесных дырок много больше концентрации неравновесных дырок ( $p_0 \gg p'$ ), мы не будем учитывать вклад неравновесных дырок в процессе рекомбинации, учтем в дальнейшем только равновесные дырки. Вычислим число рекомбинирующих электрон-дырочных пар в секунду. Очевидно, что это число пропорционально концентрации избыточных электронов  $n'$  и равновесных дырок  $p_0$

$$R = n' p_0 A, \quad (7.35)$$

где  $A$  — коэффициент пропорциональности. Соотношение (7.35) всегда справедливо при малых концентрациях избыточных электронов  $n'$ . Тогда изменение концентрации избыточных электронов со временем определяется дифференциальным уравнением

$$\frac{dn'}{dt} = -n' p_0 A. \quad (7.36)$$

В данном случае изменение концентрации избыточных электронов  $n'$  происходит за счет рекомбинации с дырками, поэтому в уравнении

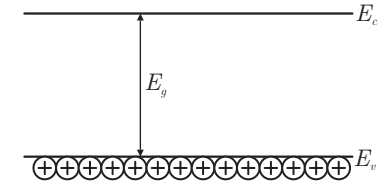


Рис. 7.5. Зонная диаграмма полупроводника  $p$ -типа

(7.36) можно ввести время их жизни  $\tau_{\text{ж}} = 1/p_0 A$ . Тогда уравнение (7.36) приобретает вид

$$\frac{dn'}{dt} = -\frac{1}{\tau_{\text{ж}}} n'. \quad (7.37)$$

Следовательно, после выключения источника света концентрация избыточных носителей убывает, согласно (7.37), по экспоненциальному закону

$$n' = n'_0 \exp\left(-\frac{t}{\tau_{\text{ж}}}\right), \quad (7.38)$$

где  $n'_0$  — концентрация избыточных электронов в момент времени  $t = 0$  (это момент времени выключения источника света). Время жизни  $\tau_{\text{ж}}$  тем короче, чем больше концентрация основных носителей  $p_0$ . Оказывается, что время жизни бывает разным для разных полупроводников. Например, в непрямозонных полупроводниках Ge и Si время жизни  $\tau_{\text{ж}} \sim 10^{-6} - \alpha \cdot 10^{-6}$  с, где  $\alpha < 1$ . Для прямозонных полупроводников типа GaAs  $\tau_{\text{ж}} \sim 10^{-8} - 10^{-10}$  с. Тем не менее, время жизни намного больше времени релаксации  $\tau$  ( $\tau \sim 10^{-12}$  с). Оказывается, что время жизни — это очень важная характеристика полупроводника. Для оптоэлектронных приборов эта характеристика играет большую роль — она определяет быстродействие оптических приборов.

Далее, очень интересно рассмотреть случай, когда концентрация избыточных носителей уже не является малой, то есть рассмотреть нелинейную теорию рекомбинации. Ниже мы рассмотрим два типа такой рекомбинации.

## 7.7. Нелинейная теория рекомбинации

**7.7.1. Прямая рекомбинация электрона и дырки (бимолекулярная рекомбинация).** Этот механизм рекомбинации соответствует бимолекулярной рекомбинации из теории газов. Несколько слов о микроскопическом подходе к такому процессу рекомбинации. Сначала ответим на вопрос, почему рекомбинация электрона и дырки происходит так медленно (напомним, что это самый медленный процесс в полупроводнике). При прямой рекомбинации электрона и дырки освобождается энергия порядка ширины запрещенной зоны  $E_g$ . Возникает вопрос, кому отдать эту энергию. Самый простой процесс — это излучение фотона, когда электрон-дырочная пара рекомбинирует и излучает квант света. Время излучательной рекомбинации, то есть время таких квантовых переходов, — это величина для прямозонных полупроводников порядка  $10^{-8}$  с. Это так называемое спонтанное излучение.

Возможен еще один канал, когда энергия отдается решетке; при этом необходимо излучить несколько фононов. Энергия фонона, как правило, порядка нескольких сотых электрон вольта (эВ). Поэтому при таком квантовом переходе необходимо, чтобы произошло излучение

большого количества фононов (больше десятка). Но, как следует из квантовой механики, такой многофононный процесс маловероятен.

Итак, говоря о прямой рекомбинации электрон-дырочных пар, будем иметь в виду излучательную рекомбинацию (рис. 7.6). Пусть в полупроводнике создана большая концентрация избыточных носителей. Наряду с рекомбинацией имеет место еще и тепловая генерация избыточных носителей. Как правило, тепловая генерация носителей идет медленно, но она существует.

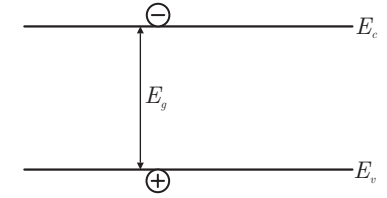


Рис. 7.6. Прямая излучательная рекомбинация электрона и дырки в полупроводнике

Рассмотрим прямую рекомбинацию и найдем число пар, рекомбинирующих в единицу времени. Очевидно, что число пар пропорционально концентрации электронов в зоне проводимости  $n$  и дырок в валентной зоне  $p$ :

$$R = Anp. \quad (7.39)$$

Здесь  $A$  — коэффициент пропорциональности. Рассмотрим теперь скорость генерации  $G$ . Это число электронов и дырок, которые генерируются в единицу времени. Это число зависит от температуры и не зависит от концентрации электронов и дырок  $n$  и  $p$  (если носители не вырождены). Оказывается, что постоянную генерации  $G$  и коэффициент  $A$  в выражении для  $R$  (см. (7.39)) можно связать между собой, если речь идет о прямом и обратном процессе. При этом существует принцип детального равновесия — в термодинамическом равновесии прямой и обратный процессы равны друг другу.

Важно, что для каждого процесса генерации и рекомбинации это равенство в термодинамическом равновесии должно выполняться отдельно:

$$G = R. \quad (7.40)$$

Из принципа детального равновесия, согласно (7.39) и (7.40), получаем

$$G = An_0 p_0, \quad (7.41)$$

где  $n_0$  и  $p_0$  — концентрации электронов и дырок в термодинамическом равновесии. Для невырожденного полупроводника в термодинамическом равновесии имеет место соотношение

$$n_0 p_0 = n_i^2, \quad (7.42)$$

где  $n_i$  — собственная концентрация носителей. В результате для скорости генерации получаем

$$G = An_i^2. \quad (7.43)$$

Изменение концентрации носителей со временем описывается уравнением баланса

$$\frac{dn}{dt} = G - R = -A(np - n_i^2). \quad (7.44)$$

В тепловом равновесии  $dn/dt = 0$ , как и должно быть. Если есть внешний источник носителей заряда, то к уравнению (7.44) необходимо добавить еще и третье слагаемое. Положим в (7.44)  $n = n_0 + n'$  и будем считать, что  $n'$  мало ( $n' \ll n_0$ ). Рассмотрим полупроводник  $p$ -типа и пренебрежем избыточными дырками (то есть  $p = p_0$ ). Тогда уравнение (7.44) превращается в закон линейной рекомбинации

$$\frac{dn'}{dt} = -Ap_0n' = -\frac{n'}{\tau_{ж}}. \quad (7.45)$$

Здесь

$$\frac{1}{\tau_{ж}} = Ap_0.$$

**7.7.2. Мономолекулярная рекомбинация. Статистика рекомбинации Шокли–Рида.** Рассмотрим более сложный случай рекомбинации через глубокие центры (ловушки). Экспериментально показано, что время жизни носителей заряда сильно зависит от концентрации примесей и дефектов в полупроводниках. Примеси сокращают время жизни избыточных носителей, увеличивая скорость рекомбинации. Все это происходит при условии, что атом примеси создает уровни в запрещенной зоне (речь идет о том, что уровни глубокие) (рис. 7.7). При наличии уровней в запрещенной зоне электрон из зоны проводимости может захватиться на примесный центр. Затем через некоторое время электрон рекомбинирует с дыркой, которая всплывает к центру из валентной зоны (см. рис. 7.7). Такой процесс рекомбинации является более эффективным, так как при каждом захвате отданная энергия меньше  $E_g$ , таким образом, энергия  $E_g$  отдается по частям. Примеси, через посредство которых происходит рекомбинация электронов и дырок, называются центрами рекомбинации, или глубокими ловушками. Экспериментально такого рода процессы были обнаружены при прохождении тока через  $p - n$ -переход. Этот процесс рекомбинации является самым явным конкурирующим процессом прямой рекомбинации носителей.

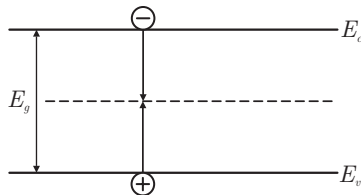


Рис. 7.7. Рекомбинация электронов и дырок через локальные уровни

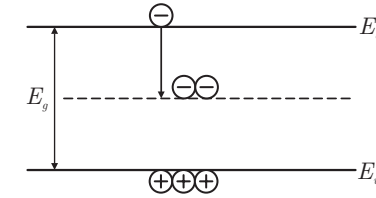


Рис. 7.8. Процесс захвата электронов из зоны проводимости на пустой центр рекомбинации (ловушку)

Рассмотрим более детально рекомбинацию Шокли–Рида через глубокую примесь. В этом случае носители заряда могут «погибать» не парами, а поодиночке. При этом уравнения для скорости рекомбинации электронов и дырок нужно рассматривать отдельно. Пусть в полупроводнике имеются примеси, уровни энергии которых расположены глубоко в запрещенной зоне (см. рис. 7.8). Часть этих уровней заполнена электронами, а часть — пустые. Обозначим число уровней через  $N$ , а долю заполненных уровней (примесей) — через  $f$ . Тогда число заполненных уровней равно  $Nf$ , а число пустых примесей (центров) равно  $N(1 - f)$ . Нужно найти вероятность рекомбинации электронов, иными словами, число актов захвата электронов на примеси в секунду. Эта вероятность пропорциональна числу электронов, то есть концентрации электронов в зоне проводимости  $n$  и числу свободных ловушек  $N(1 - f)$ :

$$R_n = \alpha_n n N(1 - f), \quad (7.46)$$

где  $\alpha_n$  — коэффициент пропорциональности. Аналогичным образом можно написать выражение для скорости генерации электронов с уровней ловушки в зону проводимости. Скорость генерации  $G_n$  пропорциональна числу заполненных ловушек  $Nf$ :

$$G_n = \beta_n Nf. \quad (7.47)$$

Здесь, по аналогии с (7.46),  $\beta_n$  — коэффициент пропорциональности.

Вероятность захвата дырок на ловушке  $R_p$  пропорциональна концентрации дырок в валентной зоне  $p$  и числу заполненных электронами ловушек  $Nf$ :

$$R_p = \alpha_p p Nf, \quad (7.48)$$

где  $\alpha_p$  — коэффициент пропорциональности.

Рассмотрим механизм генерации дырок. Каким образом дырки могут появиться в валентной зоне? Это такой процесс, когда электрон может быть выбит тепловыми колебаниями из валентной зоны на ловушки. Скорость генерации в этом случае пропорциональна числу пустых ловушек  $N(1 - f)$ , которыми может захватиться электрон:

$$G_p = \beta_p N(1 - f). \quad (7.49)$$



Здесь  $\beta_p$  — коэффициент пропорциональности. Мы нашли интересные нас выражения для скоростей рекомбинации-генерации, которые содержат четыре коэффициента:  $\alpha_n$ ,  $\beta_n$ ,  $\alpha_p$  и  $\beta_p$ . Из принципа детального равновесия находим связь между ними. В термодинамическом равновесии числа прямых и обратных процессов равны друг другу:

$$R_n = G_n, \quad \alpha_n n_0 N (1 - f_0) = \beta_n N f_0. \quad (7.50)$$

Здесь  $n_0$  и  $f_0$  — термодинамические равновесные значения концентрации электронов в зоне проводимости и степени заполнения ловушек электронами. Аналогичным образом для дырок имеем

$$R_p = G_p; \quad \alpha_p p_0 N f_0 = \beta_p N (1 - f_0), \quad (7.51)$$

где  $p_0$  — равновесное значение концентрации дырок в валентной зоне. Из (7.50) и (7.51) мы можем выразить коэффициенты  $\beta_{n,p}$  через  $\alpha_{n,p}$ :

$$\begin{aligned} \beta_n &= \alpha_n n_0 \frac{1 - f_0}{f_0}, \\ \beta_p &= \alpha_p p_0 \frac{f_0}{1 - f_0}. \end{aligned} \quad (7.52)$$

Равновесные концентрации электронов  $n_0$  и дырок  $p_0$  нам известны:

$$n_0 = N_c \exp\left(-\frac{\chi_c}{k_B T}\right), \quad (7.53)$$

где  $\chi_c$  — расстояние от уровня Ферми  $\mu$  до дна зоны проводимости (см. рис. 7.9),  $N_c$  — эффективное число состояний в зоне проводимости. Степень заполнения уровня ловушки  $f_0$  мы уже вычисляли (см. главу 5):

$$f_0 = \frac{1}{\frac{1}{g} \exp\left(\frac{E - \mu}{k_B T}\right) + 1}. \quad (7.54)$$

Здесь  $E$  — энергия на ловушке;  $\mu$  — уровень Ферми и  $g$  — кратность вырождения ловушки. В результате

$$\frac{1 - f_0}{f_0} = \frac{1}{g} \exp\left(\frac{E - \mu}{k_B T}\right). \quad (7.55)$$

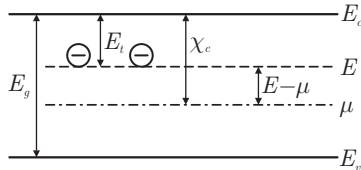


Рис. 7.9. Схема заполнения уровня ловушки электронами в зависимости от положения энергетического уровня дефектов  $E_t$ , когда уровень химического потенциала  $\mu$  находится ниже уровня дефектов

Например, для коэффициента  $\beta_n$  получаем, что

$$\beta_n = \frac{\alpha_n}{g} N_c \exp\left(-\frac{\chi_c}{k_B T}\right) \exp\left(\frac{E - \mu}{k_B T}\right). \quad (7.56)$$

Согласно рис. 7.9, разность  $(E - \mu)$  — это расстояние от уровня Ферми  $\mu$  до уровня ловушки. Если из  $\chi_c$  отнять  $(E - \mu)$ , то мы получим расстояние между уровнем ловушки и дном зоны проводимости  $E_c$ , то есть это энергия связи электрона на ловушке  $E_t$ :

$$\chi_c - (E - \mu) = E_t. \quad (7.57)$$

В результате для  $\beta_n$  получим

$$\beta_n = \frac{\alpha_n}{g} N_c \exp\left(-\frac{E_t}{k_B T}\right). \quad (7.58)$$

Величина  $N_c \exp(-E_t/k_B T)$  имеет размерность концентрации. Это есть та концентрация электронов, которая была бы в термодинамическом равновесии, если бы уровень Ферми совпадал с уровнем ловушки. Эту концентрацию обозначим через  $n_1$  <sup>1)</sup>:

$$n_1 = \frac{1}{g} N_c \exp\left(-\frac{E_t}{k_B T}\right). \quad (7.59)$$

В результате для коэффициента  $\beta_n$  получаем выражение

$$\beta_n = \alpha_n \cdot n_1. \quad (7.60)$$

Проводя аналогичные вычисления, получаем, что

$$\beta_p = \alpha_p \cdot p_1. \quad (7.61)$$

Здесь  $p_1 = g N_v \exp(-(E_g - E_t)/k_B T)$  — это та концентрация дырок, которая была бы в термодинамическом равновесии, если бы уровень Ферми совпал с уровнем примеси. Теперь мы можем найти окончательные выражения для скорости генерации и рекомбинации. Эффективная скорость генерации-рекомбинации равна разности между числом уходящих электронов (дырок)  $R_{n,p}$  и числом приходящих электронов (дырок)  $G_{n,p}$  в секунду:

$$\begin{aligned} R_n - G_n &= \alpha_n N [n(1 - f) - n_1 f], \\ R_p - G_p &= \alpha_p N [p f - p_1 (1 - f)]. \end{aligned} \quad (7.62)$$

Формулы (7.62) и лежат в основе статистики Шокли-Рида. Наибольший интерес представляет стационарный режим, когда концентрация электронов на ловушках постоянна и не меняется со временем, при этом она может быть и неравновесной. Важно то, что в стационарном состоянии число электронов, приходящих на ловушки, равно числу дырок, приходящих на ловушки. В результате, приравнявая выражения

<sup>1)</sup> Здесь учтено, что ловушка может принять только один электрон.



(7.62) друг другу, мы найдем долю заполненных уровней  $f$ . Подставляя найденное выражение для  $f$  в любое из двух уравнений (7.62), получим окончательное выражение для эффективной скорости генерации-рекомбинации

$$R_n - G_n = R_p - G_p = \frac{(np - n_i^2) \frac{1}{\tau_{n0}\tau_{p0}}}{(n + n_1) \frac{1}{\tau_{n0}} + (p + p_1) \frac{1}{\tau_{p0}}} = \frac{np - n_i^2}{(n + n_1)\tau_{p0} + (p + p_1)\tau_{n0}}. \quad (7.63)$$

Здесь введены обозначения:

$$\frac{1}{\tau_{n0}} = N\alpha_n, \quad \frac{1}{\tau_{p0}} = N\alpha_p, \quad (7.64)$$

где  $\alpha_n$  и  $\alpha_p$  называются коэффициентами захвата электронов и дырок на уровень ловушки соответственно. Тогда  $\tau_{n0}$  — это время жизни электрона в зоне проводимости, которое было бы, если бы все ловушки были пустыми, а  $\tau_{p0}$  — это время жизни дырок в валентной зоне, которое было бы, если бы все ловушки были заполнены.

Обсудим выражение, полученное для скорости генерации-рекомбинации в двух случаях.

**1) Избыточных носителей мало.** Пусть мы рассматриваем полупроводник  $p$ -типа. В этом случае концентрация дырок  $p$  велика, но концентрация избыточных электронов и дырок мала. При этом уровень химического потенциала находится между потолком валентной зоны и уровнем ловушек (см. рис. 7.10). Следовательно, все ловушки пустые. Это как раз соответствует случаю, когда время жизни электрона в зоне проводимости равно  $\tau_{n0}$  (см. (7.64) и следующий за ней текст). Тогда

$$R_n - G_n = \frac{n'}{\tau_{n0}}. \quad (7.65)$$

При выводе (7.65) мы учли, что  $p \gg n$ ,  $p_0 \gg p'$ ,  $p = p_0$  и  $n = n' + n_0$ , тогда  $np - n_i^2 = n'p_0$ .

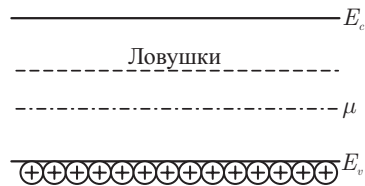


Рис. 7.10. Относительное расположение уровня химического потенциала  $\mu$  и уровня дефектов для полупроводника  $p$ -типа. Уровень  $\mu$  находится ниже уровня дефектов

**2) Избыточных носителей много.** Рассмотрим наиболее интересный случай, когда избыточных носителей много. С помощью внешнего источника накачки мы создаем много избыточных носителей, гораздо больше, чем в равновесии, и гораздо больше, чем ловушек, тогда  $n \gg n_1$ ,  $p \gg p_1$ . В этом случае должно выполняться условие нейтральности, то есть  $n \approx p$ . Этот случай соответствует сильно нелинейному режиму. Полагая в (7.63)  $n \approx p$  и  $n \gg n_1$ , получаем

$$R_n - G_n = n \frac{1}{\tau_{p0} + \tau_{n0}}. \quad (7.66)$$

В случае сильной накачки, то есть большой концентрации избыточных носителей, зависимость скорости генерации-рекомбинации от концентрации по-прежнему (см. (7.66)) является линейной и пропорциональной концентрации электронов. Однако эффективное время жизни электрона представляет собой сумму  $\tau_{p0}$  и  $\tau_{n0}$ :

$$\tau_{ж} = \tau_{p0} + \tau_{n0}. \quad (7.67)$$

Соотношение (7.67) можно интерпретировать так: для того, чтобы электрон рекомбинировал с дыркой, ему необходимо быть захваченным ловушкой; и через некоторое время ( $\tau_{p0}$ ) ловушкой захватывается дырка, после чего они рекомбинируют. Это один из случаев, когда складываются не обратные времена ( $\tau^{-1}$ ), а сами времена  $\tau$ . Экспериментально как раз удастся измерить именно  $\tau_{ж}$ .

## 7.8. Дрейфовая длина и диффузионная длина

Эффективная длина, которую проходит электрон (дырка) в электрическом поле за время жизни  $\tau_{ж}$ , называется дрейфовой длиной (это длина, на которой дрейфуют носители заряда в электрическом поле за время жизни). На самом деле дрейфует волновой пакет. Скорость дрейфа избыточных (неосновных) носителей, как было показано в § 7.3, равна

$$v_{др} = b_{бип} E. \quad (7.68)$$

Соответственно дрейфовая длина имеет вид

$$L_E = b_{бип} E \tau_{ж}. \quad (7.69)$$

Для наглядности, оценим эту длину. Пусть напряженность электрического поля равна  $E = 10$  В/см,  $b_{бип} = 3000$  см<sup>2</sup>/(В · с),  $\tau_{ж} = 10^{-6}$  с, тогда  $L_E = 3 \cdot 10^{-2}$  см. Мы получили, что дрейфовая длина имеет макроскопический масштаб. Реально  $L_E$  может достигать нескольких миллиметров, это значит, что электрон, помещенный во внешнее электрическое поле, может пройти такое расстояние, прежде чем погибнуть.

**Диффузионная длина** — это длина, на которую должны продиффундировать избыточные носители без поля, прежде чем погибнуть,

или это длина, на которой расплывается пакет избыточных носителей, прежде чем они погибнут.

Чтобы найти диффузионную длину рассмотрим следующую задачу. Пусть на границе полупроводника поддерживается избыточная концентрация носителей  $n'(0)$  (см. рис. 7.11). Носители от границы диффундируют вглубь образца и там погибают. Нужно найти зависимость концентрации от расстояния, то есть закон, по которому гибнут эти носители заряда. Для решения задачи напишем уравнение неразрывности для избыточных носителей

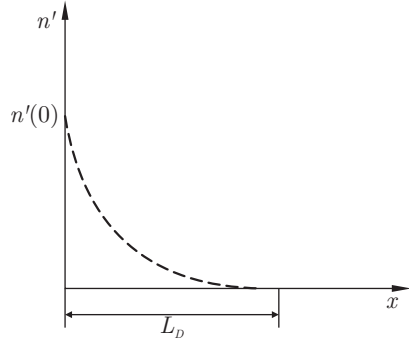


Рис. 7.11. Распределение избыточной концентрации электронов в случае их диффузии. Концентрация уменьшается в « $e$ » раз на диффузионной длине  $L_D$

$$\frac{\partial n'}{\partial t} + \text{div } \mathbf{q}_{n'} = -\frac{n'}{\tau_{\text{ж}}}. \quad (7.70)$$

В левой части (7.70) представлено изменение числа электронов в единицу времени и разность входящего и выходящего через сечение потоков частиц, справа

представлено слагаемое, ответственное за гибель электронов. Рассмотрим стационарный режим, когда  $\partial n'/\partial t = 0$ . Учтем, что для одномерного движения диффузионный поток электронов равен

$$q_{n'} = -D_a \frac{dn'}{dx}, \quad (7.71)$$

где  $D_a$  — коэффициент амбиполярной диффузии. В результате уравнение, определяющее распределение избыточных электронов, имеет вид

$$D_a \frac{d^2 n'}{dx^2} = \frac{n'}{\tau_{\text{ж}}}. \quad (7.72)$$

Уравнение (7.72) — это дифференциальное уравнение второго порядка. Оно имеет два независимых решения. Одно из решений не имеет физического смысла, так как оно соответствует увеличению концентрации носителей в глубине полупроводника, и мы его отбрасываем. Второе решение уравнения (7.72) имеет вид

$$n'(x) = n'(0) \exp\left(-\frac{x}{L_D}\right), \quad (7.73)$$

где  $L_D = (D_a \tau_{\text{ж}})^{1/2}$ . Итак, мы получили, что концентрация избыточных носителей уменьшается в « $e$ » раз на диффузионной длине  $L_D$ . Оценим численно  $L_D$ . Выше для определения времени жизни носителей мы

использовали значения  $\tau_{\text{ж}} = 10^{-6}$  с. Выразим коэффициент диффузии через подвижность согласно соотношению Эйнштейна

$$D = \frac{bk_B T}{e}. \quad (7.74)$$

Это универсальное соотношение, которое выполняется всегда при малой неравновесности. Для подвижности мы брали значение  $b = 3000 \text{ см}^2/(\text{В} \cdot \text{с})$ . Тогда при комнатной температуре коэффициент диффузии равен  $D = 10^2 \text{ см}^2/\text{с}$ . Соответственно, диффузионная длина равна  $L_D = 10^{-2} \text{ см}$ .

**Замечание.** Выражение для диффузионной длины  $L_D$  можно получить, используя простые качественные соображения. Обозначим через  $X$  полную длину, которую проходит частица за время  $t$  при диффузии:

$$X = \sum_i X_i. \quad (7.75)$$

Здесь  $X_i$  — длина элементарного акта, то есть длина свободного пробега. Если мы усредним величину  $X$ , то среднее значение равно нулю, так как числа перемещений, которые совершает частица вперед и назад, равны и  $\langle X \rangle = 0$ . Поэтому усреднять нужно не величину  $X$ , а величину  $X^2$ ; в этом случае  $\langle X^2 \rangle \neq 0$ . При усреднении  $\langle X^2 \rangle$  получится сумма квадратов ( $\langle X_i^2 \rangle$ ) и произведений типа  $\langle X_i X_j \rangle$ . Если перемещения равноправны и нескоррелированы, то  $\langle X_i X_j \rangle = \langle X_i \rangle \langle X_j \rangle = 0$ . Поэтому

$$\langle X^2 \rangle = \sum_i \langle X_i^2 \rangle. \quad (7.76)$$

Если между перемещениями есть корреляция, то  $\langle X^2 \rangle \neq \sum_i \langle X_i^2 \rangle$ , бу-

дут еще и удвоенные произведения. Таким образом,  $\langle X^2 \rangle$  имеет смысл среднеквадратичного перемещения за время  $t$ . В нашем случае все  $\langle X_i^2 \rangle$  одинаковы и мы их обозначим через  $l^2$  ( $l$  — это длина свободного пробега частицы). Количество перемещений частицы за время  $t$ , то есть количество членов в сумме (7.76), равно  $t/\tau$ , где  $\tau$  — время свободного пробега. Тогда для среднеквадратичного перемещения частицы за время  $t$  получаем

$$\langle X^2 \rangle = \frac{l^2}{\tau} t. \quad (7.77)$$

Величина  $l^2/\tau$  называется коэффициентом диффузии  $D$ . В результате мы получили

$$\sqrt{\langle X^2 \rangle} = \sqrt{Dt}. \quad (7.78)$$

Оказывается, что смещение частицы при диффузии пропорционально не времени  $t$ , а корню квадратному из времени  $\sqrt{t}$ . Учитывая эти рассуждения, мы можем теперь найти диффузионную длину, то есть

то расстояние, которое проходит частица при диффузии за время жизни  $\tau_{ж}$ :

$$L_D = \sqrt{D_a \tau_{ж}}. \quad (7.79)$$

**Эксперимент Шокли.** Кратко рассмотрим один из первых экспериментальных методов измерения времени дрейфа носителей заряда, а следовательно, и подвижности. Рассмотрим полупроводник, к которому прикладывалось внешнее электрическое поле (рис. 7.12). С помощью короткого импульса света, пропущенного через узкую щель, в узкой области создавались избыточные носители заряда (точка

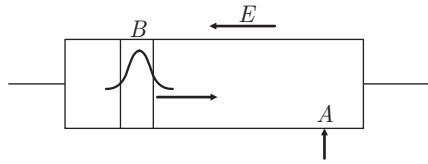


Рис. 7.12. Измерение подвижности избыточных носителей заряда в эксперименте Шокли. В точке А, где расположен  $p-n$ -переход, регистрируется ток неосновных носителей, которые дрейфуют в электрическом поле от точки В

В на рис. 7.12) и возникал волновой пакет. В электрическом поле пакет дрейфует слева направо. В точке А на рис. 7.12 расположен  $p-n$ -переход, к которому ведет зонд. Так как  $p-n$ -переход очень чувствителен к неосновным носителям, то зонд в точке А будет регистрировать неосновные носители. Когда пакет достигает точки А зонда, возникает всплеск тока. Такой эксперимент дает воз-

можность найти подвижность носителей. Для проведения такого эксперимента необходимо соблюдать следующие условия: а) расстояние ВА должно быть меньше дрейфовой длины  $L_E$ ; б) энергия кванта света  $\hbar\omega$  должна быть больше ширины запрещенной зоны полупроводника  $\hbar\omega > E_g$ ; в) время вспышки (импульса света) должно быть меньше времени жизни  $\tau_{ж}$  и времени дрейфа; г) диффузионная длина  $L_D$  должна быть больше дрейфовой длины  $L_E$ :  $L_D > L_E > l_{AB}$ . В настоящее время разработаны и другие методики по измерению подвижности.

## 7.9. Неустойчивость тока в полупроводниках

**7.9.1.  $N$ -образная характеристика. Эффект Ганна.** Рассмотрим зависимость плотности тока  $j$  от локального электрического поля  $E$ , представленную на рис. 7.13. По аналогии с зависимостями тока от напряжения такие зависимости называют дифференциальными (или локальными) вольт-амперными характеристиками. Напомним, что при наличии электрического поля, плотность тока можно записывать в виде

$$j = enbE, \quad (7.80)$$

где  $n$  — концентрация носителей;  $\mu$  — подвижность. То, что ток  $j$  (рис. 7.13) зависит от поля  $E$  нелинейно, означает, что: а) либо концентрация электронов зависит от поля  $E$ ; б) либо подвижность  $b$  зависит

от  $E$ ; в) либо  $n$  и  $b$  одновременно зависят от  $E$ . Из рис. 7.13 следует, что существует такой интервал полей  $E$ , в котором  $\partial j / \partial E < 0$ , то есть дифференциальная проводимость отрицательна (ОДП). Одна из причин

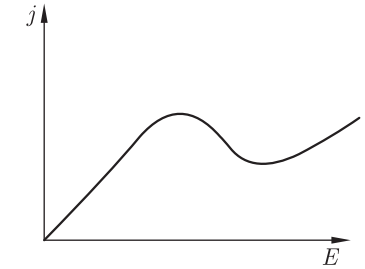


Рис. 7.13. Зависимость плотности тока  $j$  от электрического поля  $E$  для однородного полупроводника с отрицательной дифференциальной проводимостью  $N$ -типа

ОДП — это зависимость концентрации носителей заряда от электрического поля  $E$ . Как уже было отмечено выше, концентрация устанавливается в результате двух процессов — процессов генерации и рекомбинации. В термодинамическом равновесии эти два процесса уравниваются. Если имеет место разогрев электронов, то это может сильно влиять на процесс рекомбинации (например, если примесный центр отталкивающий для электрона, а энергия электрона растет, то он может все-таки захватиться на центр). Так как скорость рекомбинации при разогреве может увеличиваться, то это может привести к изменению концентрации, а тогда падает разогрев и мы получаем отрицательную дифференциальную проводимость (ОДП).

Но наиболее интересный случай ОДП — это эффект Ганна. Этот эффект был открыт американским физиком Ганном в 1963 г. при исследовании полупроводника GaAs в сильном электрическом поле. Рассмотрим качественно суть эффекта Ганна. Он связан со специфической зонной структурой GaAs. В полупроводнике GaAs основная долина для электронов расположена в точке  $\Gamma$  зоны Бриллюэна ( $k = 0$ ) (рис. 7.14). Вторая, боковая, долина расположена в точке  $L$  и чуть выше по энергии, чем долина в точке  $\Gamma$ , — на величину  $\Delta E$  (рис. 7.14). Когда происходит разогрев электронов в  $\Gamma$ -долине в электрическом поле, их энергия растет, и они начинают переходить в побочную долину  $L$ . В центральной долине  $\Gamma$  эффективная масса электрона очень мала; а в побочной долине  $L$  эффективная масса электронов, наоборот, очень велика. Поэтому с увеличением разогрева (то есть с ростом электрического поля  $E$ ) электроны уходят из основной долины  $\Gamma$  в побочную долину  $L$  и их скорость резко падает. Полное число электронов во всех долинах не меняется. Электроны в электрическом поле перераспределяются по долинам. Эта ситуация соответствует падающему участку на рис. 7.13. Второй линейный участок на рис. 7.13 означает, что электроны перешли во вторую долину  $L$  и ток опять начинает расти с ростом напряженности электрического поля  $E$ .

Рассмотрим, что произойдет в полупроводнике, если электрическое поле  $E$  таково, что мы находимся на участке ОДП. Пусть наш образец имеет размер  $d$ , а приложенное напряжение  $V$  таково, что поле  $E = V/d$  соответствует участку с  $E_1 < E < E_2$  (рис. 7.15). Состояние полупро-

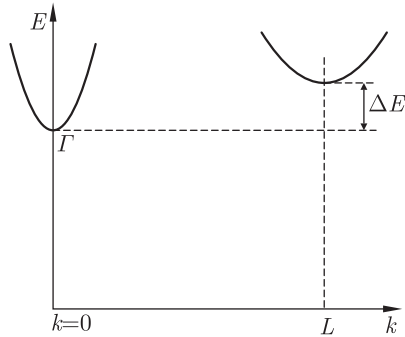


Рис. 7.14. Схематическая зависимость энергии  $E$  от волнового вектора  $k$  для зоны проводимости GaAs. В центральной долине (Г) эффективная масса электрона мала ( $m_{\Gamma} \approx 0,07m_0$ ), а в боковой долине эффективная масса велика ( $m_L \approx 1,2m_0$ )

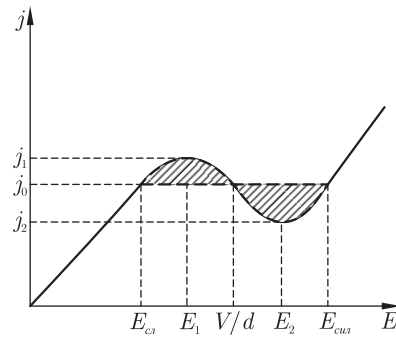


Рис. 7.15. Вольтамперная характеристика  $N$ -типа. Плотность тока является однозначной функцией поля, а поле неоднозначно: функция  $E(j)$  в области значений  $j_2 < j < j_1$  имеет три ветви. Значению плотности тока  $j = j_0$  отвечают три значения напряженности поля:  $E_{сл}$ ,  $V/d$  и  $E_{сил}$ . При  $E = E_{сл}$  и  $E = E_{сил}$  дифференциальная проводимость положительна; при  $E = V/d$  — отрицательна

водника с таким полем неустойчиво, поскольку оно соответствует отрицательной дифференциальной проводимости (см. рис. 7.13). Пример устойчивого и неустойчивого состояния дается рис. 7.16,  $a, б$  — это шарик на впадине и на горке. Состояние 7.16,  $a$  устойчиво, а 7.16,  $б$  — неустойчиво, хотя оба они стационарны. Устойчивость определяется тем, что будет происходить с системой при приложении смещения. Неустойчивость проявляется в том, что полупроводник расслаивается (рис. 7.17), возникают домены (области) сильного поля. Поле в домене соответствует значению  $E_{сил}$ , когда дифференциальная проводимость положительна (рис. 7.15). В остальной части образца (вне домена) поле меньше  $E_1$ , так что там дифференциальная проводимость снова является положительной (рис. 7.15). Таким образом, в этой части образца опять достигается устойчивость. Размер домена «выбирается» так, чтобы сумма напряжений, приходящаяся на домен  $x \cdot E_{сил}$  и на остальную область  $(d - x)E_{сл}$ , равнялась приложенному напряжению  $V$ :

$$(d - x)E_{сл} + x \cdot E_{сил} = V. \quad (7.81)$$

Если известны поля  $E_{сл}$  и  $E_{сил}$ , можно найти размер домена  $x$ . Ток в домене и вне домена одинаков. Согласно закону сохранения заряда в любом сечении образца ток должен быть одинаковым (это дает два условия при решении задачи). Есть еще одно условие — условие стационарного баланса диффузионных и дрейфовых токов в стенке



Рис. 7.16. Примеры устойчивого  $a)$  (шарик на впадине) и неустойчивого  $б)$  (шарик на горке) стационарных состояний

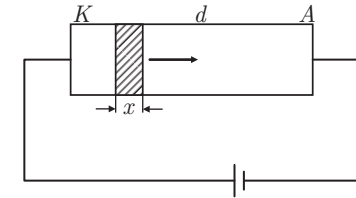


Рис. 7.17. Формирование движущегося домена сильного поля. Скорость домена равна дрейфовой скорости основных носителей заряда в слабом поле  $E_{сл}$  (соответствующем незаштрихованным областям на рисунке)

домена, разделяющей однородные состояния  $E = E_{сл}$  и  $E = E_{сил}$ . Это условие можно представить как «правило равных площадей»: площадь над прямой тока ( $j = j_0$ ) равна площади под прямой тока (рис. 7.15, заштрихованные области). Три условия позволяют решить задачу теоретически и найти три неизвестных величины: размер домена  $x$ ,  $E_{сл}$  и  $E_{сил}$ .

Домен движется в ту сторону, куда движутся электроны, и он движется со скоростью дрейфа, которая определяется слабым полем  $E_{сл}$ . Это легко доказать. Когда домен движется и выходит к аноду, то в системе исчезает область повышенного сопротивления, благодаря чему происходит сильное изменение тока, то есть всплеск тока. Это приводит к тому, что, как правило, на катоде снова зарождается новый домен, далее он доходит до анода и ток в системе растет, и все повторяется с начала. Поэтому в полупроводнике возникают колебания тока. Период колебаний тока определяется временем движения домена. Это есть размер образца  $d$ , деленный на дрейфовую скорость домена  $bE_{сл}$ ,

$$T = \frac{d}{bE_{сл}}. \quad (7.82)$$

Таким образом можно получить высокочастотные колебания (то есть колебания сантиметрового диапазона). В GaAs  $n$ -типа частота колебаний  $\omega = 2\pi bE_{сл}/d$  составляет примерно  $10^9 - 10^{10} \text{ с}^{-1}$  и выше в зависимости от длины образца. Возникновение таких колебаний в  $n$ -GaAs и подобных ему материалах называется эффектом Ганна. Эффект Ганна используется в полупроводниковой электронике для создания генераторов и усилителей СВЧ колебаний и ряда других приборов.

В однородном полупроводнике домен рождается на катоде. В неоднородном полупроводнике возможно рождение домена на неоднородностях, причем рождение неоднородного домена может происходить раньше, чем исчезнет первоначальный домен.

**7.9.2. S-образная характеристика. Тепловой пробой. Шнурование тока.** Рассмотрим качественно другой тип зависимости плотности тока от напряженности электрического поля, так называемую S-образную характеристику (рис. 7.18). Такая характеристика может возникать при разных условиях. Рассмотрим два возможных явления в полупроводниках, когда возникает S-образная зависимость плотности тока  $j$  от поля  $E$ .

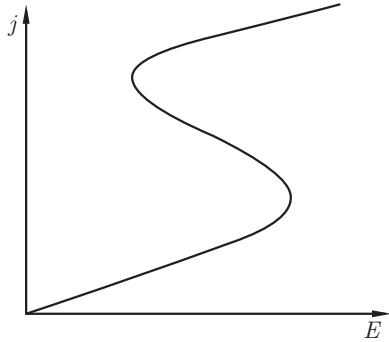


Рис. 7.18. Зависимость плотности тока  $j$  от электрического поля  $E$  для однородного полупроводника с отрицательной дифференциальной проводимостью S-типа

**а) Тепловой пробой.** Пусть по полупроводнику течет ток и выделяется джоулево тепло. Если это тепло не выводится наружу (из системы) достаточно быстро, то полупроводник нагревается. Тогда, чем больше ток, тем больше джоулев разогрев, а следовательно, тем больше концентрация носителей. Напряженность электрического поля выражается через плотность тока и концентрацию и имеет вид

$$E = \frac{j}{en\mu}. \quad (7.83)$$

При джоулевом разогреве концентрация может расти с ростом температуры быстрее, чем плотность тока, тогда поле  $E$  падает и мы получаем S-образную характеристику.

**б) Инжекция носителей из контакта. Шнурование тока.** В случае инжекции носителей из контакта механизм S-образной зависимости тока от поля аналогичен случаю теплового пробоя; по-прежнему концентрация носителей растет быстрее, чем плотность тока. Тогда, согласно (7.83), напряженность поля падает, и мы получаем зависимость тока от поля, представленную на рис. 7.18.

Что же реально в этом случае происходит в полупроводнике? Когда мы задаем плотность тока, то мы тем самым «навязываем» полупроводнику такие условия (см. горизонтальную линию тока на рис. 7.19). На рис. 7.19 точка «1» на участке отрицательного дифференциального сопротивления ВАХ неустойчива. В этом случае однородное стационарное состояние неустойчиво по отношению к образованию неоднородного распределения плотности тока в поперечном сечении полупроводника, когда возникает центральная цилиндрическая область с высокой

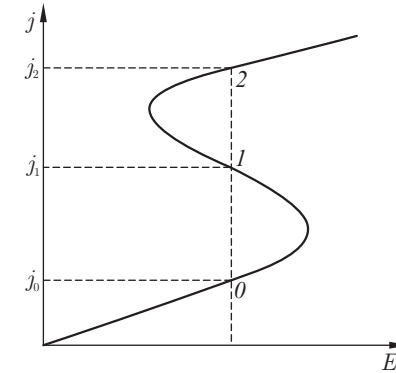


Рис. 7.19. Вольтамперная характеристика S-типа. Плотность тока не является однозначной функцией поля: в области  $E_{\text{сл}} < E < E_{\text{сил}}$  каждому значению напряженности поля отвечают три возможных значения плотности тока. Два из них ( $j_0$  и  $j_2$ ) находятся на возрастающих ветвях ВАХ (где дифференциальная проводимость положительна), одно ( $j_1$ ) — на падающей (где дифференциальная проводимость отрицательна)

плотностью тока («шнуры» тока) и окружающая ее область с низкой плотностью тока (рис. 7.20). Такое явление называется шнурованием тока. Внутри шнура плотность тока повышена, но этот ток находится на зависимости в той области, где дифференциальное сопротивление положительно, (точка «2» на рис. 7.19). Область полупроводника вне шнура тоже соответствует положительному дифференциальному сопротивлению (точка «0» на рис. 7.19). Здесь в отличие от N-образной характеристики напряженность электрического поля внутри шнура и вне его — одна и та же величина. Суммарная плотность тока внутри шнура и вне шнура равна величине внешнего тока ( $jS$ ):

$$j_1 S_1 + j_2 S_2 = jS. \quad (7.84)$$

Здесь  $S$  — площадь сечения полупроводника;  $S_2$  — площадь сечения шнура;  $S_1$  — площадь сечения остальной части полупроводника. Здесь зависимость  $j(E)$  (рис. 7.19) сходна с кривой Ван-дер-Ваальса для конденсации пара в жидкость. Однако эта аналогия неполна. В частности,

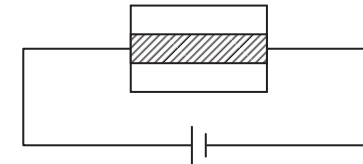


Рис. 7.20. Формирование шнура тока высокой плотности. Область, заштрихованная на рисунке, изображает сечение шнура цилиндрической формы. Именно такая конфигурация является более устойчивой



напряжение на шнуре тока (см. рис. 7.19) нельзя вычислить с помощью правила «равных площадей» для ВАХ, как в эффекте Ганна (рис. 7.15). Отметим, что шнуры тока хорошо изучены в газовом разряде<sup>1)</sup>.

**Замечания.** Условия образования той или иной ВАХ можно найти с помощью явного выражения для величины дифференциальной проводимости  $\sigma_d$ . Действительно, из рис. 7.15 и 7.19 следует, что в случае ВАХ  $N$ -типа дифференциальная проводимость меняет знак, проходя через нуль, а в случае характеристики  $S$ -типа дифференциальная проводимость меняет знак, проходя через бесконечность. Таким образом, условия возникновения падающих участков с образованием ВАХ  $N$ - и  $S$ -типа можно записать в явном виде для конкретных полупроводников.

### Задачи

**Задача 7.1.** Найти отношение концентраций электрона в побочной долине  $L$  к основной долине  $\Gamma$  ( $n_L/n_\Gamma$ -?).

*Решение.* Аналогично задаче 1 главы VI, имеем:

$$n_\Gamma = N_c^\Gamma \exp\left(-\frac{\chi_c^\Gamma}{k_B T_e^\Gamma}\right),$$

$$n_L = N_c^L \exp\left(-\frac{\chi_c^L}{k_B T_e^L}\right),$$

где  $\chi_c^L = \chi_c^\Gamma + \Delta E$ ,  $\chi_c^\Gamma$  — расстояние от уровня химического потенциала до дна зоны проводимости.

$$k_B T_e^i = k_B T + m^i (\mu^i E)^{2\alpha},$$

где  $i = \Gamma, L$ ;

$$N_c^i = 2 \left( \frac{m_c^i T_e^i}{2\pi \hbar^2} \right)^{3/2}.$$

Получаем:

$$\frac{n_L}{n_\Gamma} = \left( \frac{m_c^L T_e^L}{m_c^\Gamma T_e^\Gamma} \right) \exp\left(-\frac{\Delta E}{k_B T_e^L} + \frac{\chi_c^\Gamma}{k_B} \left( \frac{1}{T_e^\Gamma} - \frac{1}{T_e^L} \right)\right).$$

<sup>1)</sup> См. Бонч-Бруевич В. Л., Калашников С. Г. Физика полупроводников. — М.: Наука, 1977. 679 с.

## Глава 8

### ОПТИЧЕСКИЕ СВОЙСТВА ПОЛУПРОВОДНИКОВ

#### 8.1. Введение. Прямые и не прямые оптические переходы

Полупроводники обладают важным свойством — они способны поглощать электромагнитное излучение, падающее извне.

Рассмотрим полупроводник при очень низких температурах, когда свободных носителей заряда не существует. За счет каких процессов возможно поглощение электромагнитной волны? Существуют два механизма поглощения излучения (света) при таких условиях: а) поглощение света с возбуждением колебаний решетки (фононов) и б) поглощение света с возбуждением электронов.

Кратко рассмотрим первый механизм поглощения. Электромагнитная волна при поглощении может возбудить в полупроводнике фононы. Электромагнитная волна взаимодействует, в основном, с оптическими фононами. В главе 3 показано, что оптические фононы соответствуют колебаниям дипольного момента системы атомов. Электромагнитная волна, взаимодействуя с полупроводником, раскачивает дипольные (оптические) колебания и поглощается. Но, как уже было показано в главе 3, энергия фонона, например для GaAs, имеет порядок величины  $\hbar\omega_0 \approx 3 \cdot 10^{-2}$  эВ. Такая величина энергии фонона соответствует инфракрасной области спектра. Поэтому колебания кристаллической решетки (то есть оптические фононы) возбуждаются тогда, когда мы будем облучать полупроводник электромагнитным излучением инфракрасного диапазона. Таким образом, рассмотренный нами первый механизм поглощения излучения существенен только для инфракрасной области спектра.

Рассмотрим второй механизм — поглощение электромагнитного излучения за счет возбуждения электронов (рис. 8.1). В чистом (без примесей) полупроводнике (а мы рассматриваем именно такой случай) поглощение излучения и возбуждение электрона возможно только при переводе электрона из валентной зоны в зону проводимости. Это реализуется, только если энергия кванта света  $\hbar\omega \gtrsim E_g$ , где  $E_g$  — ширина запрещенной зоны полупроводника. Для света с энергией кванта  $\hbar\omega < E_g$  полупроводник прозрачен, то есть не поглощает такое излучение. На рис. 8.2 приведена зависимость коэффициента поглощения излучения

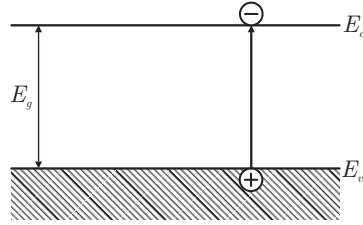


Рис. 8.1. Энергетическая диаграмма полупроводникового кристалла. Оптический переход электрона из валентной зоны в зону проводимости показан стрелкой

электронами  $\alpha$  от энергии кванта  $\hbar\omega$ . При  $\hbar\omega < E_g$  коэффициент поглощения равен нулю ( $\alpha = 0$ ), при  $\hbar\omega > E_g$  коэффициент поглощения растет. Точка  $\hbar\omega = E_g$  называется краем фундаментального поглощения излучения<sup>1)</sup>.

Нашей задачей является теоретическое описание зависимости коэффициента поглощения  $\alpha$  от частоты, см. рис. 8.2. Процесс поглощения света (рис. 8.1) описывается двумя основными законами: 1) законом сохранения квазиимпульса; 2) законом сохранения энергии. Рассмотрим подробнее, что означают эти законы для случая поглощения света электронами.

Мы для простоты рассматриваем монохроматическое излучение, имеющее определенную частоту  $\omega$ , то есть одну гармонику (много гармоник, то есть полосу излучения, мы не рассматриваем). Если энергия кванта  $\hbar\omega < E_g$ , то процесс погло-

щения невозможен. Состояние электронов в валентной зоне характеризуется точками на кривой  $E_v(k)$  (рис. 8.3). Рассмотрим переходы электронов из валентной зоны в зону проводимости при поглощении кванта света  $\hbar\omega$ . В прямозонном полупроводнике с точки зрения закона сохранения энергии один и тот же квант света может поглощаться при разных переходах (см. переходы «1» и «2» на рис. 8.3). Однако, как будет показано ниже (а именно это и наблюдается экспериментально), такие переходы («1» и «2» на рис. 8.3) запрещены из-за нарушения закона сохранения квазиимпульса. Пусть до поглощения квазиимпульс электрона равен  $\mathbf{k}$  (здесь мы положили  $\hbar = 1$ ), а импульс фотона равен  $\mathbf{\kappa}$ . После поглощения фотона мы имеем в зоне проводимости только

<sup>1)</sup> О том, какая зависимость  $\alpha(\hbar\omega)$  наблюдается экспериментально, будет говориться ниже в § 8.2.

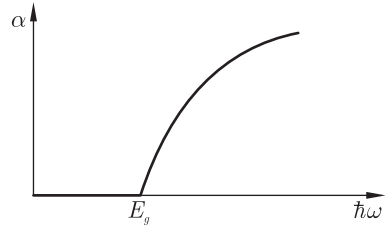


Рис. 8.2. Зависимость коэффициента поглощения света, связанного с вертикальными переходами, от энергии фотона  $\hbar\omega$

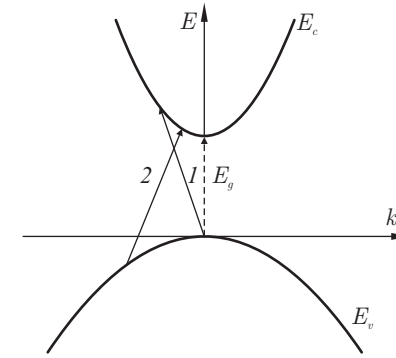


Рис. 8.3. Схема непрямых переходов в прямозонном полупроводнике. Показанные стрелкой переходы «1» и «2» запрещены законом сохранения квазиимпульса

электрон с квазиимпульсом  $\mathbf{k}'$ . Тогда закон сохранения квазиимпульса для такого процесса поглощения имеет вид

$$\mathbf{\kappa} + \mathbf{k} = \mathbf{k}'. \quad (8.1)$$

Таким образом, квазиимпульс возбужденного электрона  $\mathbf{k}'$  отличается от квазиимпульса первоначального электрона  $\mathbf{k}$  на величину импульса фотона  $\mathbf{\kappa}$ . Напомним, что волновой вектор частицы связан с длиной волны  $\lambda$  соотношением  $\kappa = 2\pi/\lambda$ . Длина волны излучения, соответствующая инфракрасной области излучения,  $\lambda \sim 10^{-4}$  см. Тогда волновой вектор фотона имеет величину порядка  $\kappa \sim 10^5$  см<sup>-1</sup>. Сравним волновой вектор фотона с волновым вектором электрона. На границе зоны Бриллюэна квазиволновой вектор электрона порядка  $k_{\text{пред}} \sim 10^8$  см<sup>-1</sup>, следовательно,  $\kappa/k_{\text{пред}} \sim 10^{-3} \ll 1$ . Таким образом, в соотношении (8.1) можно пренебречь  $\kappa$ , и тогда закон сохранения квазиимпульса для процессов поглощения света принимает вид

$$\mathbf{k} = \mathbf{k}'. \quad (8.2)$$

Соотношение (8.2) означает, что при поглощении кванта света квазиволновые вектора электронов в начальном и конечном состоянии должны совпадать. А это означает, что стрелки («1» и «2»), обозначающие переходы на диаграмме рис. 8.3 должны быть вертикальными (рис. 8.4). Поэтому, если нам задана энергия кванта  $\hbar\omega$ , то мы можем точно указать, какой электрон из валентной зоны переходит в зону проводимости. Для этого мы на шкале энергии должны сопоставить энергии  $\hbar\omega$  стрелку «↑» и показать это на диаграмме (рис. 8.4). Там, где начало и конец стрелки совпадают с  $E_v(k)$  и  $E_c(k)$  (при одном и том же  $k$ , то есть вертикально), такой электрон перейдет в зону проводимости (рис. 8.4). Такие переходы называются прямыми. Край поглощения равен  $E_g$ .



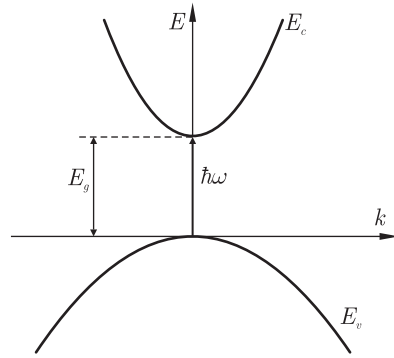


Рис. 8.4. Вертикальный переход в прямозонном полупроводнике. После поглощения фотона с энергией  $\hbar\omega$  электрон переходит из валентной зоны в зону проводимости. При этом волновой вектор электрона в начальном и в конечном состояниях одинаков

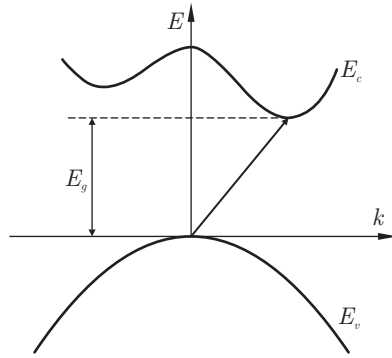


Рис. 8.5. Непрямые (невертикальные) оптические переходы в непрямоzonном полупроводнике. Непрямые переходы возможны при участии, кроме электрона и фотона, третьего тела (например, фонона)

На рис. 8.5 представлена зонная диаграмма непрямоzonного полупроводника. Для такого случая квант света с энергией  $\hbar\omega = E_g$  прямо не поглощается. Такие оптические переходы неэффективны, для них необходимо наличие (кроме электрона и фотона) третьего тела (например, фонона). Вероятность таких не прямых переходов очень мала<sup>1)</sup>. Итак, оптические переходы с энергией кванта  $\hbar\omega = E_g$  в непрямоzonных полупроводниках (рис. 8.5) не могут идти, так как они запрещены законом сохранения квазиимпульса. В этом случае край собственного поглощения не совпадает с  $E_g$ .

## 8.2. Квантовомеханический подход к задаче о поглощении света

Для вычисления коэффициента поглощения, связанного с междозонными переходами, удобно пользоваться квантовомеханическим подходом. Будем исходить из уравнения Шредингера для электрона в поле световой (электромагнитной) волны. Кроме того, мы покажем, как в квантовой механике возникает закон сохранения квазиимпульса для процесса поглощения света.

Мы выбираем такую калибровку, что скалярный потенциал  $\varphi$  электромагнитного поля равен нулю ( $\varphi = 0$ ), существует только векторный потенциал  $\mathbf{A}(\mathbf{r}, t)$ . Тогда гамильтониан электрона с зарядом  $e$  в пе-

риодическом поле кристалла  $U(\mathbf{r})$  и электромагнитном поле световой волны имеет вид

$$\hat{H} = \frac{1}{2m_0} \left( \hat{\mathbf{P}} - \frac{e}{c} \mathbf{A} \right)^2 + U(\mathbf{r}), \quad (8.3)$$

где  $m_0$  — масса свободного электрона;  $\hat{\mathbf{P}}$  — оператор импульса,  $\hat{\mathbf{P}} = -i\hbar\nabla$ . Поскольку векторный потенциал зависит от времени, мы должны воспользоваться нестационарным уравнением Шредингера

$$i\hbar \frac{\partial \psi}{\partial t} = \hat{H} \psi. \quad (8.4)$$

Далее будем считать, что векторный потенциал зависит от времени и координаты гармонически (по гармоническому закону)<sup>1)</sup>

$$\mathbf{A} = \mathbf{A}_0 \exp[i(\omega t - \mathbf{\kappa} \mathbf{r})], \quad (8.5)$$

где  $\mathbf{A}_0$  — амплитуда векторного потенциала. Когда падающий свет является слабым (его интенсивность мала), мы можем рассматривать слагаемое, пропорциональное  $\mathbf{A}$  в (8.3), как возмущение и решать задачу о поглощении света по теории возмущений. Тогда в (8.3) при возведении скобки в квадрат можно ограничиться линейными по  $\mathbf{A}$  членами. Кроме того, векторный потенциал  $\mathbf{A}$  зависит от координаты  $\mathbf{r}$  и, вообще говоря, не коммутирует с оператором импульса  $\hat{\mathbf{P}}$ . Однако для поперечной волны  $\text{div } \mathbf{A} = 0$ , и тогда можно показать, что  $\mathbf{A}$  и  $\hat{\mathbf{P}}$  коммутируют<sup>2)</sup>:

$$\mathbf{A} \hat{\mathbf{P}} = \hat{\mathbf{P}} \mathbf{A}. \quad (8.6)$$

Таким образом, гамильтониан задачи (8.3) можно переписать в виде

$$\hat{H} = \frac{\hat{\mathbf{P}}^2}{2m_0} + U(\mathbf{r}) + \hat{V}. \quad (8.7)$$

Здесь первые два члена соответствуют гамильтониану  $\hat{H}_0$  свободного электрона в периодическом поле кристалла, а слагаемое

$$\hat{V} = -\frac{e}{m_0 c} \mathbf{A} \hat{\mathbf{P}} \quad (8.8)$$

— это возмущение к  $\hat{H}_0$ , связанное с действием световой волны.

Теперь мы можем вычислить вероятность перехода электрона из валентной зоны в зону проводимости при поглощении кванта света  $\hbar\omega$ . Согласно «золотому правилу» квантовой механики число переходов

<sup>1)</sup> Заметим, что на самом деле  $\mathbf{A}$  нужно писать в вещественной форме, то есть  $\mathbf{A} = \frac{1}{2} [\mathbf{A}_0 \exp(i(\omega t - \mathbf{\kappa} \mathbf{r})) + \mathbf{A}_0^* \exp(-i(\omega t - \mathbf{\kappa} \mathbf{r}))]$ . Однако можно показать, что для наших целей достаточно выражения (8.5).

<sup>2)</sup> Коммутатор  $[\mathbf{A}, \hat{\mathbf{P}}] = i\hbar \text{div } \mathbf{A}$ .

<sup>1)</sup> Такие (непрямые) переходы будут детально рассмотрены в § 8.4.

в единицу времени в единицу объема, для которых выполняется закон сохранения энергии  $\varepsilon_2 = \varepsilon_1 + \hbar\omega$ , равно <sup>1)</sup>

$$W_{12} = \frac{2\pi}{\hbar} |\hat{V}_{12}|^2 \delta(\varepsilon_1 + \hbar\omega - \varepsilon_2). \quad (8.9)$$

Здесь  $\varepsilon_1$  — энергия электрона в начальном состоянии (в валентной зоне);  $\varepsilon_2$  — энергия электрона в конечном состоянии (в зоне проводимости);  $\hat{V}_{12}$  — матричный элемент перехода:

$$\hat{V}_{12} = -\frac{e}{m_0 c} \int \psi_1^*(\mathbf{r}) \mathbf{A} \hat{\mathbf{P}} \psi_2(\mathbf{r}) d^3 r. \quad (8.10)$$

Итак, закон сохранения энергии в (8.9) выражается через  $\delta$ -функцию. Возникает вопрос, где в (8.9) фигурирует закон сохранения импульса. Ниже мы докажем, что матричный элемент перехода  $\hat{V}_{12}$  в (8.10) равен нулю, если не выполняется закон сохранения импульса. Далее мы опускаем временной множитель у  $\mathbf{A}$ , так как от времени в (8.10) ничего больше не зависит. Поэтому при возведении в квадрат  $|\hat{V}_{12}|^2$  экспонента выпадает. Тогда векторный потенциал можно взять в виде

$$\mathbf{A} = \mathbf{A}_0 \exp(-i\boldsymbol{\kappa}\mathbf{r}).$$

Состояния электрона в зоне проводимости и валентной зоне описываются блоховскими волновыми функциями

$$\psi_1(\mathbf{r}) = e^{i\mathbf{k}_1\mathbf{r}} U_{v\mathbf{k}_1}(\mathbf{r}), \quad (8.11)$$

$$\psi_2(\mathbf{r}) = e^{i\mathbf{k}_2\mathbf{r}} U_{c\mathbf{k}_2}(\mathbf{r}). \quad (8.12)$$

Здесь « $v$ » обозначает валентную зону, а « $c$ » — зону проводимости,  $\mathbf{k}_1$  — квазиволновой вектор электрона в валентной зоне,  $\mathbf{k}_2$  — квазиволновой вектор электрона в зоне проводимости,  $U_{v\mathbf{k}_1}(\mathbf{r})$  и  $U_{c\mathbf{k}_2}(\mathbf{r})$  — блоховские амплитуды, относящиеся к валентной зоне и зоне проводимости соответственно, причем каждая из них является периодической функцией с периодом решетки. Подставив явные выражения для  $\mathbf{A}$ ,  $\psi_1$  и  $\psi_2$  в матричный элемент  $\hat{V}_{12}$ , получаем

$$\hat{V}_{12} = \frac{i\hbar e}{m_0 c} \mathbf{A}_0 \int \exp\left[-i(\mathbf{k}_1 + \boldsymbol{\kappa})\mathbf{r}\right] U_{v\mathbf{k}_1}(\mathbf{r}) \nabla (e^{i\mathbf{k}_2\mathbf{r}} U_{c\mathbf{k}_2}(\mathbf{r})) d^3 r. \quad (8.13)$$

Интеграл в этом выражении берется по всему объему кристалла. Покажем, что этот интеграл равен нулю, если не выполняется закон сохранения квазиимпульса (8.1). Кристалл состоит из элементарных ячеек (рис. 8.6). Разобьем интегрирование по всему кристаллу на интегралы в пределах элементарных ячеек. Рассмотрим произвольную точку в  $n$ -й

<sup>1)</sup> В (8.9) мы предполагаем, что в начальный момент времени состояние  $\varepsilon_1$  занято электроном, а  $\varepsilon_2$  — свободно. Тогда процессы излучения фотонов можно не рассматривать.

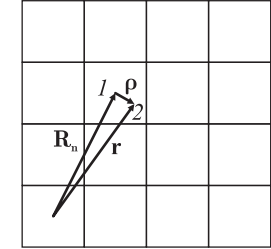


Рис. 8.6. К выводу выражения (8.15) для матричного элемента  $V_{12}$ . Вектор  $\mathbf{R}_n$  соединяет центр нулевой ячейки (которая принята за начало координат) и центр  $n$ -й ячейки. Для вектора  $\mathbf{r}$  можно записать:  $\mathbf{r} = \mathbf{R}_n + \boldsymbol{\rho}$

элементарной ячейке (точка «1», рис. 8.6) и точку в центре  $n$ -й ячейки (точка «2»). Тогда вектор  $\mathbf{r}$ , идущий из начала координат  $O$  в точку «1» (точка  $O$  — центр элементарной ячейки), мы можем выразить через вектор, идущий из начала координат в центр  $n$ -й ячейки  $\mathbf{R}_n$ :

$$\mathbf{r} = \mathbf{R}_n + \boldsymbol{\rho}. \quad (8.14)$$

Если учесть (8.14) и подставить это в  $\hat{V}_{12}$ , то интеграл в (8.13) можно разбить на сумму интегралов, каждый из которых вычисляется в пределах элементарной ячейки:

$$\hat{V}_{12} = \frac{i\hbar e}{m_0 c} \mathbf{A}_0 \sum_n \exp[i(\mathbf{k}_2 - \mathbf{k}_1 - \boldsymbol{\kappa})\mathbf{R}_n] \times \int_{\text{по } n\text{-й ячейке}} \exp[-i(\mathbf{k}_1 + \boldsymbol{\kappa})\boldsymbol{\rho}] U_{v\mathbf{k}_1}(\boldsymbol{\rho}) \nabla (e^{i\mathbf{k}_2\boldsymbol{\rho}} U_{c\mathbf{k}_2}(\boldsymbol{\rho})) d^3 \rho. \quad (8.15)$$

В (8.15) мы учли, что блоховские функции — периодические функции координат с периодом кристалла ( $\mathbf{R}_n$  — векторы элементарной ячейки). Оператор  $\nabla$  действует на переменную  $\boldsymbol{\rho}$ . Интеграл в (8.15) от  $n$  не зависит, это означает, что пропала индивидуальность элементарной ячейки, тогда этот интеграл одинаков для каждой ячейки кристалла. Таким образом, интеграл в (8.15) можно вынести из-под знака суммы. В результате, матричный элемент равен

$$V_{12} = \mathcal{P}_{12} \sum_n e^{i\mathbf{q}\mathbf{R}_n}, \quad (8.16)$$

где  $\mathbf{q} = \mathbf{k}_2 - \mathbf{k}_1 - \boldsymbol{\kappa}$ ,

$$\mathcal{P}_{12} = \frac{i\hbar e}{m_0 c} \mathbf{A}_0 \int \exp[-i(\mathbf{k}_1 + \boldsymbol{\kappa})\boldsymbol{\rho}] U_{v\mathbf{k}_1}(\boldsymbol{\rho}) \nabla (e^{i\mathbf{k}_2\boldsymbol{\rho}} U_{c\mathbf{k}_2}(\boldsymbol{\rho})) d^3 \rho.$$

Покажем, что если  $\mathbf{q} \neq 0$ , то есть если закон сохранения квазиимпульса не выполняется, то сумма по  $n$  в (8.16) равна нулю. Для простоты рассмотрим одномерный случай. Для одномерной цепочки  $R_n = n \cdot a$ , где  $a$  — постоянная решетки. Тогда матричный элемент пропорционален

$$\hat{V}_{12} \sim \sum_n e^{iqna}. \quad (8.17)$$

Сумма берется по всем элементарным ячейкам. При вычислении суммы мы можем воспользоваться хорошо разработанным приемом. Пусть кристалл (то есть цепочка) имеет длину  $L$  (рис. 8.7). Если в (8.17) мы рассмотрим сумму от нуля до  $N - 1$ , тогда мы переберем все  $N$  ячеек этой цепочки. В результате мы получаем сумму геометрической прогрессии

$$\sum_{n=0}^{N-1} e^{iqna} = \sum_{n=0}^{N-1} (e^{iqa})^n = \frac{1 - (e^{iqa})^N}{1 - e^{iqa}} = \frac{1 - e^{iqL}}{1 - e^{iqa}}. \quad (8.18)$$

В силу периодических граничных условий Борна–Кармана все волновые вектора дискретны и все волновые функции выбираются так, чтобы они были периодичны с периодом  $L$ , тогда  $\exp(iqL) = 1$ .

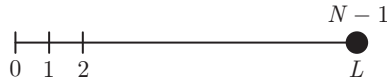


Рис. 8.7. Одномерная цепочка одинаковых атомов конечной длины  $L$

Получаем, что числитель в (8.18) равен нулю, так что равна нулю вся сумма в  $\hat{V}_{12}$ . Но если и знаменатель в (8.18) обращается в нуль, то вычислять сумму таким образом нельзя, поскольку получается неопределенность «ноль на ноль». Раскрывая неопределенность получим, что сумма в (8.17) равна

$$\sum_{n=0}^{N-1} e^{iqna} = \begin{cases} 0, & \text{если } e^{iqa} \neq 1, \\ N, & \text{если } e^{iqa} = 1, \end{cases} \quad (8.19)$$

где  $N$  — полное число элементарных ячеек в нормировочном объеме. Рассмотрим подробнее полученный результат (8.19). Сумма отлична от нуля, если  $e^{iqa} = 1$ . Это означает, что либо  $q = 0$ , либо  $q = (2\pi/a)m$ , где  $(2\pi/a)m$  — вектор обратной решетки. Таким образом, для трехмерного случая можно сформулировать правило: матричный элемент  $\hat{V}_{12}$  (см. (8.16)) не равен нулю, если

$$\mathbf{k}_2 - \mathbf{k}_1 - \boldsymbol{\kappa} = \mathbf{G}, \quad (8.20)$$

где  $\mathbf{G}$  — вектор обратной решетки. То, что закон сохранения импульса выполняется с точностью до вектора обратной решетки означает, что в полупроводнике мы имеем дело не с плоской, а с блоховской волной, при этом  $\mathbf{k}$  — это не волновой вектор электрона, а квазиволновой вектор (квазиимпульс). Решетка может получать и отдавать импульсы, соответствующие векторам обратной решетки. Такие процессы, когда переданный решетке импульс равен вектору обратной решетки, называются процессами переброса.

Для нас при исследовании оптических свойств полупроводника процессы переброса неважны. Мы указали (см. главу 2), что если волновой вектор лежит в первой зоне Бриллюэна и к нему прибавить

вектор обратной решетки, то он выходит за пределы первой зоны Бриллюэна. В закон сохранения квазиимпульса (см. (8.20)) входит разность двух векторов  $\mathbf{k}_2 - \mathbf{k}_1$ , каждый из которых лежит в первой зоне Бриллюэна, поэтому эта разность всегда попадает в первую зону Бриллюэна. Таким образом, разность  $\mathbf{k}_2 - \mathbf{k}_1$  не может быть равна вектору обратной решетки  $\mathbf{G}$ ,  $\mathbf{k}_2 - \mathbf{k}_1 \neq \mathbf{G}$ . Выше также было показано, что волновой вектор фотона  $\boldsymbol{\kappa}$  мал по сравнению с  $\mathbf{k}_1$  и  $\mathbf{k}_2$ . Тогда закон сохранения импульса для оптических переходов принимает вид

$$\mathbf{k}_2 = \mathbf{k}_1. \quad (8.21)$$

Итак, мы получили, что при поглощении света возможны только вертикальные (прямые) переходы электрона из валентной зоны в зону проводимости.

Теперь перейдем к вычислению вероятности оптического перехода электрона из валентной зоны в зону проводимости. Для этого мы воспользуемся выражением (8.9). Так как при оптическом переходе импульс сохраняется, удобно переписать (8.9) в виде

$$W_{\mathbf{k},c\mathbf{k}} = \frac{2\pi}{\hbar} |V_{12}|^2 \delta(E_{c\mathbf{k}} - E_{v\mathbf{k}} - \hbar\omega), \quad (8.22)$$

где  $E_{c\mathbf{k}}$  — энергия конечного состояния электрона;  $E_{v\mathbf{k}}$  — энергия начального состояния. Коэффициент поглощения, обусловленный прямыми межзонными переходами, пропорционален полной вероятности перехода  $W$  электрона из одного состояния в другое. Для определения полного числа переходов в единицу времени  $W$  надо просуммировать  $W_{\mathbf{k},c\mathbf{k}}$  по начальным и конечным состояниям, то есть по всем  $\mathbf{k}$ . Тогда коэффициент поглощения  $\alpha$  равен <sup>1)</sup>

$$\alpha = \frac{W}{Nv} = \frac{1}{Nv} \sum_{\mathbf{k}} W_{\mathbf{k},c\mathbf{k}}. \quad (8.23)$$

Здесь  $N$  — это концентрация фотонов (то есть число фотонов  $1 \text{ см}^3$ ), которая связана с модулем векторного потенциала  $A_0$  как

$$N = \left(\frac{\omega}{v}\right)^2 \frac{|A_0|^2}{2\pi\hbar\omega},$$

$v = c/n$ ,  $c$  — скорость света в вакууме;  $n$  — показатель преломления полупроводника.

В дальнейшем нас будет интересовать зависимость коэффициента поглощения от частоты падающего света  $\omega$ . Для этого вычислим коэффициент поглощения для случая, когда энергия кванта света  $\hbar\omega$

<sup>1)</sup> Здесь мы учли, что число поглощенных в  $1 \text{ см}^2$  в  $1 \text{ с}$  фотонов равно полному числу переходов  $W$ . Коэффициент поглощения  $\alpha$  равен полному числу переходов, деленному на плотность потока фотонов  $Nv$  (то есть число фотонов, проходящих сквозь  $1 \text{ см}^2$  в  $1 \text{ с}$ ).

чуть превышает ширину запрещенной зоны  $E_g$ , то есть здесь мы рассматриваем поглощение света вблизи порога

$$\hbar\omega - E_g \ll E_g. \quad (8.24)$$

В этом случае матричный элемент перехода  $V_{12}$  можно считать константой, не зависящей от волнового вектора  $\mathbf{k}$ . Тогда зависимость коэффициента поглощения от частоты определяется выражением

$$\alpha \approx \frac{2\pi}{\hbar} \frac{|V_{12}|^2}{Nv} \sum_{\mathbf{k}} \delta(E_{c\mathbf{k}} - E_{v\mathbf{k}} - \hbar\omega). \quad (8.25)$$

Далее, нам нужно вычислить сумму по всем  $\mathbf{k}$ . Рассмотрим подробнее, что представляет собой такая сумма. При выбранном значении волнового вектора  $\mathbf{k}$  происходит вертикальный переход из валентной зоны в зону проводимости (то есть поглощение фотона) без изменения волнового вектора электрона (см. рис.8.8). Разность энергий электронов в валентной зоне и в зоне проводимости, то есть  $E_{c\mathbf{k}} - E_{v\mathbf{k}}$ , можно разбить на три части (рис.8.8): 1) ширину запрещенной зоны  $E_g$ ; 2) кинетическую энергию электрона в зоне проводимости  $\varepsilon_c(\mathbf{k})$ ; и 3) кинетическую энергию дырки в валентной зоне  $\varepsilon_h(\mathbf{k})$ , тогда

$$E_{c\mathbf{k}} - E_{v\mathbf{k}} = \varepsilon_c(\mathbf{k}) + \varepsilon_h(\mathbf{k}) + E_g. \quad (8.26)$$

Тогда из (8.25) и (8.26) получим

$$\alpha \approx B \sum_{\mathbf{k}} \delta(\varepsilon_c(\mathbf{k}) + \varepsilon_h(\mathbf{k}) + E_g - \hbar\omega), \quad (8.27)$$

где  $B = \frac{2\pi}{\hbar} \frac{|V_{12}|^2}{Nv}$ .

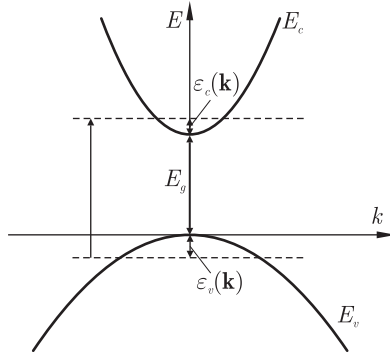


Рис. 8.8. Вертикальный переход электрона из валентной зоны в зону проводимости. Энергию перехода можно разбить на три части: ширину запрещенной зоны полупроводника  $E_g$ ; кинетическую энергию электрона в зоне проводимости  $\varepsilon_c(\mathbf{k})$ ; и кинетическую энергию дырки в валентной зоне  $\varepsilon_h(\mathbf{k})$  (см. (8.26))

Рассмотрим физический смысл полученного выражения (8.27) для коэффициента поглощения  $\alpha$ . Выше, при формулировке задачи о поглощении света, мы рассматривали переход электрона из валентной зоны в зону проводимости. Исходя из выражения (8.27) задачу о поглощении света можно сформулировать иначе: в полупроводнике ничего не было — был вакуум, и фотон с энергией  $\hbar\omega$  породил электрон-дырочную пару, при этом энергия фотона  $\hbar\omega$  расходуется на преодоление ширины запрещенной зоны полупроводника  $E_g$  и на сообщение электрону и дырке кинетических энергий  $\varepsilon_c(\mathbf{k})$  и  $\varepsilon_h(\mathbf{k})$  соответственно. Далее, для простоты рассматриваем сферические зоны для электронов и дырок (то есть изотропные параболические спектры). В результате коэффициент поглощения можно представить в виде

$$\alpha \approx B \sum_{\mathbf{k}} \delta\left(\frac{\hbar^2 \mathbf{k}^2}{2m_e} + \frac{\hbar^2 \mathbf{k}^2}{2m_h} + E_g - \hbar\omega\right) \approx B \sum_{\mathbf{k}} \delta[\varepsilon_\mu(\mathbf{k}) - (\hbar\omega - E_g)]. \quad (8.28)$$

Здесь  $\varepsilon_\mu(\mathbf{k}) = \frac{\hbar^2 \mathbf{k}^2}{2\mu}$  — энергия пары;  $\mu = \frac{m_e m_h}{m_e + m_h}$  — приведенная масса электрона и дырки,  $m_e$  и  $m_h$  — эффективная масса электрона и дырки соответственно. Напомним, что суммирование в (8.28) производится по всем состояниям  $\mathbf{k}$ . В главе 2 (колебания и волны, формула (9)) было показано правило преобразования суммы в интеграл

$$\sum_{\mathbf{k}} \varphi(\mathbf{k}) \rightarrow V \int \frac{\varphi(\mathbf{k})}{(2\pi)^3} d^3 k. \quad (8.29)$$

Если воспользоваться (8.29), мы можем в (8.28) перейти от суммирования по  $\mathbf{k}$  к интегрированию по  $\mathbf{k}$  или по энергии  $\varepsilon_\mu$ . Тогда коэффициент поглощения пропорционален

$$\alpha \propto \int \delta[\varepsilon_\mu(\mathbf{k}) - (\hbar\omega - E_g)] \rho_\mu(\varepsilon_\mu) d\varepsilon_\mu, \quad (8.30)$$

где  $\rho_\mu(\varepsilon_\mu) d\varepsilon_\mu$  — число состояний в интервале энергии  $d\varepsilon_\mu$ , то есть  $\rho_\mu$  — плотность состояний с приведенной массой  $\mu$ :

$$\rho_\mu(\varepsilon_\mu) = \frac{\mu^{3/2}}{\sqrt{2} \pi^2 \hbar^2} \varepsilon_\mu^{1/2}. \quad (8.31)$$

С учетом  $\delta$ -функции интеграл в (8.30) легко вычисляется, и мы получим

$$\alpha \propto \rho_\mu(\varepsilon_\mu)|_{\varepsilon_\mu = \hbar\omega - E_g}. \quad (8.32)$$

Итак, мы получили важный результат: коэффициент поглощения света пропорционален приведенной плотности состояний электронов и дырок при значении энергии  $\varepsilon_\mu = \hbar\omega - E_g$ . Учитывая зависимость плотности состояний от энергии (8.31), получаем окончательное выражение

для зависимости коэффициента поглощения от частоты света вблизи порога <sup>1)</sup>

$$\alpha \propto \sqrt{\hbar\omega - E_g}. \quad (8.33)$$

В результате зависимость коэффициента поглощения от частоты описывается корневым законом (рис. 8.9). Если  $\hbar\omega < E_g$ , то коэффициент поглощения  $\alpha \approx 0$ . Следует напомнить, что наш расчет справедлив вблизи  $E_g$ , то есть когда  $\hbar\omega - E_g \ll E_g$  (см. (8.24)). Какова же на самом деле зависимость коэффициента поглощения от частоты света? Оказывается, что экспериментальная зависимость коэффициента поглощения  $\alpha$  от частоты не такая, как на рис. 8.9. Край фундаментального поглощения очень чувствителен к чистоте материала (то есть к наличию примесей) и к температуре (из-за появления фононов). В примесном («грязном») полупроводнике зависимость коэффициента поглощения  $\alpha$  от частоты имеет вид такой, как на рис. 8.10. При этом говорят, что есть «хвост» коэффициента поглощения при частотах  $\hbar\omega < E_g$ , при этом, чем выше температура, тем больше хвост. Край фундаментального поглощения не резкий.

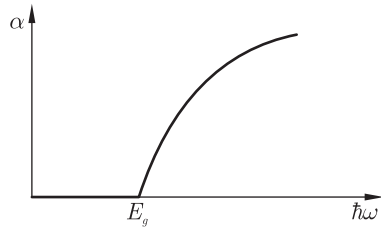


Рис. 8.9. Теоретическая зависимость коэффициента поглощения полупроводника от энергии фотона  $\hbar\omega$ . Корневой закон (8.33) следует из выражения для плотности состояний (8.31)

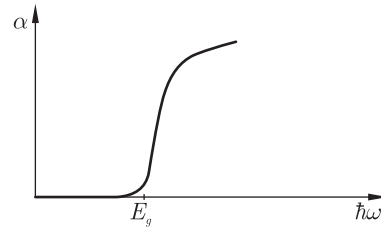


Рис. 8.10. Зависимость коэффициента поглощения примесного полупроводника от энергии кванта света  $\hbar\omega$  при конечной температуре. Наличие «хвоста» коэффициента поглощения при частотах  $\hbar\omega < E_g$  свидетельствует о наличии примесей и фононов в полупроводнике

Экспериментально даже в чистом материале наблюдается поглощение при  $E < E_g$  (см. рис. 8.10). Зависимость коэффициента поглощения  $\alpha$  от частоты такая, как на рис. 8.11, была предсказана Я. И. Френкелем в 1932 г. Немонотонная зависимость  $\alpha$  при  $\hbar\omega < E_g$  связана с экситонным поглощением. Позже, для полупроводников, экситонное погло-

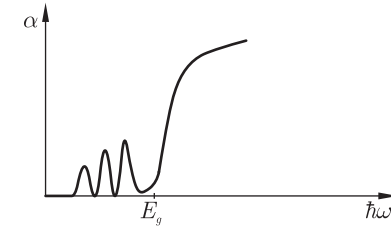


Рис. 8.11. Качественная зависимость коэффициента поглощения чистого полупроводника от частоты при низких температурах. Ряд пиков ниже края поглощения ( $E_g$ ) соответствует переходам электрона из валентной зоны на экситонные уровни, расположенные в запрещенной зоне ниже края зоны проводимости  $E_c$

щение света было детально исследовано Н. Моттом (1937). Теоретические предсказания Я. И. Френкеля и Н. Мотта были подтверждены экспериментально Е. Ф. Гроссом в 1951 г. Первые экспериментальные исследования экситонного поглощения света были выполнены на полупроводнике закиси меди (CuO). Экспериментально для коэффициента поглощения была получена целая серия пиков поглощения, расположенных ниже края фундаментального поглощения (рис. 8.11). В нашем теоретическом рассмотрении коэффициент поглощения ниже края фундаментального поглощения вообще равен нулю. Значит, наше рассмотрение не является полным. Чего же мы не учли при решении задачи о фундаментальном поглощении света полупроводником?

Вначале, до освещения полупроводника светом, валентная зона была полностью заполнена, а зона проводимости была пустой (вакуумное состояние). В конце, когда произошел переход (то есть поглощение света), появилась пара частиц: электрон в зоне проводимости и дырка в валентной зоне, поэтому выше мы и отметили, что энергия кванта света  $\hbar\omega$  расходуется на преодоление ширины запрещенной зоны  $E_g$  и идет на кинетические энергии электрона и дырки  $\varepsilon_c(\mathbf{k})$  и  $\varepsilon_h(\mathbf{k})$ . Но при этом мы не учли потенциальную энергию взаимодействия пары частиц, поскольку электрон и дырка — заряженные частицы (дырка имеет положительный заряд). Поэтому потенциальная энергия их взаимодействия отрицательная, и электрон-дырочную пару легче возбуждать. В этом случае закон сохранения энергии принимает вид

$$\hbar\omega = \varepsilon_c(\mathbf{k}) + \varepsilon_h(\mathbf{k}) + E_g + U. \quad (8.34)$$

Здесь  $U$  — потенциальная энергия кулоновского взаимодействия электрона и дырки, причем  $U$  — отрицательна ( $U < 0$ ). Поэтому для рождения электрон-дырочной пары нужно затратить меньше энергии, следовательно, становится возможным поглощение света с энергией кванта  $\hbar\omega$ , меньшей, чем ширина запрещенной зоны  $E_g$  ( $\hbar\omega < E_g$ ). Из этого следует, что при поглощении света рождается пара электрон-дыр-

<sup>1)</sup> Отметим, что этот результат верен только для разрешенных оптических переходов, то есть для тех, у которых  $\mathcal{P}_{12}$  в (8.16) не равен нулю. Это возможно, если блоховские функции валентной зоны и зоны проводимости имеют разную симметрию ( $P$  и  $S$  соответственно).

ка, не свободная, а связанная, а для рождения связанных пар нужно затратить меньше энергии.

Связанная электрон-дырочная пара называется экситоном (в данном случае экситоном Ванье–Мотта). Рассмотрим в каком состоянии после поглощения фотона могут оказаться электрон и дырка, то есть экситон. Гамильтониан электрона и дырки с учетом их кулоновского взаимодействия имеет вид

$$\hat{H} = \frac{\hat{\mathbf{p}}_e^2}{2m_e} + \frac{\hat{\mathbf{p}}_h^2}{2m_h} - \frac{e^2}{\kappa|\mathbf{r}_e - \mathbf{r}_h|}. \quad (8.35)$$

Здесь  $\kappa$  — статическая диэлектрическая постоянная полупроводника;  $\mathbf{r}_e$  и  $\mathbf{r}_h$  — радиус-векторы электрона и дырки соответственно. Гамильтониан (8.35) с точностью до переобозначений совпадает с гамильтонианом атома водорода <sup>1)</sup>.

В (8.35) можно перейти к системе центра тяжести системы и можно ввести приведенную массу электрона и дырки. Центр тяжести экситона, согласно закону сохранения импульса, должен быть равным волновому вектору света  $\mathbf{k}$ , следовательно, экситон рождается практически покоящимся, так как  $\hbar\mathbf{k} \ll \mathbf{p}_e, \mathbf{p}_h$ . Поэтому энергия относительно движения экситона — это энергия атома водорода

$$E_n = -\frac{\mu e^4}{2\hbar^2 \kappa^2} \frac{1}{n^2} = -\frac{R_{ex}}{n^2}, \quad (8.36)$$

где  $R_{ex}$  — экситонный ридберг;  $\mu$  — приведенная масса электрона и дырки,  $n = 1, 2, 3, \dots$ . Теперь мы можем выяснить, при каких значениях энергии кванта света  $\hbar\omega$  возможно поглощение. Энергия  $\hbar\omega$  тратится на преодоление ширины запрещенной зоны  $E_g$  за вычетом энергии связи пары (экситона), то есть

$$\hbar\omega = E_g - \frac{R_{ex}}{n^2}. \quad (8.37)$$

Получается, что ниже края фундаментального поглощения должна существовать серия водородоподобных уровней (рис. 8.11). Именно такая картинка и наблюдается на опыте.

Можно было бы ожидать, что такую зависимость коэффициента поглощения от частоты мы получили бы при переходе электрона на уровень донорной примеси в легированном полупроводнике. Коэффициент поглощения света донором должен быть пропорциональным концентрации доноров. И если донорных примесей в полупроводнике очень мало, то поглощения света донорными примесями не будет. Экситонное поглощение света в полупроводнике наиболее эффективно наблюдается в чистых полупроводниках.

<sup>1)</sup> См. Ландау Л. Д., Лифшиц Е. М. Квантовая механика. — М.: Наука, 1974. 752 с.

Рассмотрим поглощение света, если энергия кванта будет равна  $\hbar\omega > E_g$ . Это соответствует случаю положительной энергии экситона. Дело в том, что наряду со связанными состояниями экситона, можно рассматривать и несвязанные состояния, которым соответствуют положительные значения энергии. Такие состояния экситона соответствуют сплошному спектру, то есть электроны и дырки уже не связаны кулоновским притяжением и после их образования двигаются почти независимо. Кулоновское взаимодействие приводит к тому, что вероятность найти электрон и дырку рядом оказывается больше, чем для невзаимодействующей пары. Можно показать, что для несвязанных состояний вероятность поглощения пропорциональна волновой функции электрон-дырочной пары при  $\mathbf{r}_e = \mathbf{r}_h$  <sup>1)</sup>:

$$W \propto |\psi|_{\mathbf{r}_e=\mathbf{r}_h}^2. \quad (8.38)$$

Таким образом, в этом случае задача может быть решена аналитически и вероятность поглощения света  $W$  вычислена точно <sup>2)</sup>.

Тогда

$$\alpha = C\sqrt{\hbar\omega - E_g} \cdot \eta(\hbar\omega), \quad (8.39)$$

$$\eta(\hbar\omega) = \frac{\Gamma}{1 - \exp(-\Gamma)}, \quad \Gamma = 2\pi\sqrt{\frac{R_{ex}}{\hbar\omega - E_g}},$$

где  $C$  — некоторая константа;  $\eta$  называют фактором Зоммерфельда. Выше без учета кулоновского взаимодействия электрона и дырки мы получили (см. (8.33)), что

$$\alpha = C\sqrt{\hbar\omega - E_g}. \quad (8.40)$$

Если мы рассмотрим край поглощения, то есть  $\hbar\omega = E_g$ , то, согласно (8.40),  $\alpha = 0$ . В данном случае, то есть при учете кулоновского взаимодействия электрона и дырки, на краю поглощения коэффициент  $\alpha$  (согласно (8.39)) не равен нулю. Кроме того, в точке  $\hbar\omega = E_g$  коэффициент поглощения  $\alpha$  испытывает скачок (см. рис. 8.11). Таким образом, вблизи края поглощения кулоновское взаимодействие между электроном и дыркой существенно влияет на величину коэффициента поглощения света.

Далеко от края поглощения, когда энергия кванта  $\hbar\omega$  гораздо больше  $E_g$  ( $\hbar\omega \gg E_g$ ), кулоновское взаимодействие менее существенно. Тогда в (8.39) экспоненту можно разложить в ряд и мы получаем, что  $\alpha \approx C\sqrt{\hbar\omega - E_g}$ , то есть возвращаемся к формуле (8.40). Этот закон (корневая зависимость  $\alpha$  от частоты  $\omega$ ) наступает вдали от края поглощения, примерно при  $\hbar\omega - E_g \approx 50R_3$ .

<sup>1)</sup> Явный вид волновой функции атома водорода в нуле можно найти в книге Ландау Л. Д., Лифшица Е. М. Квантовая механика. — М.: Наука, 1974. 752 с.

<sup>2)</sup> См. Ансельм А. И. Введение в теорию полупроводников. — М.: Наука, 1978. 616 с.



Итак, экспериментально наблюдаемую зависимость коэффициента поглощения от частоты света можно объяснить, только если учитывать кулоновское взаимодействие электрона и дырки.

**Экситоны.** Первоначально экситон рассматривался как бестоковое возбуждение в кристалле, обладающее некоторым квазиимпульсом и энергией поступательного движения. Основы теории экситонов впервые были сформулированы Я. И. Френкелем (1931). Согласно идее, предложенной Френкелем, экситон рассматривался как сильно связанная система — как возбужденное состояние отдельного атома, которое может передаваться от одного атома к другому вследствие связи между соседними атомами (рис. 8.12). Пусть в одном из атомов (например, первом) мы перевели электрон на какой-то возбужденный уровень. Далее, возбужденный электрон переходит в основное состояние и освободившаяся энергия передается электрону соседнего атома; соседний, возбужденный, атом, снова переходит в основное состояние и передает энергию электрону следующего атома и т. д. В результате по кристаллу распространяется волна возбуждения, но электроны остаются на своих атомах. Такие возбуждения (экситоны) наблюдаются в щелочно-галогидных кристаллах (NaCl, NaBr, KCl и KBr), и в этом случае они называются экситонами малого радиуса.

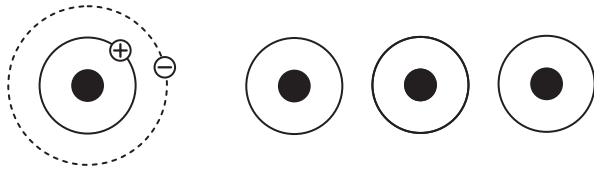


Рис. 8.12. Схематическая модель экситона Френкеля. Возбуждение, то есть связанная пара электрон–дырка, локализованная на каком-то атоме, перемещается от одного атома к другому вследствие связи между соседними атомами

Позже Ванье и Мотт рассмотрели аналог этого случая (экситона Френкеля) в кристаллах со слабой связью и возможность образования связанного состояния электрона и дырки. Пусть электрон находится в зоне проводимости кристалла, а дырка — в валентной зоне. Электрон и дырка притягиваются в силу кулоновского взаимодействия. Кулоновский потенциал в этом случае имеет вид  $U(r) = -e^2/\kappa r$ , где  $r$  — расстояние между частицами, а  $\kappa$  — диэлектрическая проницаемость кристалла. Связанное состояние электрона и дырки (экситона) (см. рис. 8.13) будет иметь полную энергию, меньшую, чем ширина запрещенной зоны. Такая система электрон–дырка (экситон) очень похожа на атом водорода, если энергетические спектры для электрона и дырки сферически симметричны. Энергетические уровни экситона вблизи дна зоны проводимости можно, таким образом, описывать несколько модифицированной формулой Бора для атома водорода

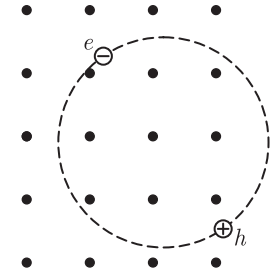


Рис. 8.13. Экситон Ванье–Мотта — связанная пара электрон–дырка, которая может свободно перемещаться в кристалле. Такой экситон рассматривается как слабо связанная система, причем расстояние между электроном и дыркой считается большим по сравнению с постоянной решетки кристалла

$$E_n = E_g - \frac{\mu e^4}{2\hbar^2 \kappa^2 n^2}. \quad (8.41)$$

Здесь  $n$  — главное квантовое число, а  $\mu$  — приведенная масса, определяемая соотношением

$$\frac{1}{\mu} = \frac{1}{m_e} + \frac{1}{m_h}, \quad (8.42)$$

где  $m_e$  — эффективная масса электрона, а  $m_h$  — эффективная масса дырки. Такие экситоны принято называть экситонами Ванье–Мотта, или экситонами большого радиуса (рис. 8.13).

Экситоны можно наблюдать экспериментально по линиям поглощения света. Например, для экситона Ванье–Мотта наблюдается серия водородоподобных линий вблизи края фундаментального поглощения полупроводника (рис. 8.14). Основное энергетическое состояние экситона мы получим, полагая в формуле (8.41)  $n = 1$ . Эта энергия (при  $n = 1$ ) соответствует потенциалу ионизации, необходимому для

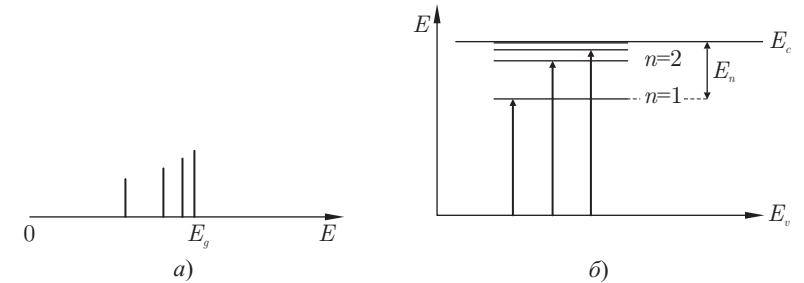


Рис. 8.14. Схема водородоподобных энергетических уровней экситона. Оптические переходы из валентной зоны показаны стрелками, самая длинная из которых соответствует потенциалу ионизации экситона; эта энергия равна ширине запрещенной зоны  $E_g$ . Энергия связи экситона по отношению к свободному электрону и свободной дырке равна  $E_n$



разделения экситона, который находится в основном состоянии, на составляющие его частицы — электроны и дырки.

**Замечание 1.** Вернемся к задаче о поглощении света в полупроводниках. Как было показано выше, без учета кулоновского взаимодействия между электроном и дыркой коэффициент поглощения имеет пороговую зависимость от частоты

$$\alpha \propto (\hbar\omega - E_g)^{1/2}. \quad (8.43)$$

Поэтому можно сказать, что при учете кулоновского взаимодействия корень квадратный в (8.43) умножается на множитель, который называется зоммерфельдовским множителем (см. (8.39)). Тонкая структура спектра поглощения света в полупроводниках, наблюдаемая ниже края поглощения, ниже  $E_g$ , связана с экситонным механизмом поглощения света (рис. 8.14).

**Замечание 2.** Интересно отметить, что когда было введено понятие экситона в твердом теле, то сразу же возникла аналогия с физикой элементарных частиц. При нулевой температуре валентная зона полупроводника полностью заполнена, а зона проводимости пуста. Это и есть не что иное, как *вакуумное состояние*. С ростом температуры (или за счет какого-нибудь другого процесса) в полупроводнике возникают электроны и дырки (рис. 8.15, а).

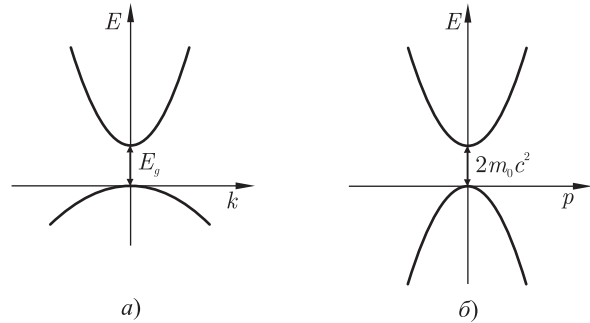


Рис. 8.15. Сравнение энергетических спектров: электрона и дырки в полупроводнике (а) и электрона и позитрона в релятивистской квантовой механике (б)

В физике элементарных частиц мы имеем электрон и позитрон, энергетические зоны которых симметричны (рис. 8.15, б). Поэтому электрон и позитрон имеют одинаковые массы. Запрещенная зона, разделяющая электронные и позитронные состояния, — это щель шириной  $2m_0c^2$ , где  $m_0$  — масса свободного электрона, а  $c$  — скорость света. В этом случае с помощью кванта света можно создать электрон-позитронную пару. На самом деле один фотон не может создать такую пару. Для выполнения закона сохранения импульса необходимо участие третьей частицы, например, еще одного фотона. Таким обра-

зом, два фотона уже могут возбуждать электрон-позитронную пару. Из физики элементарных частиц известно, что электрон и позитрон образуют связанное состояние, называемое позитронием<sup>1)</sup>.

Рассматривая приближение слабой связи для полупроводников, мы получили релятивистский спектр для электронов (см. (4.102)). В полупроводнике аналогом позитрония является экситон Ванье–Мотта. Таким образом, аналогия полупроводника с элементарными частицами является достаточно глубокой.

В заключение данного раздела мы еще раз обратимся к механизму поглощения света в полупроводниках. Мы уже выяснили, что облучение полупроводника светом порождает электрон-дырочные пары. Если энергия кванта света  $\hbar\omega$  чуть меньше, чем  $E_g$ , то рождаются экситоны. Если же  $\hbar\omega > E_g$ , то образуются свободные электроны и дырки. Ясно, что при низких температурах (когда  $k_B T$  меньше энергии ионизации экситона  $E_1$ ) свободные электроны и дырки связываются в экситоны. Впоследствии они могут аннигилировать (то есть рекомбинировать, излучая, например, квант света). Экситон в полупроводнике живет довольно долго (имеет большое время жизни, например, для GaAs время жизни экситона порядка  $10^{-9}$  с). В результате, при облучении полупроводника светом, можно создать и накопить очень много экситонов. Если экситонов много (а экситон — это аналог атома водорода), возникает аналог водородного газа. При понижении температуры и увеличении плотности газа экситонов могут образоваться молекулы — биэкситоны (аналог молекулы водорода). Известно, что водород при понижении температуры и увеличении плотности сжимается (превращается в жидкость). Таким образом, при понижении температуры и увеличении плотности возникает замечательное явление, когда экситонный газ сжимается и образуются капли — возникает так называемая электрон-дырочная жидкость (Л. В. Келдыш, 1969).

### 8.3. Поглощение света свободными носителями заряда

Если в полупроводнике имеются свободные носители заряда (электроны и дырки), то световая волна, падающая на кристалл, может индуцировать внутризонные переходы, при этом свет может поглощаться носителями заряда. Такое поглощение играет важную роль для полупроводников, когда энергия кванта света  $\hbar\omega$  меньше ширины запрещенной зоны  $E_g$ . В этом случае фундаментальное поглощение отсутствует, если пренебречь экситонным эффектом.

Выясним теперь, каким образом электромагнитная волна влияет на электроны (дырки). Могут ли электроны (или дырки) поглощать такое

<sup>1)</sup> Берестецкий В. Б., Лифшиц Е. М., Питаевский Л. П. Квантовая электродинамика. — М.: Наука, 1980. 704 с.

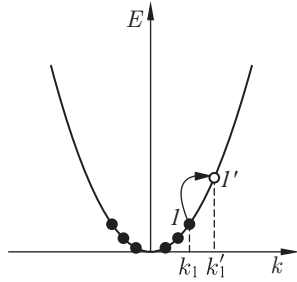


Рис. 8.16. Изменение квазиимпульса электрона вследствие изменения его энергии за счет поглощения кванта света  $\hbar\omega$ . Такие переходы происходят с испусканием или поглощением фонона

излучение? Оказывается, поглощать его непосредственно они не могут. Дело в том, что при поглощении кванта света электрон перешел бы из состояния «1» в состояние «1'» и тогда его квазиимпульс сильно изменился бы (рис. 8.16). Но мы уже выяснили, что переход электрона из одного состояния в другое с изменением квазиимпульса запрещен законом сохранения квазиимпульса. (В случае межзонных переходов электрон, поглощая квант света, переходит из валентной зоны в зону проводимости вертикально вверх с изменением энергии, но без изменения импульса.) Оказывается, внутризонное поглощение света электроном все же возможно при условии взаимодействия электрона (кроме фотона) еще с третьим телом (например с примесью). Ниже мы рассмотрим внутризонное поглощение света, используя классический подход.

Будем считать, что электрон движется в высокочастотном электромагнитном поле. При движении электрона в периодическом поле без рассеяния его средняя энергия не изменяется. Рассмотрим уравнение движения электрона в таком поле

$$m \frac{d\mathbf{v}}{dt} = e\mathbf{E}(t). \quad (8.44)$$

Здесь  $\mathbf{E}(t)$  — поле монохроматической электромагнитной волны,  $\mathbf{E}(t) = \mathbf{E}_0 \cos \omega t$ . Тогда уравнение (8.44) легко решается и мы получаем

$$\mathbf{v} = \frac{e\mathbf{E}_0}{m\omega} \sin \omega t. \quad (8.45)$$

Энергия, поглощаемая электронами в единице объема в единицу времени (то есть мощность), равна

$$W = e\mathbf{E}\mathbf{v}n \propto \sin \omega t \cos \omega t \propto \sin 2\omega t. \quad (8.46)$$

Отсюда видно, что при усреднении выражения (8.46) по периоду получается нуль. Таким образом, поглощение света свободными носителями заряда, как уже было отмечено выше, невозможно. Такое поглоще-

ние становится возможным, если электрон испытывает столкновения. Согласно решению (8.45) электрон при движении без столкновений совершает колебательное движение и отстает по фазе на  $\pi/2$  от поля  $\mathbf{E}$ .

С учетом столкновений фаза его скорости сбивается, и тогда уравнение движения для электрона имеет вид уравнения Лоренца–Друде (см. (6.6) гл. 6)

$$m \frac{d\mathbf{v}}{dt} = e\mathbf{E} - \frac{m\mathbf{v}}{\tau}. \quad (8.47)$$

Здесь  $\tau$  — характерное время релаксации электронов. Для решения уравнения (8.47) представим электрическое поле в комплексной форме

$$\mathbf{E} = \frac{1}{2} (\mathbf{E}_0 e^{-i\omega t} + \mathbf{E}_0^* e^{i\omega t}). \quad (8.48)$$

Тогда плотность тока  $j$ , выраженная через решение уравнения (8.47), будет

$$\mathbf{j} = en\mathbf{v} = \frac{1}{2}\sigma(\omega)\mathbf{E}_0 e^{-i\omega t} + \frac{1}{2}\sigma^*(\omega)\mathbf{E}_0^* e^{i\omega t}, \quad (8.49)$$

где

$$\sigma(\omega) = \frac{\sigma_0}{1 - i\omega\tau}. \quad (8.50)$$

Здесь  $\sigma_0 = \frac{ne^2}{m}\tau$  — значение проводимости в постоянном электрическом поле. Выражение (8.50) называется высокочастотной проводимостью.

Энергия, поглощаемая в единице объема полупроводника в единицу времени,

$$W = \mathbf{j}\mathbf{E} = \frac{1}{4}\sigma(\omega)E_0^2 e^{-2i\omega t} + \frac{1}{4}\sigma^*(\omega)E_0^{*2} e^{2i\omega t} + \frac{1}{2}[\sigma(\omega) + \sigma^*(\omega)]\frac{1}{2}|E_0|^2. \quad (8.51)$$

Усредним выражение (8.51) по периоду поля. Первые два члена при таком усреднении дают нуль, и, таким образом, средняя поглощаемая мощность поля в единице объема равна

$$\overline{W} = \sigma'(\omega)\overline{E^2}, \quad (8.52)$$

где  $\sigma'(\omega)$  — вещественная часть высокочастотной проводимости

$$\sigma'(\omega) = \frac{\sigma_0}{1 + \omega^2\tau^2}, \quad (8.53)$$

и  $\overline{E^2}$  — средний квадрат электрического поля волны.  $\overline{E^2} = \frac{1}{2}|E_0|^2$  согласно (8.48). Таким образом, рассеяние электронов приводит к внутризонному поглощению света.

### 8.4. Непрямые межзонные переходы

Рассмотрим межзонное поглощение света в непрямозонных полупроводниках, таких, как германий и кремний. Для этих полупроводников (Ge, Si) энергетические зоны имеют вид, представленный на рис. 4.16 и 4.17. Самый глубокий минимум в зоне проводимости расположен в точке, не совпадающей с положением максимума энергии в валентной зоне в точке  $\Gamma$ , при  $\mathbf{k} = 0$ . В германии минимумы  $L$  в зоне проводимости расположены на границе зоны Бриллюэна в направлении  $\langle 111 \rangle$ , а в кремнии — на оси  $\Delta$  в направлениях  $\langle 100 \rangle$ . В дальнейшем, для определенности, мы будем рассматривать не прямые переходы на примере германия.

Переход электрона из валентной зоны (вблизи точки  $\Gamma$ ) в зону проводимости (вблизи  $L$ ) за счет поглощения кванта света  $\hbar\omega > E_g$  невозможен; в этом случае закон сохранения волнового вектора (квазиимпульса) выполняться не может (рис. 8.17), так как электрону пришлось бы сильно изменить свой импульс. Поэтому для поглощения кванта света электроном и для перехода из точки  $\Gamma$  в точку  $L$  требуется взаимодействие с третьим телом для того, чтобы выполнялся закон сохранения квазиимпульса (например, это может быть рассеяние на примесях или на фононах). Однако известно, что процесс рассеяния электрона на примеси с существенным изменением величины квазиимпульса маловероятен, поэтому примеси не эффективны для такого процесса рассеяния.

Оказывается, в качестве третьего тела выступают фононы. Поэтому, если одновременно с поглощением фотона электрон поглотит (испустит) фонон с волновым вектором  $\mathbf{q} = \pm(\mathbf{k}_L - \mathbf{k}_\Gamma)$ , то закон сохранения импульса будет выполняться. Здесь  $\mathbf{k}_\Gamma$  и  $\mathbf{k}_L$  — волновые векторы электрона в начальном и конечном состояниях (напомним, что волновой вектор фотона  $\mathbf{k}$  существенно меньше, чем  $\mathbf{k}_L - \mathbf{k}_\Gamma$ , поэтому импульсом фотона можно пренебречь).

При низких температурах фононов практически не существует. Поэтому при этих условиях процессы рассеяния электрона с поглощением фонона отсутствуют. Однако процесс рассеяния электрона с испусканием фонона возможен, когда электрон поглощает квант света и переходит в зону проводимости. В этом случае энергия кванта света  $\hbar\omega$  должна превышать ширину запрещенной зоны  $E_g$  на величину энергии фонона  $\hbar\omega_q$ . Возникает вопрос о том, какого типа фонон излучается электроном. Напомним, что мы рассматриваем переход электрона из точки  $\Gamma$  в минимум зоны проводимости, который находится на грани-

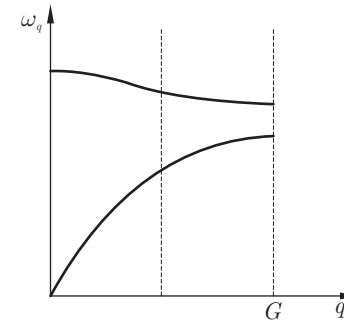


Рис. 8.18. Схематический вид акустической и оптической ветви фононов для некоторого направления в кристалле

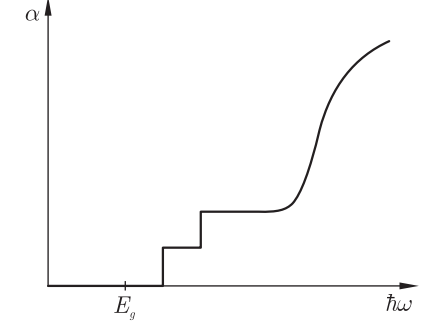


Рис. 8.19. Схематическая зависимость коэффициента поглощения от энергии фотона. Первая ступенька соответствует испусканию малоэнергичного фонона, вторая — испусканию второго фонона. Затем следует участок, соответствующий прямому переходу электрона из валентной зоны в зону проводимости (фундаментальное поглощение)

це зоны Бриллюэна. Это означает, что импульс электрона при этом изменяется на величину порядка  $G$ , где  $G$  — вектор обратной решетки. Поэтому фононы, которые могут возбуждаться, имеют импульс  $q$  порядка  $G$  (см. рис. 8.18). Тогда для коэффициента поглощения можно наблюдать зависимость от энергии, представленную на рис. 8.19. Первая ступенька на рис. 8.19 соответствует испусканию самого малоэнергичного фонона, вторая ступенька — испусканию второго фонона. Затем начинается фундаментальное поглощение, то есть прямой переход электрона из валентной зоны в зону проводимости.

Рассмотрим более детально квантовую теорию не прямых оптических переходов. Для этого можно воспользоваться «золотым» правилом квантовой механики для вероятности перехода

$$W_{12} = \frac{2\pi}{\hbar} |V_{12}|^2 \delta(E_1 - E_2). \quad (8.54)$$

Здесь  $V_{12}$  — матричный элемент перехода из начального ( $E_1$ ) в конечное ( $E_2$ ) состояние. Под знаком  $\delta$ -функции стоит разность энергий начального и конечного состояний. В начальном состоянии у нас есть электрон в валентной зоне и фотон, который он поглощает; в конечном состоянии мы имеем электрон в зоне проводимости и фонон, который он излучил. При вычислении вероятности не прямых переходов можно воспользоваться теорией возмущений. В теории возмущений такие оптические переходы из начального состояния  $|v, \mathbf{k}_\Gamma\rangle$  в конечное состояние  $|c, \mathbf{k}_L\rangle$  выступают как процессы второго порядка, осуществляемые

через промежуточные состояния  $|\mathbf{k}_n\rangle$ . При этом матричный элемент перехода равен

$$V_{12} = \sum_n \frac{V_{2n} V_{n1}}{E_n - E_1} + \sum_{n'} \frac{V_{2n'} V_{n'1}}{E_2 - E_{n'}}. \quad (8.55)$$

Здесь  $n, n'$  — номера промежуточных состояний.

Из (8.55) следует, что квантовый переход электрона из валентной зоны в зону проводимости происходит в два этапа. Первый этап — это переход электрона из начального состояния валентной зоны «1» в промежуточное состояние « $n$ » зоны проводимости. Такой переход описывается матричным элементом  $V_{n1}$ . Второй этап — это переход электрона из промежуточного состояния « $n$ » в конечное состояние «2» зоны проводимости. Этот переход описывается матричным элементом  $V_{2n}$ . На рис. 8.20 представлены энергетические зоны Ge и два канала перехода электрона из начального состояния «1» в валентной зоне (вблизи  $\mathbf{k} \approx 0$ ) в конечное состояние «2» в зону проводимости (вблизи точки  $L$ ). В первом канале на пути  $1n2$  электрон переходит из начального состояния «1» в промежуточное состояние « $n$ », поглощая фотон (прямой межзонный переход), а затем переходит из промежуточного состояния « $n$ » в конечное состояние «2», испуская фотон с волновым вектором  $\mathbf{q} = -(\mathbf{k}_L - \mathbf{k}_\Gamma) \approx -\mathbf{k}_L$ . Первый этап  $1n$  соответствует виртуальному переходу электрона из начального состояния в промежуточное, при этом энергия для виртуального перехода не сохраняется. Однако импульс при виртуальных переходах сохраняется. *Итак, теория квантовых переходов утверждает, что закон сохранения энергии должен выполняться только для начального состояния 1 и конечного состояния 2<sup>1)</sup>. Закон сохранения квазиимпульса имеет место и для промежуточных (виртуальных) переходов.*

Рассмотрим второй канал перехода электрона из валентной зоны в зону проводимости. На пути  $1n'2$  электрон из валентной зоны с волновым вектором  $\mathbf{k} \approx \mathbf{k}_L$  поглощает фотон и переходит в зону проводимости в окрестность точки  $L$ , а затем другой электрон в валентной зоне вблизи точки  $\Gamma$  с волновым вектором  $\mathbf{k} \approx 0$  испускает фотон с волновым вектором  $\mathbf{q} = -(\mathbf{k}_L - \mathbf{k}_\Gamma)$ , переходя на место дырки в точке  $\mathbf{k} \approx \mathbf{k}_L$ . Второй канал рекомбинации можно представить себе происходящим с дыркой. Дырка в зоне проводимости поглощает фотон и переходит из состояния «2» в « $n'$ » (это эквивалентно переходу электрона из  $n'$  в 2), а затем дырка поглощает фотон и переходит из состояния « $n'$ » в состояние «1». Отметим, что оба перехода дырки эквивалентны переходу электрона из состояния «1» в «2».

Итак, переход электрона из состояния «1» в промежуточное состояние « $n$ » обусловлен электрон-фотонным взаимодействием. Матричный

<sup>1)</sup> Фактически, в течение короткого промежутка времени, когда идет переход, соотношение неопределенностей допускает «несохранение» энергии.

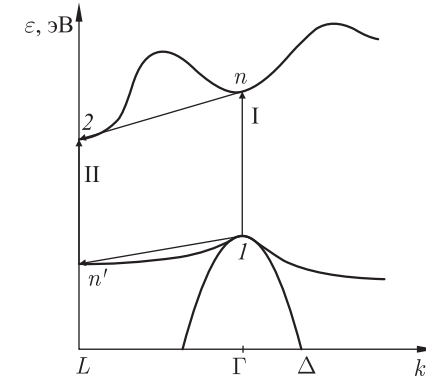


Рис. 8.20. Непрямые переходы в зонной структуре германия. I — это переход электрона, связанный с поглощением фотона, из валентной зоны (1) в зону проводимости ( $n$ ) и следующий за ним переход внутри зоны проводимости (из  $n$  в 2) при поглощении или испускании фонона. II — это переход дырки, связанный с поглощением фотона, из зоны проводимости (2) в валентную зону ( $n'$ ) и следующий за ним переход дырки внутри валентной зоны (из  $n'$  в 1) при поглощении или испускании фонона

элемент для такого перехода обозначим через  $M_{n1}^{\text{фот}}$ . Переход электрона из промежуточного состояния « $n$ » в конечное состояние «2» обусловлен электрон-фононным взаимодействием. Матричный элемент такого перехода обозначим через  $M_{2n}^{\text{фон}}$ . В результате полное число таких переходов в единицу времени в единице объема через промежуточные состояния « $n$ » и « $n'$ » равно

$$W = \frac{2\pi}{\hbar} \sum_{1,2} \left\{ \left| \sum_n \frac{M_{2n}^{\text{фон}} \cdot M_{n1}^{\text{фот}}}{\varepsilon_n - \varepsilon_1 - \hbar\omega} + \sum_{n'} \frac{M_{n'1}^{\text{фон}} \cdot M_{2n'}^{\text{фот}}}{\varepsilon_2 - \varepsilon_{n'} - \hbar\omega} \right|^2 \right\} \times \delta(\varepsilon_2 - \varepsilon_1 - \hbar\omega + \hbar\omega_q). \quad (8.56)$$

Здесь суммирование проводится по начальным (1), конечным (2) и промежуточным ( $n$ ) и ( $n'$ ) состояниям электрона, а  $\varepsilon_2$  — это энергия электрона вблизи точки  $L$ .

Исходя из определения коэффициента поглощения света (8.23) получим

$$\alpha_{\text{непр}} = \frac{1}{Nv} W = \frac{2\pi}{\hbar} \frac{1}{Nv} \times \sum_{1,2} \left\{ \left| \sum_n \frac{M_{2n}^{\text{фон}} \cdot M_{n1}^{\text{фот}}}{\varepsilon_n - \varepsilon_1 - \hbar\omega} \right|^2 + \left| \sum_{n'} \frac{M_{n'1}^{\text{фон}} \cdot M_{2n'}^{\text{фот}}}{\varepsilon_2 - \varepsilon_{n'} - \hbar\omega} \right|^2 \right\} \times \delta(\varepsilon_2 - \varepsilon_1 - \hbar\omega + \hbar\omega_q). \quad (8.57)$$

Напомним, что знаки « $\pm$ » перед  $\hbar\omega_q$  соответствуют поглощению и испусканию фонона соответственно<sup>1)</sup>.

Отметим, что с ростом температуры эффективность процессов поглощения фононов при не прямых оптических переходах увеличивается, поэтому появится «хвост» коэффициента поглощения при энергии  $E < E_g$ . Фактически край поглощения сдвигается в длинноволновую область.

### 8.5. Индуцированное (вынужденное) излучение. Формулы Планка и Эйнштейна

Кристаллические твердые тела могут быть использованы для создания квантовых генераторов СВЧ диапазона (мазеры) и генераторов в оптической области спектра (лазеры). Принцип работы и мазера и лазера основан на явлении вынужденного (индуцированного) излучения света. Как возникает вынужденное излучение фотонов в среде? Для ответа на данный вопрос рассмотрим вывод формулы Планка для излучения черного тела согласно Эйнштейну.

Кратко напомним историю решения проблемы излучения абсолютно черного тела. Среди возможных видов тепловых излучений важное место занимает такое излучение, которое находится в термодинамическом равновесии с излучающими телами. Такое излучение принято называть излучением абсолютно черного тела. Энергия излучения черного тела в полости объемом  $V$  в интервале частот  $\omega$ ,  $\omega + d\omega$  дается формулой Планка

$$dE_\omega = \frac{V\hbar}{\pi^2 c^3} \frac{\omega^3 d\omega}{e^{\hbar\omega/k_B T} - 1}. \quad (8.58)$$

При малых частотах ( $\hbar\omega \ll k_B T$ ) формула Планка (8.58) переходит в формулу Рэлея–Джинса для энергии черного излучения. В обратном предельном случае больших частот ( $\hbar\omega \gg k_B T$ ) из формулы (8.58) следует закон смещения Вина. Из (8.58) следует, что спектральная плотность черного излучения, отнесенная к единице объема, равна

$$\frac{1}{V} \frac{dE_\omega}{d\omega} = \frac{\hbar}{\pi^2 c^3} \omega^3 N_\omega. \quad (8.59)$$

Здесь

$$N_\omega = \frac{1}{e^{\hbar\omega/k_B T} - 1} \quad (8.60)$$

— это универсальная функция. Для вывода формулы (8.58) Планк вынужден был предположить, что свет излучается порциями; таким образом Планк ввел понятие кванта света — фотона. Поскольку черное излучение можно рассматривать как газ, состоящий из фотонов, то

<sup>1)</sup> Все детали вычисления  $\alpha^{\text{непр}}$  можно найти в книге: *Маделунг О.* Теория твердого тела. — М.: Наука, 1980. 416 с.

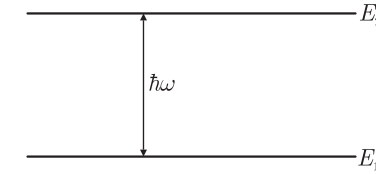


Рис. 8.21. Схематическое представление двухуровневой системы

величина  $N_\omega$  имеет смысл среднего числа фотонов с частотой  $\omega$  при данной температуре.

Эйнштейн получил выражение для среднего числа фотонов  $N_\omega$  исходя из других соображений. Рассмотрим двухуровневую систему (например, атом). Расстояние между уровнями равно  $\hbar\omega$  (рис. 8.21). Пусть таких одинаковых двухуровневых систем, то есть атомов, будет много. Заселенности уровней  $E_1$  и  $E_2$  обозначим через  $n_1$  и  $n_2$  соответственно, число фотонов обозначим через  $N$ . В стационарных условиях число атомов на каждом энергетическом уровне остается постоянным и равным равновесному значению ввиду баланса между процессами поглощения и излучения кванта света. Рассмотрим число переходов атома в единицу времени из состояния  $E_1$  в состояние  $E_2$  с поглощением кванта света. Это число переходов пропорционально числу квантов света  $N$ , числу атомов в состоянии «1», то есть  $n_1$ , и дается выражением

$$n_{12} = B_\pi N n_1. \quad (8.61)$$

Фотон может стимулировать переход не только из состояния «1» в состояние «2», но и из состояния «2» в состояние «1». Это так называемые индуцированные переходы. В этом случае имеющийся фотон индуцирует излучение еще одного фотона, следовательно, атом переходит из состояния «2» в состояние «1» под действием уже имеющегося фотона. Число переходов в единицу времени для такого процесса равно

$$n_{21} = B_\pi N n_2. \quad (8.62)$$

Здесь  $B_\pi$  и  $B_\pi$  — постоянные. Кроме того, имеют место еще спонтанные переходы из состояния «2» в состояние «1» с испусканием фотона. Число таких переходов в единицу времени равно

$$n_{\text{сп}} = A n_2, \quad (8.63)$$

где  $A$  — некоторая постоянная.

В состоянии термодинамического равновесия число переходов из состояния «1» в состояние «2» равно числу переходов из состояния «2» в состояние «1»:

$$B_\pi N n_1 = B_\pi N n_2 + A n_2. \quad (8.64)$$



Решая это уравнение относительно числа фотонов, находим

$$N = \frac{1}{\frac{B_n n_1}{A n_2} - \frac{B_n}{A}}. \quad (8.65)$$

Поскольку система находится в состоянии термодинамического равновесия, то заселенности уровней подчиняются больцмановскому распределению, так что

$$\frac{n_1}{n_2} = \exp\left(\frac{E_2 - E_1}{k_B T}\right) = \exp\left(\frac{\hbar\omega}{k_B T}\right). \quad (8.66)$$

Здесь мы учли, что  $E_2 - E_1 = \hbar\omega$  (см. рис. 8.21). В результате, для  $N$  получаем

$$N = \frac{1}{\frac{B_n}{A} \exp\left(\frac{\hbar\omega}{k_B T}\right) - \frac{B_n}{A}}. \quad (8.67)$$

Сравнивая это выражение с формулой Планка (8.60), мы приходим к выводу, что

$$B_n = B_i = A. \quad (8.68)$$

Это означает, что темп переходов с поглощением и с индуцированным излучением одинаков. Отсюда следует важный вывод: можно добиться такого условия, чтобы свет в среде не поглощался, а усиливался. Для этого нужно, чтобы заселенность  $n_2$  была больше  $n_1$ ,

$$n_2 > n_1. \quad (8.69)$$

Это условие означает избыток атомов на верхнем уровне  $n_2$  по сравнению с нижним  $n_1$ . Из сказанного ясно, что  $n_2$  — это число атомов, в принципе, способных испустить фотон, а  $n_1$  — это число атомов, способных поглотить фотон. При тепловом равновесии  $n_2 < n_1$  и какое-либо результирующее излучение в системе отсутствует. Однако, когда созданы такие неравновесные условия, что  $n_2 > n_1$ , имеет место испускание излучения.

Основная проблема, возникающая при создании мазеров и лазеров, состоит в отыскании способов создания инверсной заселенности, то есть избыточной населенности верхнего уровня по сравнению с нижним. В двухуровневой системе нельзя получить инверсную заселенность при любой интенсивности возбуждения (максимум, что можно получить в этом случае, — это равенство заселенностей,  $n_2 = n_1$ ).

Поэтому одним из возможных способов создания инверсной заселенности является использование системы, состоящей из трех уровней (рис. 8.22). В такой системе с тремя уровнями энергии можно получить инверсию населенностей, насыщая мощным источником излучения переходы между уровнями  $E_1$  и  $E_3$  (то

есть создавая ситуацию с  $n_3 = n_1$ ). Далее, электроны за счет спонтанного излучения могут переходить на уровень  $E_2$ . Таким путем мож-

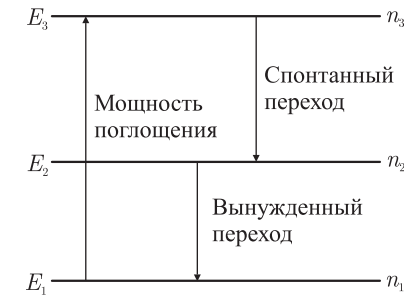


Рис. 8.22. Энергетическая диаграмма трехуровневого лазера. За счет поглощения кванта света электрон переходит с основного уровня  $E_1$  на уровень  $E_3$ . Из возбужденного состояния  $E_3$  электрон быстро переходит на уровень  $E_2$ . Между уровнями  $E_2$  и  $E_1$  создается инверсная заселенность. Переход  $E_2 \rightarrow E_1$  — вынужденный переход

но создать инверсию населенностей между уровнями  $E_2$  и  $E_1$  (в этом случае можно переводить все электроны с уровня  $E_1$  на уровень  $E_2$ ). Такая ситуация возникает, когда  $\tau_{32} \ll \tau_{21}$ ; где  $\tau_{32}$  — время спонтанных переходов с уровня  $E_3$  на уровень  $E_2$ ,  $\tau_{21}$  — время перехода с уровня  $E_2$  на уровень  $E_1$ . Таким образом, лазерное излучение возникает на переходе  $E_2 \rightarrow E_1$ . Именно на таком принципе работает рубиновый лазер.

**8.5.1. Усиление света в полупроводниках. Лазер на гомопереходе.** Идея применения полупроводников для генерации электромагнитного излучения была сформулирована в 1958–1959 гг. Н. Г. Басовым и А. М. Прохоровым с сотрудниками. В 1961 г. они же предложили использовать для получения лазерного эффекта инжекцию носителей заряда в вырожденных  $p$ – $n$ -переходах. Отличительной особенностью полупроводниковых лазеров является использование в них оптических переходов с участием состояний зоны проводимости и валентной зоны. Речь идет об усилении света в области края фундаментального поглощения полупроводника. Как уже было показано выше, свет с энергией кванта  $\hbar\omega > E_g$ , где  $E_g$  — ширина запрещенной зоны полупроводника, будет поглощаться. Спрашивается, что нужно выполнить в полупроводнике, чтобы свет с  $\hbar\omega > E_g$  не поглощался, а усиливался. В полупроводнике в состоянии термодинамического равновесия свет всегда будет поглощаться, так как заселенность зоны проводимости  $f_c$  всегда меньше заселенности валентной зоны  $f_v$ . Отношение заселенностей в термодинамическом равновесии подчиняется статистике Больцмана

$$\frac{f_c}{f_v} \propto \exp\left(-\frac{\hbar\omega_{cv}}{k_B T}\right). \quad (8.70)$$

Здесь  $\hbar\omega_{cv} = E_c - E_v = E_g$ . Итак, для получения индуцированного излучения необходимо, чтобы заселенность уровней была сильно неравновесной. Возникает вопрос, сколько должно быть носителей при



этом в зоне проводимости и валентной зоне, чтобы заселенность была инверсной. Пусть в полупроводнике мы создали (например за счет инжекции) неравновесные электроны и дырки. Далее, произойдет их энергетическая релаксация: электроны с высоких уровней переходят на дно зоны проводимости, а дырки всплывают к потолку валентной зоны. Предположим при этом, что рекомбинация электронов и дырок происходит медленно. Тогда электроны и дырки примут отдельно равновесные — фермиевские распределения по энергии. Чем отличается равновесное распределение по энергии электронов и дырок отдельно от термодинамического равновесия? Это различие по количеству носителей. В термодинамическом равновесии электроны и дырки рекомбинируют и их количество мало. То, что электроны и дырки отдельно равновесны по энергии, означает, что их квазиуровни Ферми ( $E_{F_e}$  и  $E_{F_h}$ ) разные. В термодинамическом равновесии их квазиуровни Ферми одинаковы. Таким образом, в каждой зоне создано равновесное по энергии распределение носителей, однако зоны не находятся в состоянии термодинамического равновесия друг с другом<sup>1)</sup>.

Степень заполнения некоторого уровня электрона энергией  $E_2$  в зоне проводимости определяется функцией Ферми–Дирака и равна

$$f_2 = \frac{1}{\exp \left[ \frac{(E_2 - E_{F_e})}{k_B T} \right] + 1}. \quad (8.71)$$

Аналогично, степень заполнения какого-то уровня энергии  $E_1$  электрона в валентной зоне равна

$$f_1 = \frac{1}{\exp \left[ \frac{(E_1 - E_{F_h})}{k_B T} \right] + 1}. \quad (8.72)$$

Условие инверсии для уровней с энергией  $E_2$  и  $E_1$  состоит в выполнении неравенства

$$f_2(E_2) - f_1(E_1) > 0. \quad (8.73)$$

Легко видеть, что условие (8.73) выполняется, если

$$E_{F_e} - E_{F_h} > E_2 - E_1. \quad (8.74)$$

Получается, что разность квазиуровней Ферми должна быть больше разности энергий уровней  $E_2 - E_1$ , но эта разность равна энергии излучаемого кванта  $E_2 - E_1 = \hbar\omega$ . Минимально возможная энергия излучаемого кванта — это ширина запрещенной зоны  $\hbar\omega = E_g$ . В ре-

<sup>1)</sup> Это связано с тем, что энергетическая релаксация носителей внутри зоны идет значительно быстрее, чем межзонаная рекомбинация ( $\tau_e \ll \tau_i$ ),  $\tau_i$  — время жизни.

зультате условие инверсной заселенности (8.74) для какой бы то ни было энергии излучения может быть записано в виде

$$E_{F_e} - E_{F_h} > E_g. \quad (8.75)$$

Важно подчеркнуть, что условие (8.75) соответствует случаю, когда квазиуровень Ферми (либо электронов, либо дырок) лежит в разрешенной зоне, а это означает, что эти носители вырождены. Однако создать в полупроводнике вырожденные электроны или дырки не так-то просто. Такая задача может быть решена за счет инжекции носителей заряда через  $p - n$ -переход.

Напомним, что материал  $n$ -типа — это полупроводник, легированный донорной примесью, а материал  $p$ -типа — это полупроводник, легированный акцепторной примесью. Если между материалами  $n$ -типа и  $p$ -типа образуется переход, то величины  $E_{F_e}$  и  $E_{F_h}$  одинаковы для всей среды, заселенность состояний вблизи  $p - n$ -перехода показана на рис. 8.23 (это равновесная ситуация).

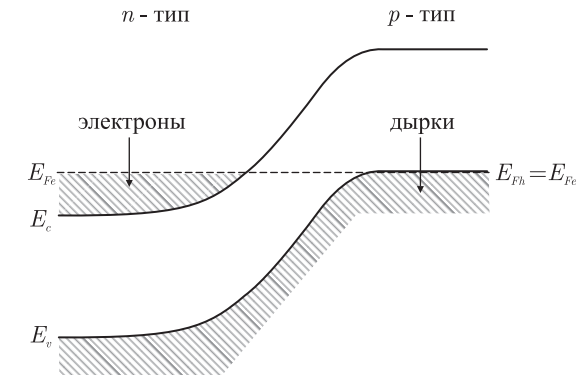


Рис. 8.23. Энергетическая диаграмма для полупроводника с  $p - n$ -переходом в случае теплового равновесия

Если к  $p - n$ -переходу приложено напряжение  $U > 0$  (прямое смещение — «+» на  $p$ -область и «-» на  $n$ -область), то при прохождении электрона через переход из области  $n$  в область  $p$  его энергия уменьшается на величину  $eU$ . Поэтому при наличии прямого смещения нужно изменить рис. 8.23 так, чтобы энергия в  $p$ -области уменьшалась по отношению к энергии в  $n$ -области на величину  $eU$ . На рис. 8.24 представлена энергетическая диаграмма, соответствующая прямому смещению. Приложение напряжения прямого смещения приводит к инжекции неосновных носителей через  $p - n$ -переход, при этом электроны инжектируются в  $p$ -область, а дырки — в  $n$ -область. В результате вблизи  $p - n$ -перехода (больше в  $p$ -области) существует область толщиной порядка 1–10 мкм, в которой зона проводимости заполнена электронами, а валентная зона занята дырками. Это область

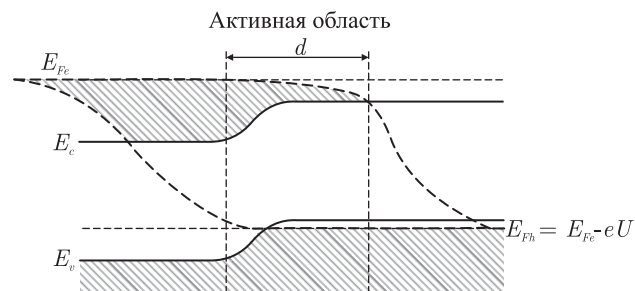


Рис. 8.24. Энергетическая диаграмма инжекционного полупроводникового лазера при смещении в прямом направлении. Заметим, что в материале  $p$ -типа существует инверсная заселенность в непосредственной близости к  $p-n$ -переходу ( $E_{Fe} - E_{Fn} > E_g$ )

порядка диффузионных длин электронов и дырок. В результате в этой области возникает инверсная заселенность, так как расстояние между квазиуровнями Ферми электронов и дырок больше  $E_g$ .

Таким образом, если  $p-n$ -переход смещен в прямом направлении, зонная структура принимает вид, такой, как на рис. 8.24. Очевидно, что толщина активной области  $d$  (то есть области, в которой имеется инверсная заселенность) приблизительно равна среднему расстоянию, проходимому электроном до рекомбинации с дыркой. Согласно теории диффузии (см. главу 7) толщина  $d$  дается выражением  $d = (D\tau)^{1/2}$ , где  $D$  — коэффициент диффузии, а  $\tau$  — среднее время жизни неосновного носителя до рекомбинации. В GaAs мы имеем  $D \approx 10 \text{ см}^2/\text{с}$ , а  $\tau \approx 10^{-9} \text{ с}$ , так что  $d \approx 1 \text{ мкм}$ . Наконец, заметим, что, поскольку  $E_{Fe} - E_{Fn} = eU$ , то из соотношения (8.75) следует, что  $U \approx E_g/e$ . В случае GaAs мы имеем  $U \approx 1,5 \text{ В}$ .

Первым материалом, который был использован для создания лазера на  $p-n$ -переходе, был арсенид галлия GaAs. Он излучает в близкой инфракрасной области спектра на длине волны, приблизительно равной  $8383 \text{ Å}$  ( $1,48 \text{ эВ}$ ).

На рис. 8.25 приведена типичная конструкция лазера на  $p-n$ -переходе на основе GaAs. Заштрихованная область представляет собой активный слой лазера. Чтобы обеспечить необходимую для генерации обратную связь электромагнитной волны, две выходные плоскости лазера делают параллельными друг другу, обычно посредством скалывания вдоль кристаллографических плоскостей. Эти две плоскости обычно называются зеркалами лазера. Часто на эти зеркала не наносятся отражающие покрытия. Действительно, так как показатель преломления у полупроводника очень большой (например, для GaAs  $n = 3,6$ ), то на поверхности раздела полупроводник–воздух при френелевском отражении получается достаточно высокий коэффициент отражения, порядка 35% (для GaAs). Как уже было отмечено выше, толщина активной области в перпендикулярном

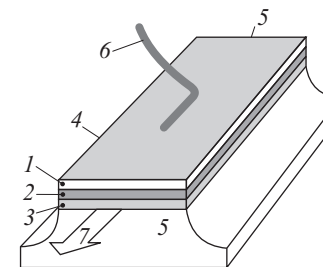


Рис. 8.25. Типичная конструкция лазера на  $p-n$ -переходе. Заштрихованная область (2) представляет собой активный слой лазера. Характерная ширина лазера — 100–200 мкм, а длина — 200–500 мкм. 1 — область  $p$ -типа; 2 — активный слой; 3 — область  $n$ -типа; 4 — слой металла; 5 — сколотая грань; 6 — электрический контакт; 7 — лазерный пучок

к  $p-n$ -переходу направлению составляет около 1 мкм. Однако вследствие дифракции поперечный размер лазерного пучка в этом направлении больше толщины активной области и составляет величину порядка 5 мкм. Следовательно, лазерный пучок довольно далеко проникает в  $p$  и  $n$ -области, где он испытывает сильное поглощение. Это является главной причиной того, что пороговая плотность тока при комнатной температуре в лазере на гомопереходе оказывается высокой, порядка  $10^5 \text{ А/см}^2$  для GaAs. Вследствие этого такой лазер не может работать в непрерывном режиме при комнатной температуре. При комнатной температуре возможна только импульсная генерация. Однако пороговая плотность тока в лазере на гомопереходе быстро (экспоненциально) уменьшается с понижением рабочей температуры. Вследствие этого лазеры на гомопереходе могут работать в непрерывном режиме только при низких температурах. Это является серьезным недостатком данного типа лазеров, что наложило ограничения на возможности их практического применения.

**8.5.2. Лазер на гетеропереходе.** Впервые непрерывная генерация при комнатной температуре была получена в ФТИ им. А. Ф. Иоффе на полупроводниковом лазере с двойным гетеропереходом <sup>1)</sup>. Одиночный гетеропереход представляет собой контакт между двумя различными полупроводниками (с разной шириной запрещенной зоны, например GaAs и  $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ) <sup>2)</sup> (рис. 8.26).

Мы рассмотрим кратко двойной гетеропереход, так как именно он применяется для создания полупроводниковых лазеров, работающих

<sup>1)</sup> Алферов Ж. И., Андреев В. М., Портной Е. Л., Трукан М. К. Инжекционные лазеры на основе гетеропереходов в системе AlAs–GaAs с низким порогом генерации при комнатной температуре // ФТП. 1969. Т. 3. С. 1328–1332.

<sup>2)</sup> Примечание:  $\text{Al}_x\text{Ga}_{1-x}\text{As}$  обозначает твердый раствор, в котором  $x$  — доля атомов Al, замещающих атом Ge.

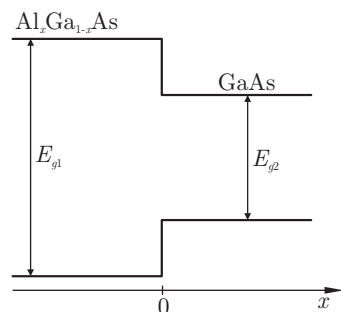


Рис. 8.26. Зонная диаграмма одиночного гетероперехода состоящего из двух различных материалов AlGaAs и GaAs

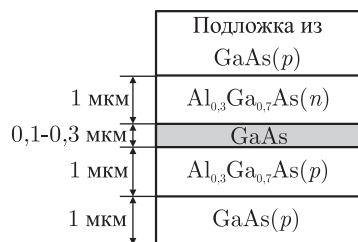


Рис. 8.27. Схематическое представление полупроводникового лазера с двойным гетеропереходом. Активная область (заштрихованная область) представляет собой слой из GaAs  $n$ -типа

в непрерывном режиме при комнатной температуре. Чтобы проиллюстрировать его свойства, на рис.8.27 приведен пример лазерной структуры с двойным гетеропереходом. В этом лазере реализованы два гетероперехода:  $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$  ( $p$ -типа)–GaAs и GaAs– $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$  ( $n$ -типа). Активная область представляет собой тонкий слой GaAs (толщиной порядка 0,1–0,3 мкм; на рис.8.27 — это заштрихованная область). В такой лазерной структуре пороговую плотность тока при комнатной температуре можно уменьшить примерно на два порядка (то есть до  $\sim 10^3 \text{ А/см}^2$ ) по сравнению с лазером на гомопереходе. Таким образом, лазер на двойном гетеропереходе может работать в непрерывном режиме при комнатной температуре. Резкое уменьшение пороговой плотности тока происходит благодаря совместному действию трех следующих факторов.

1) Показатель преломления GaAs больше показателя преломления  $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$  ( $n_1 \approx 3,6$ ), что приводит к образованию оптической волноводной структуры (рис. 8.28, а). Отсюда следует, что активная область (GaAs) является оптически более плотной средой и что лазерный пучок будет теперь сосредоточен главным образом в слое GaAs, то есть в области, в которой имеется усиление (рис. 8.28, б).

2) Ширина запрещенной зоны  $E_{g1}$  в GaAs (около 1,5 эВ) меньше, чем ширина запрещенной зоны  $E_{g2}$   $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$  (около 1,8 эВ). Поэтому на обоих гетеропереходах образуются энергетические барьеры, которые эффективно удерживают инжектированные электроны и дырки в активном слое (рис. 8.28, в). Таким образом, для данной плотности тока концентрация электронов и дырок в активной области возрастает, а значит, увеличивается и усиление.

3) Поскольку  $E_{g2}$  значительно больше, чем  $E_{g1}$ , то лазерный пучок с частотой  $\omega = E_{g1}/\hbar$  почти не поглощается в  $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ . Поэтому крылья поперечного профиля волны, заходящие как в  $p$ , так и в  $n$ -области (рис. 8.28, б), не испытывают там сильного поглощения.

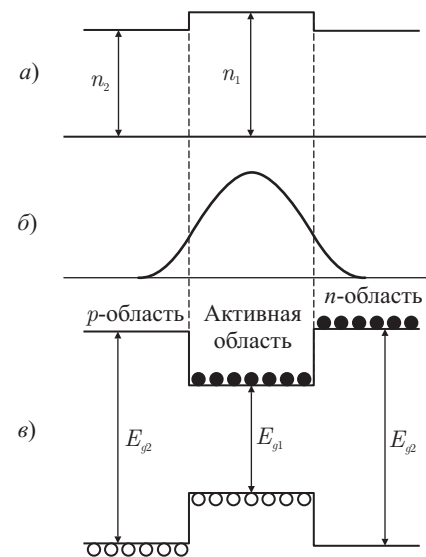


Рис. 8.28. Фрагмент лазера на двойном гетеропереходе: а) профиль показателя преломления; б) поперечное сечение лазерного пучка; в) зонная структура полупроводника с двойным гетеропереходом. Черные кружки обозначают электроны, светлые кружки — дырки

**8.5.3. Лазер на квантовых ямах.** В заключение рассмотрим кратко свойства современных полупроводниковых лазеров. Экспериментально и теоретически было показано, что с уменьшением ширины активной области  $d$  пороговая плотность тока лазера на двойном гетеропереходе вначале уменьшается, достигая минимума, а затем увеличивается. Увеличение порогового тока наступает, когда толщина  $d$  становится очень малой и поле волны уже не удерживается внутри активного слоя (см. рис.8.28, а). При этом крылья пучка света испытывают существенные потери в  $p$ - и  $n$ -областях гетероперехода. Тем не менее, в современных полупроводниковых лазерах удастся уменьшить пороговую плотность тока более чем на порядок по сравнению с лазерами на двойном гетеропереходе. Это достигается за счет того, что создаются лазеры с раздельными оптическим и электронным ограничением (рис. 8.29). В таком лазере ширина активной области  $d$  (то есть области локализации электронов и дырок) существенно меньше поперечного размера волновода  $\tilde{d}$ . Если ширина активной области  $d$  становится порядка длины волны де-Бройля электрона (дырки), то спектр носителей заряда квантуется. Такие тонкие слои полупроводника получили название квантовых ям, а лазеры, включающие такие тонкие слои, получили название лазеров на квантовых ямах. Лазеры на квантовых ямах имеют достаточно низкие пороговые токи и обладают предельно высокой мощностью излучения по сравнению с лазерами на двойном

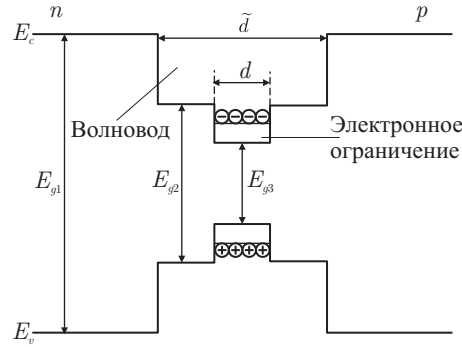


Рис. 8.29. Зонная структура полупроводникового лазера на квантовых ямах с раздельным оптическим и электронным ограничением

гетеропереходе. В настоящее время такие лазеры являются лучшими (по многим параметрам) среди полупроводниковых лазеров<sup>1)</sup>.

### 8.6. Влияние магнитного поля на межзонное поглощение света

Согласно классической механике на электрон, движущийся со скоростью  $\mathbf{v}$  в магнитном поле с напряженностью  $\mathbf{H}$ , действует сила Лоренца

$$\mathbf{F} = \frac{e}{c}[\mathbf{v}\mathbf{H}]. \quad (8.76)$$

Здесь  $e$  — это заряд электрона, а  $c$  — скорость света. Сила  $\mathbf{F}$  перпендикулярна скорости электрона и, следовательно, не изменяет его энергии. В результате электрон движется по спирали с осью вдоль магнитного поля и с постоянной составляющей скорости  $\mathbf{v}_{\parallel}$  вдоль поля. Мы будем интересоваться в дальнейшем только движением электрона в плоскости, перпендикулярной магнитному полю  $\mathbf{H}$ . Проекция его траектории на плоскость, перпендикулярную полю, является окружностью с ларморовским радиусом

$$R = \frac{mv_{\perp}c}{eH}. \quad (8.77)$$

Здесь  $m$  — масса электрона;  $v_{\perp}$  — компонента скорости, перпендикулярная магнитному полю. Период обращения электрона вокруг оси поля — это

$$T = \frac{2\pi R}{v_{\perp}} = \frac{2\pi mc}{eH}. \quad (8.78)$$

<sup>1)</sup> Quantum Well Lasers / Ed. by P.S. Zory, Jr. — NY: Academic Press, Inc., 1993.

Круговая частота вращения

$$\omega_H = \frac{2\pi}{T} = \frac{eH}{mc} \quad (8.79)$$

и называется ларморовской, или циклотронной, частотой. Циклотронная частота зависит только от массы электрона и от величины магнитного поля  $H$ . Эта частота, что очень важно, не зависит от энергии электрона. Быстрые электроны вращаются по ларморовской орбите большого радиуса, а медленные электроны — по орбите малого радиуса, однако частота вращения одна и та же.

Квантовая теория движения электрона в магнитном поле впервые была развита Л. Д. Ландау. Энергия электрона, отвечающая финитному движению в плоскости, перпендикулярной магнитному полю  $\mathbf{H}$ , квантуется. Вдоль магнитного поля электрон движется равномерно и прямолинейно. Этому движению соответствует сплошной спектр. Пусть магнитное поле описывается векторным потенциалом  $\mathbf{A}$  (где  $\mathbf{H} = \text{rot } \mathbf{A}$ ) и направлено вдоль оси  $z$ . Выберем калибровку Ландау, согласной которой векторный потенциал имеет компоненты

$$A_x = -yH, \quad A_y = A_z = 0. \quad (8.80)$$

Тогда функция Гамильтона электрона (без учета спина) в магнитном поле выражается через оператор импульса  $\hat{\mathbf{p}}$  электрона и векторный потенциал  $\mathbf{A}$ :

$$\hat{H} = \frac{1}{2m} \left[ \left( \hat{p}_x + \frac{e}{c}Hy \right)^2 + \hat{p}_y^2 + \hat{p}_z^2 \right]. \quad (8.81)$$

Этому гамильтониану соответствует уравнение Шредингера

$$\frac{1}{2m} \left[ \left( \hat{p}_x + \frac{e}{c}Hy \right)^2 + \hat{p}_y^2 + \hat{p}_z^2 \right] \psi(x, y, z) = \varepsilon \psi(x, y, z). \quad (8.82)$$

Отметим, что магнитное поле считается постоянным, однородным и достаточно слабым, поэтому его можно учитывать в рамках метода эффективной массы. Поэтому в уравнении (8.82)  $m$  — эффективная масса. Из вида гамильтониана (8.81) следует, что  $\hat{H}$  коммутирует с  $\hat{p}_x^2$  и  $\hat{p}_z^2$ , то есть  $x$  и  $z$  компоненты обобщенного импульса сохраняются. Тогда волновую функцию электрона следует искать в виде

$$\psi(x, y, z) = \chi(y) \exp \left[ \frac{i}{\hbar} (p_x x + p_z z) \right]. \quad (8.83)$$

Здесь  $p_x$  и  $p_z$  — собственные значения соответствующих компонент оператора импульса, они пробегает все значения от  $-\infty$  до  $\infty$ .

Подставляя (8.83) в (8.82), получаем для  $\chi(y)$  обыкновенное дифференциальное уравнение второго порядка

$$\frac{d^2\chi}{dy^2} + \frac{2m}{\hbar^2} \left[ \varepsilon - \frac{p_x^2}{2m} - \frac{m}{2} \omega_H^2 (y - y_0)^2 \right] \chi = 0, \quad (8.84)$$

где

$$y_0 = -\frac{cp_x}{eH}. \quad (8.85)$$

Уравнение (8.84) совпадает по форме с уравнением Шредингера для линейного гармонического осциллятора с минимумом потенциала в точке  $y_0$ . Воспользовавшись этой аналогией, можно сразу написать выражение для энергии электрона в магнитном поле

$$\varepsilon_n = \left( n + \frac{1}{2} \right) \hbar\omega_H + \frac{\hbar^2 k_z^2}{2m}, \quad (8.86)$$

где  $n = 0, 1, 2, \dots$ . Здесь  $k_z$  — продольная компонента волнового вектора электрона. Итак, мы получили, что поперечная часть энергии квантуется. Соответствующие уровни энергии называются уровнями Ландау<sup>1)</sup>. Расстояние между уровнями Ландау равно  $\hbar\omega_H$ . При  $n = 0$  мы имеем  $E_0 = \hbar\omega_H/2$ , то есть энергия поперечного движения не равна нулю. Здесь ситуация аналогична гармоническому осциллятору. Конечное значение для энергии означает, что осциллятор не может покоиться при  $n = 0$ . Этот результат вытекает из соотношения неопределенностей Гейзенберга.

Спектр разрешенных значений энергии электрона в магнитном поле представляет собой совокупность парабол  $\frac{\hbar^2 k_z^2}{2m}$ , сдвинутых друг относительно друга по оси энергии на расстояние, равное  $\hbar\omega_H$  (рис. 8.30). Так же, как и уровни Ландау, каждое состояние на параболе вырождено. Кратность вырождения уровней Ландау определяется числом возможных значений  $p_x$ , при которых равновесное положение осциллятора  $y_0$  находится внутри полупроводника. Пусть полупроводник имеет форму параллелепипеда со сторонами  $L_x$ ,  $L_y$ ,  $L_z$ . Тогда условие того, что осциллятор находится в пределах нашей системы, имеет вид

$$0 < |y_0| < L_y.$$

Следовательно, возможные значения  $p_x$  согласно (8.85) должны лежать в интервале

$$0 < p_x < \frac{eH}{c} L_y. \quad (8.87)$$

<sup>1)</sup> Следует отметить, что все эти выводы справедливы для изотропного спектра электрона.

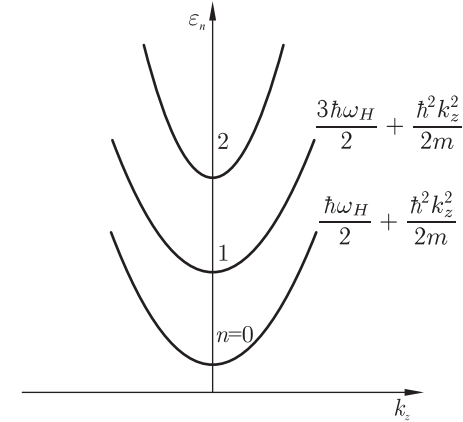


Рис. 8.30. Зонная диаграмма, изображающая магнитные уровни электрона при  $k_z \neq 0$  и различных значениях  $n$

Так как шаг дискретности  $p_x$  равен  $\frac{2\pi\hbar}{L_x}$ , то число возможных значений  $p_x$  и, следовательно, кратность вырождения уровня равны

$$\frac{eH}{2\pi\hbar c} L_x L_y. \quad (8.88)$$

Отметим одно очень важное обстоятельство. Энергия электрона в магнитном поле непрерывным образом зависит только от одной координаты импульса  $p_z$ . Вследствие этого система электронов в магнитном поле имеет многие черты одномерной системы. В связи с этим плотность электронных состояний на каждой параболе может быть описана так же, как и в случае одномерного кристалла длиной  $L_z$ . Воспользуемся граничными условиями Борна–Кармана. Тогда разрешенные значения волнового вектора  $k_z$  равны

$$k_z = \frac{2\pi}{L_z} m, \quad (8.89)$$

где  $m = 1, 2, 3, \dots$  — целое число. Расстояние между «дискретными» значениями  $k_z$  равно

$$\delta k_z = \frac{2\pi}{L_z}. \quad (8.90)$$

Таким образом, на каждое разрешенное значение волнового вектора в одномерном  $k_z$ -пространстве приходится интервал  $\delta k_z$ . Нетрудно посчитать, сколько разрешенных значений волновых векторов  $N(k_z)$  находится в достаточно большом интервале  $dk_z$ . Для этого необходимо разделить  $dk_z$  на один интервал  $\delta k_z$ :

$$N(k_z) = \frac{L_z}{2\pi} dk_z. \quad (8.91)$$



Далее, нам нужно найти число состояний с энергиями от  $\varepsilon_z$  до  $\varepsilon_z + d\varepsilon_z$ , где  $\varepsilon_z = \hbar^2 k_z^2 / 2m$ . Число состояний в интервале энергий  $d\varepsilon_z$  (по определению) равно  $\rho(\varepsilon_z) d\varepsilon_z$ . Это же число равно  $N(k_z)$ . Тогда число состояний, приходящихся на единичный интервал энергии, принимает вид

$$\rho(\varepsilon_z) = \frac{\sqrt{m} L_z}{2\sqrt{2} \pi \hbar \sqrt{\varepsilon_z}}. \quad (8.92)$$

Далее, необходимо учесть, что в магнитном поле каждое состояние с данным  $k_z$  имеет кратность вырождения (см. (8.88)). Принимая во внимание это обстоятельство, получаем выражение для числа состояний на  $n$ -й параболе Ландау на единицу объема, приходящегося на единичный интервал энергии,

$$\tilde{\rho}_n(\varepsilon) = \frac{eH}{2c} \frac{\sqrt{2m}}{(2\pi\hbar)^2} \frac{1}{\left[ \varepsilon_n - \hbar\omega_H \left( n + \frac{1}{2} \right) \right]^{1/2}}. \quad (8.93)$$

Здесь мы выразили  $\varepsilon_z$  через  $\varepsilon_n$  согласно (8.86). Чтобы получить полную плотность состояний  $\rho_H(\varepsilon_n)$  электронной системы в магнитном поле для данного значения энергии  $\varepsilon_n$ , достаточно просуммировать выражение (8.93) для  $\tilde{\rho}_n(\varepsilon)$  по всем параболам, для которых аргумент  $\varepsilon_n - \hbar\omega_H \left( n + \frac{1}{2} \right)$  является положительной величиной. Таким образом, полная плотность состояний  $\rho_H(\varepsilon_n)$  в магнитном поле для электронной системы с изотропным квадратичным законом дисперсии без учета спинового расщепления уровней Ландау может быть записана в виде

$$\rho_H(\varepsilon_n) = \frac{m^{3/2}}{\sqrt{2}} \frac{\omega_H}{(2\pi\hbar)^2} \sum_n \left[ \varepsilon_n - \hbar\omega_H \left( n + \frac{1}{2} \right) \right]^{-1/2}. \quad (8.94)$$

Максимальное значение  $n$  для данного уровня энергии  $\varepsilon_n$  определяется из условия положительности подкоренного выражения (8.94).

Итак, мы получили, что, когда  $\varepsilon_n = \hbar\omega_H \left( n + \frac{1}{2} \right)$  (то есть на дне каждой зоны Ландау), плотность состояний  $\rho_H$  обращается в бесконечность, поскольку плотность состояний имеет корневую особенность при  $\varepsilon_z = 0$ . Полная плотность состояний показана на рис.8.31. На этом же рисунке пунктиром изображена зависимость зонной плотности состояний  $\rho(\varepsilon)$  от энергии при  $H = 0$ . Эти особенности плотности состояний в магнитном поле приводят к важным следствиям. Чтобы в этом убедиться, рассмотрим поглощение света прямозонным полупроводником в квантующем магнитном поле.

В зоне проводимости и в валентной зоне при наличии магнитного поля появляются уровни (зоны) Ландау. Ширина запрещенной зоны полупроводника в магнитном поле эффективно увеличивается на величину  $\frac{\hbar\omega_H^e}{2} + \frac{\hbar\omega_H^h}{2}$ , где  $\omega_H^e = \frac{eH}{m_e c}$ ,  $\omega_H^h = \frac{eH}{m_h c}$ , а  $m_e$  и  $m_h$  —

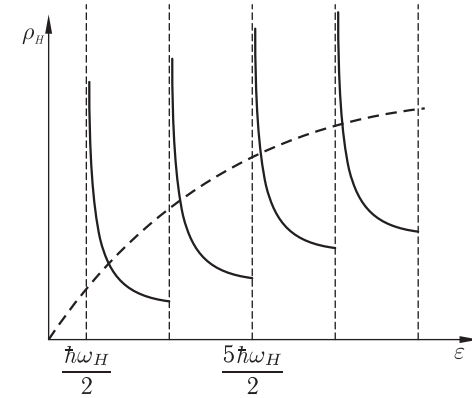


Рис. 8.31. Плотность состояний  $\rho_H(\varepsilon)$  как функция энергии  $\varepsilon$  для электронов в магнитном поле. При  $k_z = 0$  плотность состояний имеет корневую особенность в каждой точке, где энергия  $\varepsilon_n = \left( n + \frac{1}{2} \right) \hbar\omega_H$ . Пунктирная кривая соответствует зависимости плотности состояний от энергии при  $H = 0$

эффективные массы электрона и дырки соответственно. Это означает, что край фундаментального поглощения света смещается в коротковолновую сторону (рис. 8.32). Однако оптические переходы с поглощением кванта света по-прежнему будут вертикальными. Мы будем отсчитывать энергию от вершины валентной зоны при  $H = 0$ . Для прямых оптических переходов с поглощением кванта света частоты  $\omega$  закон сохранения энергии определяется соотношением

$$E_g + \hbar\omega_H^c \left( n_c + \frac{1}{2} \right) + \frac{\hbar^2 k_z^2}{2m_c} + \hbar\omega_H^h \left( n_h + \frac{1}{2} \right) + \frac{\hbar^2 k_z^2}{2m_h} = \hbar\omega, \quad (8.95)$$

где  $E_g$  — ширина запрещенной зоны полупроводника в отсутствие магнитного поля. Для определения коэффициента поглощения света нужно вычислить матричный элемент перехода электрона из валентной зоны ( $v$ ) в зону проводимости ( $c$ )

$$V_{vc} = \int \psi_c^* \hat{H}' \psi_v d^3r, \quad (8.96)$$

где гамильтониан возмущения  $\hat{H}'$  при наличии поля световой волны и постоянного магнитного поля  $H$  имеет вид

$$\hat{H}' = -\frac{e}{m_0 c} \mathbf{A}'(\mathbf{r}, t) \left( \hat{\mathbf{p}} - \frac{e}{c} \mathbf{A} \right). \quad (8.97)$$

Здесь  $\mathbf{A}'(\mathbf{r}, t)$  — векторный потенциал электромагнитной волны, зависящий от времени и координаты по гармоническому закону. Воспользовавшись (8.83), напомним выражение для волновой функции электрона



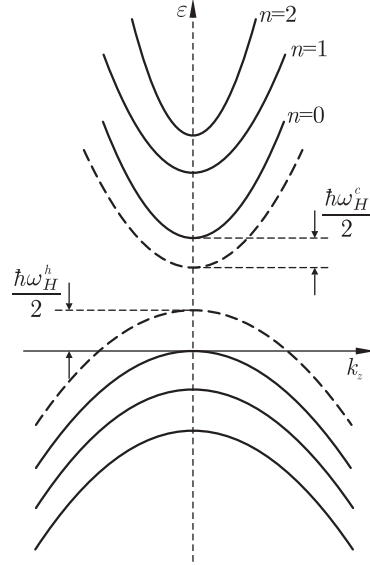


Рис. 8.32. Энергетический спектр электронов и дырок в магнитном поле для разных значений  $n_e$  и  $n_h$ . Пунктирные кривые соответствуют зонной структуре полупроводника при  $H = 0$ . Отмечено эффективное увеличение ширины запрещенной зоны полупроводника в магнитном поле на величину  $(\hbar\omega_H^e + \hbar\omega_H^h)/2$

в кристалле в присутствии магнитного поля:

$$\psi_c(x, y, z) = u_{c0}(\mathbf{r}) \exp \left[ \frac{i}{\hbar} (p_x x + p_z z) \right] \cdot \chi_n(y). \quad (8.98)$$

Здесь  $u_{c0}(\mathbf{r})$  — блоховская амплитуда волновой функции электрона в точке  $\mathbf{p} = 0$ . Выражение, аналогичное (8.98), имеет место и для волновой функции носителей в валентной зоне  $\psi_v$ . Отметим, что волновые функции  $\psi_c$  и  $\psi_v$  содержат функции  $u_{c0}(r)$  и  $u_{v0}(r)$ , быстро меняющиеся на протяжении постоянной решетки кристалла, и медленно меняющиеся функции  $\chi_n(y) \exp \left[ \frac{i}{\hbar} (p_x x + p_z z) \right]$ . Поэтому удобно переходить в (8.96) от интегрирования по всей области кристалла к интегрированию по объемам элементарных ячеек кристалла  $\Omega_0$ . В результате матричный элемент оптического перехода дается выражением

$$V_{vc} = -\frac{e}{m_0 c} A_0 (\mathbf{e} \mathbf{P}_{vc}) \delta_{p_x p'_x} \delta_{p_z p'_z} \delta_{n_e n_h}. \quad (8.99)$$

Здесь  $A_0$  — амплитуда векторного потенциала световой волны;  $\mathbf{e}$  — единичный вектор электрического поля  $\mathbf{E}$  волны,

$$\mathbf{P}_{vc} = -\frac{i\hbar}{\Omega_0} \int d^3 r u_{c0}^*(\mathbf{r}) \nabla u_{v0}(\mathbf{r}). \quad (8.100)$$

Из выражения (8.99) для матричного элемента следует, что оптические переходы являются вертикальными, о чем свидетельствуют символы Кронекера  $\delta_{p_x p'_x}$  и  $\delta_{p_z p'_z}$ . Кроме того, мы пришли к тому, что оптические переходы возможны между уровнями Ландау для электронов и дырок с одинаковыми квантовыми числами ( $n_e = n_h$ ). Это так называемое правило отбора для оптических переходов в квантующем магнитном поле. В выражение (8.99) для матричного элемента  $V_{vc}$  из-за ортогональности функций  $u_{c0}(\mathbf{r})$  и  $u_{v0}(\mathbf{r})$  не вошла часть гамильтониана  $\hat{H}'$ , пропорциональная  $(\mathbf{A}' \cdot \mathbf{A})$ .

Полное число оптических переходов из валентной зоны в зону проводимости в единицу времени в единице объема с учетом закона сохранения энергии равно

$$W_{vc} = \frac{2\pi}{\hbar} \sum_{n, p_x, p_z} |V_{vc}|^2 \delta(\varepsilon_c - \varepsilon_v - \hbar\omega). \quad (8.101)$$

Подставим в (8.101) явное выражение для  $V_{vc}$  из (8.99), разделим полученное выражение на плотность потока фотонов  $N \cdot v = Nc/n$ , где  $N$  — плотность фотонов,  $c$  — скорость света в вакууме,  $n$  — показатель преломления полупроводника. В результате получим выражение для коэффициента поглощения света в квантующем магнитном поле

$$\begin{aligned} \alpha(H) &= \frac{W_{vc}}{Nv} = \frac{2\pi}{\hbar} \frac{e^2}{m_0^2 c^2} \frac{|A_0|^2}{Nv} |\mathbf{e} \cdot \mathbf{P}_{vc}|^2 \times \\ &\times \sum_n \int \frac{dp_x}{2\pi\hbar} \int \frac{dp_z}{2\pi\hbar} \delta \left[ \hbar\omega - E_g - \frac{p_z^2}{2\mu} - \hbar\omega_H^\mu \left( n + \frac{1}{2} \right) \right] = \\ &= \frac{e^2}{\hbar c} \frac{|\mathbf{e} \cdot \mathbf{P}_{vc}|^2}{\sqrt{2} n \cdot m_0} \frac{\omega_H^\mu}{\omega} \frac{\mu^{3/2}}{m_0 \hbar} \sum_n \left[ \hbar\omega - E_g - \hbar\omega_H^\mu \left( n + \frac{1}{2} \right) \right]^{-1/2}. \end{aligned} \quad (8.102)$$

Здесь мы учли, что

$$\frac{1}{\mu} = \frac{1}{m_e} + \frac{1}{m_h} \quad \text{и} \quad \omega_H^\mu = \frac{eH}{\mu c}.$$

Отметим, что выражение, полученное для коэффициента поглощения в магнитном поле, справедливо для таких значений номера  $n$ , для которых подкоренное выражение в (8.102) положительно,

$$\hbar\omega - E_g - \frac{p_z^2}{2\mu} - \hbar\omega_H^\mu \left( n + \frac{1}{2} \right) > 0. \quad (8.103)$$

Значения  $n$ ,  $H$  и  $\omega$ , для которых неравенство (8.103) переходит в равенство, определяют сингулярные точки коэффициента поглощения. Следовательно, при увеличении магнитного поля  $H$  коэффициент поглощения  $\alpha(H)$  является осциллирующей функцией магнитного поля. Поэтому в спектре поглощения наблюдается множество пиков (рис. 8.33). Расстояние между пиками равно  $\hbar(\omega_H^e + \omega_H^h) = \hbar\omega_H^\mu$ . Этот

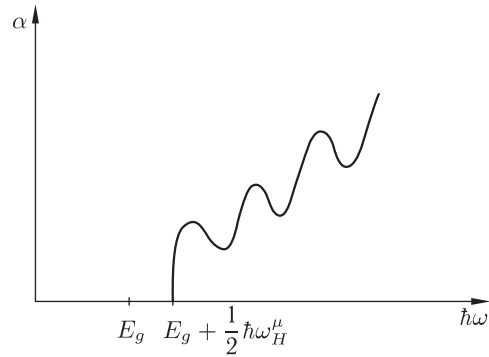


Рис. 8.33. Частотная зависимость коэффициента поглощения света электронами в магнитном поле согласно (8.102). В магнитном поле край полосы поглощения смещается на  $\hbar\omega_H^\mu/2$ . Согласно (8.102) коэффициент поглощения стремится к бесконечности, когда энергия фотона становится равной минимальной энергии одного из переходов в сумме. При учете рассеяния, коэффициент поглощения состоит из последовательности конечных пиков поглощения, то есть приобретает осциллирующий характер

эффект при низких температурах и больших магнитных полях наблюдается в ряде полупроводников. Он называется осцилляторным магнитно-абсорбционным эффектом, или магнетопоглощением. По осцилляциям коэффициента поглощения можно экспериментально измерить приведенную эффективную массу электрона и дырки  $\mu$ .

Если построить графики зависимости положения этих пиков (максимумов) от напряженности магнитного поля  $H$ , то мы должны получить серию прямых линий, сходящихся к значению  $\hbar\omega = E_g$  и имеющих наклон  $(n + 1/2)\hbar\omega_c^\mu$  (см. формулу (8.103)) (рис. 8.34). Такое именно поведение измеренных значений энергии и было найдено экспериментально. Экстраполяция экспериментальных кривых к  $k \rightarrow 0$  дает точное значение ширины запрещенной зоны полупроводника при  $\mathbf{k} = 0$ .

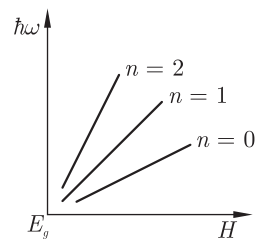


Рис. 8.34. Зависимость положения пиков поглощения от магнитного поля  $H$ . Для каждого максимума зависимость между  $\hbar\omega$  и  $H$  линейна. По наклону прямой  $n = 0$  можно определить эффективную массу электрона. Точность этого метода выше, чем какого-либо другого метода

На опыте коэффициент поглощения в точках, где выражение (8.103) переходит в равенство, имеет конечное значение. Это связано с тем, что в реальном кристалле взаимодействие электронов с фононами и дефектами решетки приводит к размыванию уровней Ландау, ширина уровня  $\Delta\varepsilon \gtrsim \hbar/\tau$ , где  $\tau$  — характерное время рассеяния частиц. Размывание же уровней Ландау, в свою очередь, приводит к размыванию пиков коэффициента поглощения.

Экспериментальная зависимость коэффициента поглощения от магнитного поля часто бывает более сложной, чем на рис. 8.33. Кривая поглощения состоит не только из пиков, но и из ступенек. Это связано с тем, что наряду с прямыми разрешенными переходами, рассмотренными выше, могут существовать не прямые переходы с участием фононов и запрещенные переходы, которые ведут себя по-разному в зависимости от взаимной ориентации поляризации света  $\mathbf{e}$  и магнитного поля  $\mathbf{H}$ <sup>1)</sup>.

**Замечание.** Осцилляторный магнитно-абсорбционный эффект подобен циклотронному резонансу (см. главу 9), поскольку и в том и в другом эффектах происходит переход электрона между уровнями Ландау. Однако при циклотронном резонансе происходит переход электронов между уровнями Ландау одной и той же зоны, а при магнетопоглощении — между уровнями Ландау различных зон. Энергии фотонов, необходимых для таких переходов, отличаются на величину порядка ширины запрещенной зоны полупроводника.

В заключение отметим, что раздел физики, изучающий оптические явления в магнитном поле, называется магнетооптикой.

## 8.7. Влияние электрического поля на край фундаментального поглощения света. Эффект Келдыша–Франца

При помещении полупроводника во внешнее (постоянное и однородное) электрическое поле изменяются и спектры поглощения, и спектры излучения света. Теоретически и экспериментально показано, что постоянное электрическое поле  $\mathcal{E}$ , приложенное к полупроводнику, приводит к сдвигу края фундаментального поглощения света в сторону меньших энергий квантов (рис. 8.35). Как продемонстрировано в главе 6, уровни энергии электрона в электрическом поле (штарковская лестница) не наблюдаются, поскольку частота осцилляций  $e\mathcal{E}a/\hbar$  (где  $a$  — постоянная решетки) при обычных полях много меньше, чем обратное время релаксации электрона  $\tau^{-1}$ . В этом случае влияние

<sup>1)</sup> См. Ансельм А. И. Введение в теорию полупроводников. — М.: Наука, 1978.

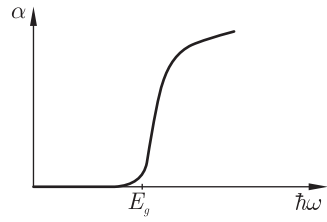


Рис. 8.35. Влияние электрического поля на коэффициент поглощения вблизи края фундаментального поглощения, обусловленного прямыми переходами. При увеличении напряженности электрического поля имеет место сдвиг края собственного поглощения в сторону меньших энергий фотонов

ную энергию электрона во внешнем поле можно рассматривать как почти постоянную, и тогда мы можем представить собственные значения энергии в виде

$$E_i = E_i(\mathbf{k}) + |e|\mathcal{E}x. \quad (8.106)$$

Здесь  $i = c, v$ , где значки  $c$  и  $v$  отвечают зоне проводимости и валентной зоне соответственно. Таким образом, локально здесь сохраняется представление о зонном энергетическом спектре электрона (дырки). Однако на больших интервалах  $x$  изменение второго слагаемого в (8.106) становится заметным. Это означает, что зоны кристалла, оставаясь практически неизменными по форме, должны наклоняться так, как это показано на рис. 8.36. На этом рисунке изображены только границы двух зон — проводимости и валентной (по вертикальной оси отложена энергия электрона  $E$ , а по горизонтальной — пространственная координата  $x$ ).

На рис. 8.36 представлена энергия электрона во внешнем электрическом поле (см. (8.106)). Если энергия кванта света  $\hbar\omega$  меньше ширины запрещенной зоны  $E_g$  ( $\hbar\omega < E_g$ ), то на первый взгляд кажется, что электрону не хватает энергии, чтобы перейти из валентной зоны в зону проводимости. Однако благодаря тому, что электрон может туннелировать под барьер, вероятность обнаружить его в запрещенной зоне становится отличной от нуля при наличии электрического поля, при этом электрон может перейти из валентной зоны в зону проводимости даже при условии, что  $\hbar\omega < E_g$  (см. рис. 8.36). Следует также отметить, что ширина запрещенной зоны  $E_g$  в каждой точке  $x$  одна и та же. Поэтому рассматриваемые переходы не являются прямыми в координатном пространстве.

электрического поля можно учитывать по теории возмущений. Тогда потенциальная энергия электрона имеет вид

$$U(\mathbf{r}) = V(\mathbf{r}) + |e|\mathcal{E}x. \quad (8.104)$$

Здесь  $V(\mathbf{r})$  — потенциальная энергия электрона в кристалле;  $e < 0$  — заряд электрона, а электрическое поле  $\mathcal{E}$  направлено вдоль оси  $x$ . Поскольку напряженность внешнего электрического поля  $\mathcal{E}$ , действующего на электрон, удовлетворяет неравенству

$$|e|\mathcal{E}a \ll E_g, \quad (8.105)$$

то второе слагаемое в (8.104) весьма мало по сравнению с первым. Следовательно, локально (в каждой не слишком большой области кристалла) потенциаль-

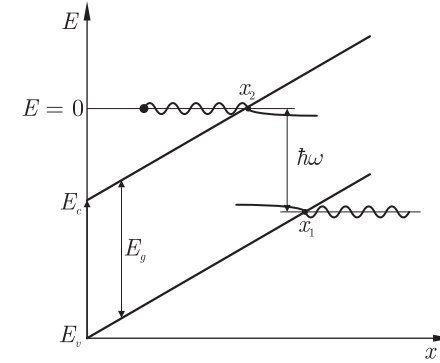


Рис. 8.36. Зависимость энергии электрона от пространственной координаты  $x$ . В электрическом поле энергетические зоны электрона наклоняются так, как показано на рисунке. Благодаря волновым свойствам, электрон может посредством туннельного эффекта проникнуть вглубь запрещенной зоны. Вероятность туннелирования экспоненциально растет с ростом напряженности электрического поля. Значения  $x_1$  и  $x_2$  — это классические точки поворота электрона. Также отмечены «хвосты» волновых функций электрона и оптический переход между ними

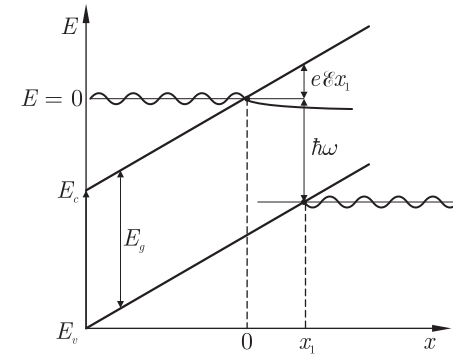


Рис. 8.37. Зависимость энергии электрона от координаты  $x$ . Стрелкой показан переход электрона из валентной зоны и «хвостом» волновой функции электрона из зоны проводимости. На рисунке также отмечен дефект энергии кванта равный  $|e|\mathcal{E}x_1$

Рассмотрим физический механизм сдвига края фундаментального поглощения света в электрическом поле. Согласно классической механике точки  $x_1$  и  $x_2$  на рис. 8.36 — это точки поворота электрона. Это те точки, в которых частица остановилась бы, после чего начала бы двигаться в обратном направлении. Точки поворота определяются из равенства нулю импульса электрона

$$p(x) = [2mE - U(x)]^{1/2} = 0. \quad (8.107)$$

Здесь  $E$  — полная энергия электрона;  $U(x) = |e|\mathcal{E}x$  — потенциальная энергия электрона в электрическом поле. Но согласно квантовой механике электрон может туннелировать под барьером, то есть электрон можно обнаружить в запрещенной зоне, и если энергия фотона  $\hbar\omega < E_g$ , то квант света может индуцировать оптический переход между «хвостами» волновых функций электрона (см. рис. 8.36).

Рассмотрим вероятность таких индуцированных переходов между хвостами волновых функций, «протянутых» под барьерами. Сразу можно понять, что вероятность таких переходов должна быть пропорциональной перекрытию волновых функций электрона в валентной зоне (состояние «1») и электрона в зоне проводимости (состояние «2»). Поскольку волновая функция электрона убывает по мере проникновения вглубь запрещенной зоны, нас больше всего интересует экспоненциальная часть вероятности  $W$ , то есть, насколько сильно перекрываются волновые функции электрона и дырки в запрещенной зоне. Предположим для простоты, что эффективная масса дырки много больше эффективной массы электрона,  $m_h \gg m_e$ . Тогда можно считать, что туннелирует под барьер только электрон в зоне проводимости.

Итак, пусть дырка в валентной зоне имеет энергию  $E_1$  (см. рис. 8.37, линия до пересечения с линией  $E_v$ ) и проникновением ее волновой функции в запрещенную зону пренебрегаем, а электрон в зоне проводимости также имеет энергию  $E$  и проникает под барьер (в запрещенную зону). Таким образом, вероятность туннелирования будет пропорциональна волновой функции электрона в точке  $x_1$ . Итак,

$$W \propto |\psi_c(x_1)|^2. \quad (8.108)$$

Поскольку речь идет о туннелировании электрона под барьером, то  $\psi_c$  — это волновая функция туннелирующего электрона и она, согласно квазиклассическому приближению, пропорциональна туннельной экспоненте:

$$\psi_c(x_1) \propto \exp \left[ -\frac{1}{\hbar} \int_0^{x_1} |p(x)| dx \right]. \quad (8.109)$$

В уравнении (8.109) мы выписали лишь самую сильную, экспоненциальную зависимость волновой функции от координаты. Сюда входит модуль  $p(x)$ , так как кинетическая энергия электрона под барьером отрицательна и функция  $p(x)$  — чисто мнимая:

$$|p(x)| = \sqrt{2m(U(x) - E)}. \quad (8.110)$$

В этой области (под барьером)  $E < U(x)$ . Поместим начало отсчета энергии  $E = 0$  в точку  $x_1$ . Поскольку потенциальная энергия электрона в электрическом поле  $U(x) = |e|\mathcal{E}x$ , то для показателя экспоненты,

входящего в (8.109), имеем

$$\int_0^{x_1} |p(x)| dx = \int_0^{x_1} \sqrt{2m(U(x) - E)} dx = \frac{2}{3} (2m|e|\mathcal{E})^{1/2} x_1^{3/2}. \quad (8.111)$$

Так как в точке  $x = x_1$  энергия  $E = 0$ , то расстояние (по энергии) от линии  $E = 0$  до пересечения с линией  $E_c$  равно  $|e|\mathcal{E}x_1$  и (см. рис. 8.37)

$$|e|\mathcal{E}x_1 = E_g - \hbar\omega. \quad (8.112)$$

Величина  $|e|\mathcal{E}x_1$  называется дефицитом энергии кванта. Подставим  $x_1$  из (8.112) в (8.111), тогда

$$\int_0^{x_1} |p(x)| dx = \frac{2}{3} \frac{(2m)^{1/2}}{e\mathcal{E}} (E_g - \hbar\omega)^{3/2}. \quad (8.113)$$

Это квазиклассический расчет показателя экспоненты. Квазиклассическое приближение справедливо, когда показатель экспоненты велик. Здесь именно этот случай и имеет место. Мы пока вычислили только основную, экспоненциальную, зависимость вероятности от энергии кванта и соответственно от напряженности электрического поля. Подставим (8.113) в (8.109) и возведем волновую функцию электрона в квадрат. Тогда для вероятности перехода, согласно (8.108), получаем

$$W \propto \exp \left\{ -\frac{4}{2} \frac{(2m)^{1/2}}{e\mathcal{E}\hbar} (E_g - \hbar\omega)^{3/2} \right\}. \quad (8.114)$$

Выражение (8.114) для вероятности подтверждается опытом. Видно, что при  $\hbar\omega < E_g$ , то есть ниже края фундаментального поглощения, вероятность  $W$  отлична от нуля. Таким образом при  $\hbar\omega < E_g$  существует «хвост» коэффициента поглощения (рис. 8.35). Чем больше дефицит кванта (см. (8.113)), тем меньше поглощение в области ниже края фундаментального поглощения, то есть при  $\hbar\omega < E_g$ . Можно сказать, что при поглощении фотона с энергией  $\hbar\omega < E_g$  дефицит энергии кванта покрывается работой электрического поля. Из (8.114) также следует, что чем меньше напряженность электрического поля  $\mathcal{E}$ , тем меньше вероятность поглощения за краем фундаментального поглощения. Очевидно, что выражение (8.114) описывает поглощение излучения слева (при  $\hbar\omega < E_g$ ) за краем фундаментального поглощения. Кроме того, чем больше эффективная масса частицы, тем меньше вероятность поглощения в этой области. Выражение для вероятности (8.114) можно обобщить на случай, когда дырка также туннелирует под барьером. Тогда в (8.114) под  $m$  нужно подразумевать приведенную массу электрона и дырки  $1/m = 1/m_e + 1/m_h$ .

Найдем сдвиг края фундаментального поглощения в длинноволновую область, который происходит в электрическом поле. Этот сдвиг

качественно можно определить из условия, при котором вероятность поглощения  $W$  изменится в  $\epsilon$  раз (то есть примерно в 3 раза),

$$E_g - \hbar\omega \simeq \left[ \frac{3|e|\mathcal{E}\hbar}{2(2m)^{1/2}} \right]^{2/3}. \quad (8.115)$$

Сдвиг края фундаментального поглощения в электрическом поле впервые был рассмотрен Л.В. Келдышем и В. Францем в 1958 г. Сам эффект назван их именами.

### Замечания

**Замечание 1.** Эффект Келдыша–Франца лежит в основе метода модуляционной спектроскопии. Представим, что мы изучаем коэффициент поглощения полупроводника в широкой области спектра. Как правило, мы получаем полосы, и не так просто узнать, где находится край фундаментального поглощения. Для этого к кристаллу прикладывают переменное электрическое поле и находят разность коэффициента поглощения во внешнем поле и без поля. Эта разность больше в той области частот, где электрическое поле сильнее влияет на поглощение излучения. А электрическое поле сильнее влияет на коэффициент поглощения  $\alpha(\omega)$  вблизи края фундаментального поглощения. Тогда в той области частот, где величина

$$\Delta\alpha^{(\omega)} = \alpha_{\text{без поля}}^{(\omega)} - \alpha_{\text{с полем}}^{(\omega)} \quad (8.116)$$

имеет максимум, и находится край фундаментального поглощения полупроводника.

**Замечание 2.** Эффект Келдыша–Франца оказался важным для объяснения механизма поглощения излучения в сильнолегированных полупроводниках. Оказывается, в легированных полупроводниках коэффициент поглощения  $\alpha(\omega)$  отличен от нуля в области энергий  $\hbar\omega < E_g$ . Этот результат можно понять исходя из следующих соображений.

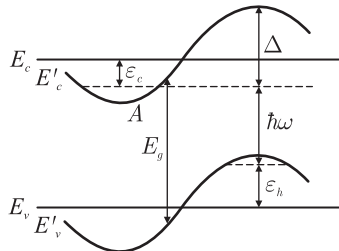


Рис. 8.38. Схема поглощения кванта света с дефицитом  $\Delta$  в примесном полупроводнике. Сплошные горизонтальные прямые соответствуют границам невозмущенных зон проводимости и валентной. Изогнутые линии — границы зон проводимости и валентной, искривленные потенциалом примесей. Стрелкой отмечен возможный оптический переход электрона из валентной зоны в зону проводимости при поглощении кванта света  $\hbar\omega$

Если в полупроводнике имеются заряженные примеси (доноры и акцепторы), случайно расположенные в пространстве, то они создают случайные (флуктуирующие) электрические поля. В результате при наличии флуктуирующего потенциала примесей зоны полупроводника искривляются (рис. 8.38).

Как видно из рис. 8.38, если  $\Delta \equiv E_g - \hbar\omega > 0$ , то области, в которых находятся электрон и дырка, пространственно разделены. При этом область правее точки  $A$  является классически недоступной для электрона, а область левее точки  $A$  — классически недоступной для дырки. В этом случае коэффициент поглощения света содержит интеграл перекрытия волновых функций электрона и дырки, туннелирующих под барьером. Таким образом, при  $\Delta > 0$  коэффициент поглощения  $\alpha(\Delta)$  экспоненциально мал, но отличен от нуля. По существу, это есть эффект Келдыша–Франца, но не во внешнем электрическом поле, а в случайном поле заряженных примесей.

**Замечание 3. Диагональное туннелирование в  $p-n$ -переходе в электрическом поле.** Рассмотрим еще один аспект эффекта Келдыша–Франца на примере излучательной рекомбинации в  $p-n$ -переходе. На рис. 8.39 приведена упрощенная диаграмма  $p-n$ -перехода, смещенного в прямом направлении. Как происходит излучательная рекомбинация в такой системе? В этом случае возможен канал прямой излучательной рекомбинации, когда электроны инжектируются в  $p$ -область (стрелка 1 на рис. 8.39) и там рекомбинируют с дыркой, излучая квант света с энергией  $\hbar\omega = E_g$ . Аналогично, дырки инжектируются в  $n$ -область (стрелка 2) и рекомбинируют с электронами. И в первом, и во втором случае излучение происходит на частоте  $\omega = E_g/\hbar$ . Экспериментально, наряду с краевой люминесценцией (когда  $\hbar\omega = E_g$ ), наблюдается (в вырожденных материалах) линия люминесценции на частоте  $\omega < E_g/\hbar$ , которая смещается в сторону больших  $\omega$  при увеличении напряжения прямого смещения. Легко понять, что в таком  $p-n$ -переходе (при прямом смещении) возможен еще один канал

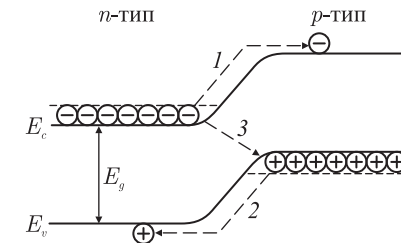


Рис. 8.39. Зонная диаграмма  $p-n$ -перехода, смещенного в прямом направлении. Наряду с прямой рекомбинацией электрона из зоны проводимости с дыркой из валентной зоны, возможен еще один канал рекомбинации — диагональная туннельная рекомбинация (стрелка 3 на рисунке). В этом случае энергия излучаемого кванта  $\hbar\omega < E_g$  и зависит от приложенного напряжения

рекомбинации — переход электрона из зоны проводимости в валентную зону (стрелка 3 на рис. 8.39). Энергия излучаемого кванта  $\hbar\omega < E_g$  зависит от приложенного напряжения. При увеличении напряжения барьеры спрямляются (уменьшаются) и энергия кванта  $\hbar\omega$  увеличивается. Такой механизм рекомбинации носителей заряда (когда электроны и дырки пространственно разделены) получил название диагонального туннелирования. Очевидно, эффект диагональной туннельной рекомбинации — это тот же эффект Келдыша–Франца<sup>1)</sup>.

### Задачи

**Задача 8.1.** Рассмотреть межзонное поглощение света в квантовой яме с бесконечными барьерами. Получить зависимость коэффициента поглощения света от частоты  $\alpha(\omega)$ . Установить, как эта зависимость переходит в  $\alpha(\omega) \propto \sqrt{\hbar\omega - E_g}$ .

*Указание.* Легко убедиться, что если барьеры в квантовой яме бесконечны, то при межзонных переходах возбуждаются электроны и дырки на одинаковых уровнях размерного квантования. Плотность состояний в каждой подзоне размерного квантования постоянна. Зависимость коэффициента поглощения от энергии будет представлять собой набор ступенек на энергиях  $\hbar^2 k^2 / 2\mu$ , где  $\mu$  — приведенная масса электрона и дырки.

**Задача 8.2.** Установить правила отбора при межзонных и межподзонных переходах в квантовой яме с бесконечно высокими барьерами для электронов и дырок (свет падает по нормали к структуре).

**Задача 8.3.** Установить зависимость коэффициента поглощения света от частоты  $\alpha(\omega)$  вблизи края поглощения для:

- 1) прямых разрешенных переходов;
- 2) прямых запрещенных переходов.

*Ответ.*

1.  $\alpha(\omega) \propto \sqrt{\hbar\omega - E_g}$ ;
2.  $\alpha(\omega) \propto (\hbar\omega - E_g)^{3/2}$ .

**Задача 8.4.** Какова частотная зависимость коэффициента поглощения света  $\alpha(\omega)$  при  $\hbar\omega > E_g$  в условиях эффекта Келдыша–Франца?

<sup>1)</sup> Формулы (8.114) и (8.115) имеют качественный характер. Более строгую теорию эффекта Келдыша–Франца можно найти в книге: Ансельм А. И. Введение в теорию полупроводников. — М.: Наука, 1978. 616 с.

## Глава 9

# КИНЕТИЧЕСКИЕ ЯВЛЕНИЯ В ПОЛУПРОВОДНИКАХ В МАГНИТНОМ ПОЛЕ

### 9.1. Циклотронный (диамагнитный) резонанс

**9.1.1. Классическое рассмотрение.** Прежде чем перейти к теоретическому исследованию кинетических явлений в полупроводниках в магнитном поле, мы познакомимся с одним из самых мощных и прямых экспериментальных методов изучения зонной структуры (в частности, определения величины эффективной массы) — методом циклотронного резонанса. Структура энергетических зон германия и кремния, описанная в главе 4, обоснована, главным образом, благодаря изучению в этих материалах циклотронного резонанса. В настоящем разделе мы рассмотрим основные физические принципы, лежащие в основе явления циклотронного резонанса, и дадим его упрощенную теорию.

Рассмотрим динамику движения электрона с изотропной эффективной массой в постоянном магнитном  $\mathbf{H}$  и высокочастотном электрическом  $\mathbf{E}$  полях. Как известно (см. §8.6), в однородном магнитном поле электрон движется по ларморовой круговой орбите в плоскости, перпендикулярной магнитному полю, с циклотронной частотой

$$\omega_H = \frac{eH}{mc}. \quad (9.1)$$

Здесь  $e$  — заряд электрона;  $m$  — эффективная масса электрона;  $c$  — скорость света. Возникает вопрос — как экспериментально обнаружить такое движение. С точки зрения классической физики электрон вращается по окружности, то есть с ускорением, поэтому он должен излучать электромагнитную энергию. С другой стороны, движение электрона по окружности с ларморовским радиусом

$$R = \frac{v_{\perp}}{\omega_H} \quad (9.2)$$

создает круговой ток (рис. 9.1), где  $v_{\perp}$  — компонента скорости, перпендикулярная магнитному полю. Каждый такой круговой ток связан с магнитным моментом соотношением

$$\mu = \frac{e}{2c} R^2 \omega_H. \quad (9.3)$$



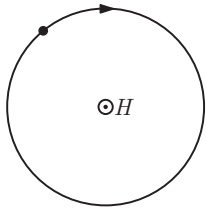


Рис. 9.1. Траектория движения электрона в магнитном поле. В случае, когда скорость электрона вдоль поля  $\mathbf{H}$  равна нулю, он движется по окружности в плоскости, перпендикулярной к полю

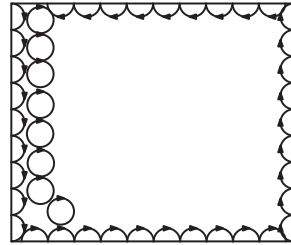


Рис. 9.2. Движение приграничных электронов вдоль стенки кристалла. Направление движения приграничных электронов противоположно циклотронному движению электронов внутри образца

Однако такой магнитный момент обнаружить экспериментально невозможно. В ограниченном образце, кроме внутренних электронов, существуют еще и приграничные электроны, которые, ударяясь о стенки кристалла, совершают (в магнитном поле) не полный ларморовский круг, а часть круга. Они двигаются скачками вдоль стенки кристалла, как видно из рис. 9.2, их движение оказывается противоположным циклотронному. Расчет показывает, что такой пристеночный (обтекающий) ток электронов направлен противоположно по отношению к току электронов внутри образца и что обтекающий (пристеночный) ток в точности компенсирует внутренний ток электронов. В результате, в образце конечных размеров в магнитном поле отсутствует суммарный магнитный момент свободных электронов. Таким образом, классическая физика утверждает, что в образце конечных размеров диамагнетизма свободных электронов не наблюдается.

Однако в квантовой механике, как было впервые показано Л. Д. Ландау, диамагнетизм свободных электронов существует (то есть существует отличный от нуля магнитный момент электронов в постоянном и однородном магнитном поле). Это связано с тем, что согласно квантовой механике волновая функция электрона на границе должна обращаться в нуль<sup>1)</sup>. Вследствие этого концентрация приграничных электронов меньше, чем объемная концентрация электронов. В результате компенсации токов (объемных и приповерхностных) не происходит, следовательно, результирующий ток порождает магнитный момент кристалла.

Рассмотрим теперь, как влияет на движение электрона высокочастотное электрическое поле  $\mathbf{E}$ , колеблющееся в плоскости, перпендикулярной к  $\mathbf{H}$ . Если частота этого поля  $\omega$  совпадает с циклотронной ча-

<sup>1)</sup> Здесь рассматривается образец конечного размера, а не система с периодическими граничными условиями.

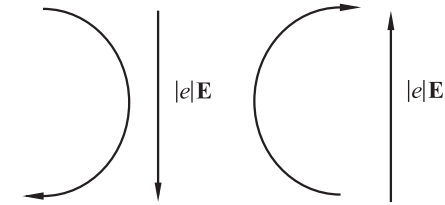


Рис. 9.3. Траектория электрона в однородном постоянном магнитном поле  $\mathbf{H}$  при действии переменного электрического поля  $\mathbf{E}$ , колеблющегося с частотой  $\omega$  в плоскости, перпендикулярной к  $\mathbf{H}$ . Электрон ускоряется электрическим полем как на одном полупериоде (а), так и на другом (б)

стотой  $\omega_H$ , то электрон будет ускоряться под действием электрического напряжения высокочастотного поля на обеих полуокружностях своего движения (см. рис. 9.3). Из рис. 9.3 видно, что на одной полуокружности (рис. 9.3, а) электрон ускоряется электрическим полем, а на другой полуокружности (то есть полупериоде) (рис. 9.3, б) электрон также ускоряется. В результате этого скорость и энергия электрона будут возрастать, и он будет двигаться по раскручивающейся спирали, как заряженная частица в циклотроне. Этот эффект, получивший название циклотронного (диамагнитного) резонанса (Дорфман и Дингль, 1951), может быть обнаружен по максимуму поглощения высокочастотного излучения. Максимум поглощения излучения будет наблюдаться как раз на частоте  $\omega = \omega_H$  (рис. 9.4). Это позволяет наиболее простым образом определить, согласно (9.1), эффективную массу электрона  $m$ . Если частота электромагнитного поля не зависит от времени ( $\omega = \text{const}$ ), то циклотронный резонанс можно наблюдать, изменяя напряженность магнитного поля  $H$  (рис. 9.5).

Явление циклотронного резонанса хорошо наблюдается экспериментально. Помимо взаимодействия с постоянным магнитным и высокочастотным полем, электрон проводимости взаимодействует еще и с колебаниями решетки (фононами) или примесными центрами, в результате чего устанавливается некоторое время свободного пробега электрона  $\tau$ . Таким образом, полуширина резонансной кривой поглощения (рис. 9.4 и 9.5) определяется частотой столкновений  $\nu = \tau^{-1}$ . Следовательно, если частота столкновений  $\nu$  больше  $\omega_H$ , то циклотронного резонанса наблюдаться не будет. Таким образом, первое условие для наблюдения циклотронного резонанса состоит в том, чтобы  $\omega_H$  было больше частоты столкновений  $\nu$ . В полупроводниках при комнатной температуре частота столкновений составляет  $\nu \sim 10^{13} \text{ с}^{-1}$ , поэтому наблюдать циклотронный резонанс при этих условиях достаточно трудно, нужны очень сильные магнитные поля. При низких температурах (ниже температуры жидкого азота  $T \approx 77 \text{ К}$ ) частота  $\nu \lesssim 10^{11} \text{ с}^{-1}$ , тогда условия наблюдения циклотронного резонанса более благоприятны. Например, в полупроводнике InSb, где эффективная масса электрона  $m \approx 0,01 m_0$ ,  $\omega_H \approx 1,6 \cdot 10^9 \cdot H \text{ с}^{-1}$  (здесь  $m_0$  — масса свободного электрона; маг-

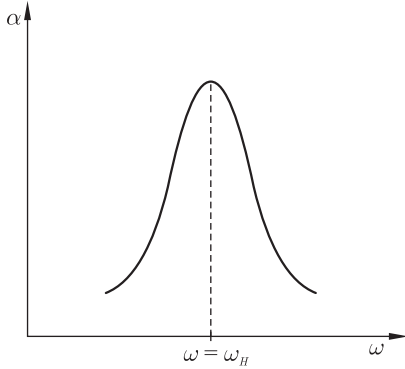


Рис. 9.4. Схематическая зависимость коэффициента поглощения излучения от частоты  $\omega$ . Максимум поглощения наблюдается на частоте, равной циклотронной частоте  $\omega = \omega_H$

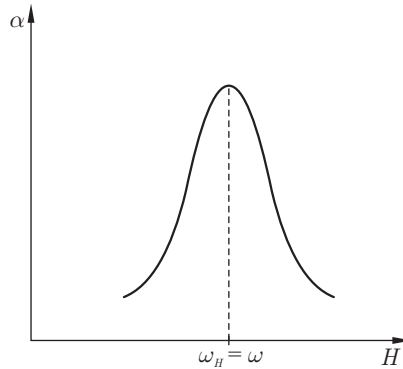


Рис. 9.5. Схематическая зависимость коэффициента поглощения излучения от магнитного поля. Максимум поглощения наблюдается при  $H = \frac{mc}{e}$

нитное поле  $H$  измеряется в эрстедах). Поэтому для наблюдения циклотронного резонанса нужны магнитные поля  $H \gtrsim (10^2 - 10^3) \text{ Э}^1$ .

Рассмотрим элементарную теорию циклотронного резонанса, считая, что электрон движется по законам классической механики. Коэффициент поглощения света  $\alpha$  пропорционален вещественной части комплексной проводимости  $\sigma'(\omega)$ . Для нахождения проводимости мы можем воспользоваться теорией Лоренца–Друде

$$\frac{d\mathbf{p}}{dt} = e\mathbf{E} + \frac{e}{c}[\mathbf{v}\mathbf{H}] - \frac{\mathbf{p}}{\tau}. \quad (9.4)$$

Это уравнение справедливо для любого закона дисперсии электрона. В (9.4) мы учли, что время свободного пробега  $\tau$  соответствует времени релаксации импульса, то есть в отсутствие внешних полей  $\mathbf{p} = \mathbf{p}_0 \exp(-t/\tau)$ , где  $\mathbf{p}_0$  — импульс электрона при  $t = 0$ .

Для простоты рассмотрим изотропный квадратичный закон дисперсии для электрона, тогда  $\mathbf{v} = \mathbf{p}/m$ . В результате уравнение (9.4) принимает вид

$$\left(\frac{1}{\tau} + \frac{d}{dt}\right) m\mathbf{v} - \frac{e}{c}[\mathbf{v}\mathbf{H}] = e\mathbf{E}. \quad (9.5)$$

Из этого уравнения следует найти скорость  $\mathbf{v}$  и подставить ее в выражение для плотности тока  $\mathbf{j}$ . Уравнение (9.5) — это линейное неоднородное уравнение. Роль вынуждающей силы играет внешнее электрическое поле  $\mathbf{E}$ . Решение неоднородного уравнения состоит из

<sup>1)</sup> В настоящее время в лабораторных условиях можно получить магнитные поля до  $10^5 \text{ Э}$ , импульсные поля — до  $10^6 \text{ Э}$ .

суммы общего решения однородного уравнения и частного решения неоднородного уравнения.

Общее решение однородного уравнения, которое зависит от начальных условий, всегда затухает со временем. Нас же интересует установившееся решение уравнения. Поэтому мы будем искать только частное решение уравнения (9.5), которое можно определить по виду неоднородности. В нашем случае поле волны зависит от времени по гармоническому закону  $\mathbf{E} \propto \exp(-i\omega t)$ , поэтому можно положить, что  $\mathbf{v} \propto \exp(-i\omega t)$ . Выберем ось  $z$  вдоль магнитного поля  $\mathbf{H}$ , а электрическое поле — в плоскости  $(x, y)$  ( $\mathbf{E} \perp \mathbf{H}$ ). В результате (9.5) превращается в систему алгебраических уравнений для комплексных амплитуд  $v_x$  и  $v_y$ :

$$\begin{aligned} \left(\frac{1}{\tau} - i\omega\right) mv_x - \frac{e}{c} H v_y &= e E_x, \\ \left(\frac{1}{\tau} - i\omega\right) mv_y + \frac{e}{c} H v_x &= e E_y. \end{aligned} \quad (9.6)$$

Систему уравнений (9.6) для  $v_x$  и  $v_y$  можно превратить в два независимых уравнения для величин  $v_{\pm} = v_x \pm i v_y$ , если второе из уравнений умножить на  $\pm i$  и сложить с первым:

$$v_{+} = \frac{\frac{e}{m} E_{+}}{\frac{1}{\tau} - i(\omega - \omega_H)}, \quad v_{-} = \frac{\frac{e}{m} E_{-}}{\frac{1}{\tau} - i(\omega + \omega_H)}. \quad (9.7)$$

Здесь  $E_{\pm} = E_x \pm i E_y$  — поле циркулярно поляризованной электромагнитной волны. Знак «+» соответствует правополяризованной волне, когда поле волны вращается в ту же сторону, что и частица с зарядом  $e$ ; знак «−» — левополяризованной волне, когда поле вращается противоположно вращению частиц (считаем, что  $\omega_H > 0$ ).

Плотность тока  $\mathbf{j}$  связана со скоростью соотношением

$$\mathbf{j} = en\mathbf{v}; \quad j_{\pm} = j_x \pm i j_y = en v_{\pm}, \quad (9.8)$$

где  $n$  — концентрация частиц. Подставляя (9.7) в (9.8), получим закон Ома для циркулярно поляризованных компонент тока

$$j_{\pm} = \frac{\sigma_0}{1 - i(\omega \mp \omega_H)\tau} E_{\pm}, \quad (9.9)$$

где  $\sigma_0 = \frac{ne^2}{m}\tau$  — статическая проводимость. Зная плотность тока, можно вычислить энергию, поглощаемую полупроводником. Джоулева энергия (средняя по времени), поглощаемая в единицу времени в еди-

нице объема полупроводника, равна <sup>1)</sup>

$$W = \frac{1}{2} \operatorname{Re}(\mathbf{j} \cdot \mathbf{E}^*). \quad (9.10)$$

В результате, поглощаемая мощность есть сумма двух компонент:

$$W = W_+ + W_-, \quad (9.11)$$

где

$$W_{\pm} = \frac{W_0}{1 + (\omega \mp \omega_H)^2 \tau^2}. \quad (9.12)$$

Здесь  $W_0 = \sigma_0 |E_x|^2$  — джоулева мощность при  $\omega = 0$ . При выводе (9.11) и (9.12) мы учли, что для циркулярно поляризованной волны  $|E_x|^2 = |E_y|^2$ . Из формул (9.9) и (9.12) видно, что выражения для тока  $j_+$  и мощности  $W_+$  расходятся бы (то есть обратились в бесконечность) при  $\omega = \omega_H$ , если бы не было рассеяния электронов ( $\tau \rightarrow \infty$ ).

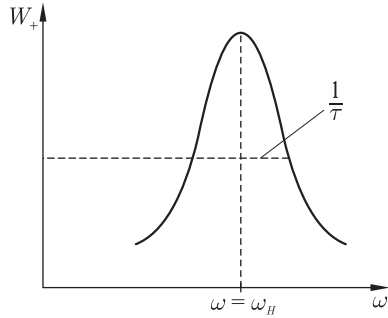


Рис. 9.6. Зависимость поглощенной мощности излучения от частоты  $\omega$ . При  $\omega_H \tau \gg 1$  поглощаемая мощность имеет острый максимум на частоте  $\omega = \omega_H$

Как уже было отмечено выше, при отсутствии соударений электрон в электрическом поле бесконечно разгоняется и стационарное состояние никогда не устанавливается. На рис. 9.6 представлена зависимость мощности  $W_+$  от частоты. Резонансная кривая (так называемый лоренцевский контур) центрирована на циклотронной частоте  $\omega_H$ . Если на половине высоты

$W_+$  отметить полуширину резонансной кривой (по оси  $\omega$ ), то ей соответствует значение  $1/\tau$ . При  $\omega_H \tau \gg 1$  поглощаемая мощность  $W_+$  имеет острый максимум как функция частоты  $\omega$ . Этот максимум тем острее, чем больше величина  $\omega_H \tau$ . Поэтому эксперименты по циклотронному резонансу надо проводить с наиболее чистыми образцами и при низких температурах. Тогда время между столкновениями  $\tau$  максимально. Схема эксперимента состоит в следующем. В заданном магнитном поле  $\mathbf{H}$  снимается частотная зависимость мощности поглощаемого излучения. Частота, при которой поглощение максимально, соответствует условию резонанса  $\omega = \omega_H$ . Зная частоту  $\omega$  и напряжен-

<sup>1)</sup> Формула (9.10) — это точное выражение, и оно справедливо независимо от тензорного вида проводимости.

ность магнитного поля  $H$ , нетрудно определить эффективную массу

$$m = \frac{eH}{c\omega}. \quad (9.13)$$

Возможна и другая постановка эксперимента. При заданной частоте  $\omega$  высокочастотного поля изменяется магнитное поле  $H$ . Максимум поглощения соответствует магнитному полю, равному

$$H = \frac{m}{e} c \omega. \quad (9.14)$$

В результате определяем эффективную массу по формуле (9.13).

**Замечание 1.** Выражение для мощности  $W_-$  при  $\omega = \omega_H$  не равно нулю. Однако при  $\omega_H \tau \gg 1$  величина  $W_- \ll W_+$ .

**Замечание 2.** Из приведенного выше расчета понятно, что если бы мы интересовались только частотой циклотронного резонанса при отсутствии столкновений, то нам достаточно было решить уравнение Ньютона, в котором оставлена только сила Лоренца:

$$m \frac{d\mathbf{v}}{dt} = \frac{e}{c} [\mathbf{v} \mathbf{H}]. \quad (9.15)$$

Это однородное уравнение имеет ненулевое решение при  $\omega = \omega_H$ . Воспользуемся этим обстоятельством и найдем решение в случае анизотропного спектра электронов с тремя различными массами по главным осям:  $m_1, m_2, m_3$ . Уравнения движения электрона в магнитном поле  $\mathbf{H}$  с проекциями  $H_1, H_2, H_3$  на главные оси согласно (9.15), при учете, что  $\mathbf{v} \propto \exp(-i\omega t)$ , имеют вид

$$\begin{aligned} -i\omega c m_1 v_1 + e(H_2 v_3 - H_3 v_2) &= 0, \\ -i\omega c m_2 v_2 + e(H_3 v_1 - H_1 v_3) &= 0, \\ -i\omega c m_3 v_3 + e(H_1 v_2 - H_2 v_1) &= 0. \end{aligned} \quad (9.16)$$

Мы получили систему трех линейных однородных уравнений. Эта система имеет отличные от нуля решения, если ее определитель равен нулю. Корни характеристического уравнения третьей степени для  $\omega$  равны

$$\begin{aligned} \omega &= 0, \\ \omega^2 &= \frac{e^2}{c^2} \left( \frac{H_1^2}{m_2 m_3} + \frac{H_2^2}{m_1 m_3} + \frac{H_3^2}{m_1 m_2} \right) \equiv \omega_H^2. \end{aligned} \quad (9.17)$$

Мы получили выражение для циклотронной частоты при произвольном направлении магнитного поля. Решение  $\omega = 0$  соответствует движению электрона вдоль магнитного поля; решение  $\omega^2 = \omega_H^2$  соответствует движению электрона в плоскости, перпендикулярной к магнитному полю. Если магнитное поле совпадает, например, с первой осью, то  $H_1 = H, H_2 = H_3 = 0$  и  $\omega_c = eH/c\sqrt{m_2 m_3}$ . Определяя циклотронную частоту (по максимуму поглощения) при разных ориентациях магнитного поля относительно кристалла, можно определить компоненты тензора эффективной массы  $m_1, m_2$  и  $m_3$ .

Как уже отмечалось выше (см. главу 6), поверхности постоянной энергии электрона в Ge и Si представляют собой эллипсоиды вращения, то есть для них справедливы соотношения  $m_1 = m_2 = m_{\perp}$  и  $m_3 = m_{\parallel}$ . Если магнитное поле  $\mathbf{H}$  образует с осью  $z$  (то есть с осью 3 в наших обозначениях) угол  $\theta$ , то  $H_3^2 = H^2 \cos^2 \theta$ ;  $H_1^2 + H_2^2 = H^2 - H_3^2 = H^2 \sin^2 \theta$ . Тогда выражение (9.17) можно записать следующим образом:

$$\omega_H^2 = \left( \frac{eH}{c} \right)^2 \left[ \frac{\cos^2 \theta}{m_{\perp}^2} + \frac{\sin^2 \theta}{m_{\parallel} m_{\perp}} \right]. \quad (9.18)$$

Измеряя  $\omega_H$  для разных углов, можно определить  $m_{\perp}$  и  $m_{\parallel}$ . Из-за того, что в  $\mathbf{k}$ -пространстве имеется несколько эквивалентных минимумов энергии, экспериментальные кривые поглощения имеют много пиков, соответствующих разным минимумам. Анализ экспериментальных данных позволяет определить, в каком направлении волнового вектора расположен минимум в зоне Бриллюэна и величины  $m_{\perp}$  и  $m_{\parallel}$ . Можно ввести *циклотронную эффективную массу*  $m_H$ , полагая, что

$$\omega_H = \frac{eH}{m_H c}. \quad (9.19)$$

Тогда, согласно (9.18),

$$\left( \frac{1}{m_H} \right)^2 = \frac{\cos^2 \theta}{m_{\perp}^2} + \frac{\sin^2 \theta}{m_{\parallel} m_{\perp}}. \quad (9.20)$$

Таким образом, цикллотронная масса  $m_H$  зависит не только от  $m_{\perp}$  и  $m_{\parallel}$ , но и от угла  $\theta$ , определяющего направление магнитного поля. На рис. 9.7 приведены экспериментальные данные по угловой зависимости цикллотронной массы  $m_H$  для германия. Из рис. 9.7 следует, что имеется только один резонансный пик поглощаемой мощности, когда магнитное поле параллельно оси  $\langle 001 \rangle$ . Это значит, что главные оси всех эллипсоидов расположены симметрично относительно этого направления, то есть главные оси направлены вдоль больших диагоналей куба  $\langle 111 \rangle$ . Это заключение подтверждается тем, что наблюдаются два пика, когда поле направлено вдоль осей  $\langle 111 \rangle$  и  $\langle 110 \rangle$ . В первом случае имеются две группы неэквивалентных эллипсоидов: один эллипсоид с осью, параллельной  $\mathbf{H}$ , и три другие эллипсоида. Во втором случае также есть две группы эллипсоидов. Отметим, что поскольку для Ge точки минимумов в  $L$  расположены на границе зоны Бриллюэна, то общее количество эллипсоидов с центрами в точке  $L$  равно четырем<sup>1)</sup>.

<sup>1)</sup> Более детальный анализ цикллотронного резонанса в полупроводниках можно найти в книгах: Ансельм А. И. Введение в теорию полупроводников. — М.: Наука, 1978; Гельмонт Б. Л., Зегря Г. Г. Основы теории полупроводников. Препринт ФТИ 1331. — Л., 1989.

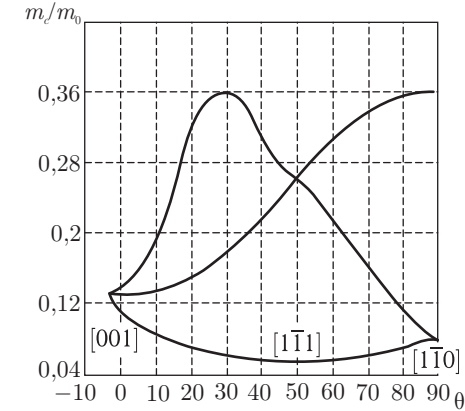


Рис. 9.7. Зависимость эффективной массы электронов в германии при  $T = 4$  К (найденное из цикллотронного резонанса) от угла между направлением магнитного поля  $\mathbf{H}$  (расположенного в плоскости  $(110)$ ) и осью  $[001]$ , лежащей в той же плоскости (Dresselhaus G., Kip A., Kittel C. // Phys. Rev. 1955. V. 100. P. 618)

В заключение этого параграфа получим общее выражение для цикллотронной массы  $m_H$ , справедливое при произвольном законе дисперсии  $\varepsilon(\mathbf{k})$ . Квазиклассические уравнения движения электрона в однородном магнитном поле (см. (4.42)) имеют вид

$$\hbar \frac{d\mathbf{k}}{dt} = \frac{e}{c} [\mathbf{vH}], \quad (9.21)$$

где скорость  $\mathbf{v}$  равна

$$\mathbf{v} = \frac{1}{\hbar} \frac{\partial \varepsilon(\mathbf{k})}{\partial \mathbf{k}}. \quad (9.22)$$

Направим ось  $z$  вдоль  $\mathbf{H}$ . Тогда из (9.21) следует, что

$$\frac{dk_z}{dt} = 0. \quad (9.23)$$

Из решения (9.23), то есть из того, что  $k_z = \text{const}$ , следует естественный вывод о том, что электрон движется вдоль магнитного поля, сохраняя проекцию квазиимпульса на направление магнитного поля  $\mathbf{H}$ . Из (9.21) и (9.22) легко получить, что

$$\frac{d\varepsilon(\mathbf{k})}{dt} = \frac{\partial \varepsilon(\mathbf{k})}{\partial \mathbf{k}} \frac{d\mathbf{k}}{dt} = \frac{e}{c} (\mathbf{v}[\mathbf{vH}]) \equiv 0. \quad (9.24)$$

Отсюда получаем, что

$$\varepsilon(\mathbf{k}) = \text{const}. \quad (9.25)$$

Следовательно, движение электрона в  $\mathbf{k}$ -пространстве происходит по кривой, определяемой пересечением поверхности постоянной энергии (9.25) с плоскостью, перпендикулярной оси  $k_z$ . Если ввести

элемент траектории, по которой происходит движение электрона в  $\mathbf{k}$ -пространстве,

$$dl = \sqrt{dk_x^2 + dk_y^2}, \quad (9.26)$$

то из (9.21) следует, что

$$\hbar \frac{dl}{dt} = \frac{e}{c} v_{\perp} H, \quad (9.27)$$

где  $v_{\perp}^2 = v_x^2 + v_y^2$ ,  $v_{\perp}$  — проекция скорости электрона на плоскость, перпендикулярную магнитному полю (то есть оси  $z$ ). Для замкнутой в  $\mathbf{k}$ -пространстве траектории период обращения электрона равен

$$T_c = \frac{2\pi}{\omega_H} = \oint dt = \frac{\hbar c}{eH} \oint \frac{dl}{v_{\perp}}. \quad (9.28)$$

В (9.28) мы подставим  $dt$  из (9.27). Далее, выразим период  $T_c$  через площадь сечения  $S$ , которая заключена внутри замкнутой траектории,

$$S = \iint dk_x dk_y = \int \frac{dk_{\perp}}{dl} dl. \quad (9.29)$$

Здесь  $k_{\perp}$  — элемент нормали к  $dl$  в плоскости  $k_z = \text{const}$ . Вектор  $\mathbf{v}$  направлен по нормали к изоэнергетической поверхности. Поэтому справедливо соотношение

$$d\varepsilon = \hbar v_{\perp} dk_{\perp}. \quad (9.30)$$

Подставляя  $ddk_{\perp}$  из (9.30) в (9.29), получим

$$S = \frac{1}{\hbar} \int d\varepsilon \oint \frac{dl}{v_{\perp}}. \quad (9.31)$$

Продифференцируем (9.31) по энергии:

$$\hbar \frac{\partial S}{\partial \varepsilon} = \oint \frac{dl}{v_{\perp}}. \quad (9.32)$$

Сравнивая (9.32) и (9.28), найдем связь между  $T_c$  и  $S$ :

$$T_c = \frac{2\pi}{\omega_H} = \frac{\hbar^2 c}{eH} \frac{\partial S}{\partial \varepsilon}. \quad (9.33)$$

Таким образом, циклотронная масса выражается через площадь сечения, заключенного внутри траектории электрона:

$$m_H = \frac{\hbar^2}{2\pi} \frac{\partial S}{\partial \varepsilon}. \quad (9.34)$$

Измеряя  $m_H$  при различных ориентациях магнитного поля относительно кристалла, в принципе можно восстановить форму поверхности

$\varepsilon(k) = \text{const}$ . Для примера рассмотрим простой случай, когда поверхность постоянной энергии представляет собой сферу, так что

$$\varepsilon = \frac{\hbar^2 k^2}{2m}, \quad (9.35)$$

где  $m$  — изотропная эффективная масса. В этом случае

$$S = \pi k_{\perp}^2 = \pi(k^2 - k_z^2) = \pi \left( \frac{2m\varepsilon}{\hbar^2} - k_z^2 \right). \quad (9.36)$$

Подставляя  $S$  в (9.34), получим  $m_H = m$ , что соответствует соотношению (9.1). Так же просто из (9.34) получить выражение (9.20) для циклотронной массы Ge. В случае сложной энергетической поверхности  $\varepsilon(\mathbf{k}) = \text{const}$  производная  $\partial S / \partial \varepsilon$ , а следовательно, и  $m_H$ , может зависеть от  $k_z$ .

В заключение также отметим, что все полученные нами выражения для циклотронной массы были найдены в квазиклассическом приближении. В сильных магнитных полях при достаточно низких температурах необходима квантовомеханическая теория, которую мы изложим ниже.

**9.1.2. Квантовая теория циклотронного резонанса. Мазер на циклотронном резонансе.** Квантовая теория циклотронного резонанса была разработана в работе Латтинжера<sup>1)</sup>. Как уже было отмечено в § 8.6, энергия электрона, отвечающая финитному движению в плоскости, перпендикулярной магнитному полю  $\mathbf{H}$ , квантуется:

$$\varepsilon_n = \left( n + \frac{1}{2} \right) \hbar \omega_H + \frac{\hbar^2 k_z^2}{2m}. \quad (9.37)$$

Вдоль магнитного поля (ось  $z$ ) электрон движется равномерно и прямолинейно. Спектр разрешенных значений энергии электрона в магнитном поле (при  $k_z = 0$ ) представляет собой совокупность уровней, сдвинутых друг относительно друга по оси энергии на расстояние, равное  $\hbar \omega_H$  (рис. 9.8). Согласно квантовой механике, процессу поглощения света электронами в магнитном поле соответствуют оптические переходы между уровнями Ландау (рис. 9.8, стрелка между уровнями  $n = 1$  и  $n = 2$ ).

В § 8.6 детально рассмотрена квантовая теория оптических переходов между уровнями Ландау для межзонных переходов. Аналогичным образом можно вычислить коэффициент поглощения света для оптических внутризонных переходов между уровнями Ландау.

Спектр электронов в магнитном поле (при  $k_z = 0$ ) имеет точно такой же вид, как и для линейного гармонического осциллятора. Для гармонического осциллятора существует правило отбора для пере-

<sup>1)</sup> *Luttinger T. M. Quantum Theory of Cyclotron Resonance in Semiconductors: General Theory // Phys. Rev. 1955. V. 102. P. 1030.*



ходов между уровнями: для эквидистантных уровней энергии переходы разрешены только между соседними (ближайшими) уровнями (как показано на рис. 9.8)<sup>1)</sup>. Из рис. 9.8 видно, что квант света  $\hbar\omega$

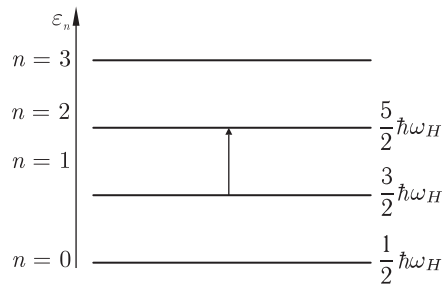


Рис. 9.8. Схематическое изображение эквидистантных уровней Ландау электрона во внешнем магнитном поле для случая  $k_z = 0$

может поглощаться электронами только тогда, когда  $\omega = \omega_H$ . Итак, циклотронный резонанс, с точки зрения квантовой механики, — это переход электрона между уровнями Ландау, при этом частота кванта света равна циклотронной частоте электрона. Важно отметить, что циклотронные переходы на одной частоте возможны только тогда, когда уровни энергии электрона эквидистантны.

Если расстояния (по энергии) между уровнями электрона не совсем одинаковы (то есть нет эквидистантности), то возможны оптические переходы на частотах  $\omega = 2\omega_H$  и  $3\omega_H$ , то есть появляются циклотронные гармоники. Какие явления могут привести к нарушению эквидистантности между уровнями энергии электрона в магнитном поле? Причин может быть несколько<sup>2)</sup>, но главной из них в полупроводниках является непараболичность спектра. Выше, рассматривая спектр электронов в магнитном поле (см. (9.37)), мы считали, что он строго параболичен. Если спектр электронов является непараболическим (см. (4.105)), эффективная масса начинает зависеть от энергии. Для непараболического спектра уровни Ландау электрона уже не будут эквидистантными; при этом в спектре циклотронного резонанса появляются циклотронные гармоники, то есть частоты  $2\omega_H$ ,  $3\omega_H$ . Кроме того, появляется и некоторая ширина линии циклотронного резонанса.

Возникает вопрос, могут ли уровни Ландау для свободного электрона быть неэквидистантными. Оказывается, что это возможно, если учитываются релятивистские эффекты. Как уже было отмечено в главе 4 (см. формулы (4.104) и (4.105)), релятивистское выражение для энергии электрона в теории относительности аналогично формуле для энергии электрона в полупроводнике, полученной в приближении слабой связи. Такой спектр можно интерпретировать, считая, что масса электрона зависит от его скорости. Для нерелятивистских скоростей эффект непараболичности мал. Для релятивистских скоростей эффект

<sup>1)</sup> См. Ландау Л. Д., Лифшиц Е. М. Квантовая механика. — М.: Наука, 1974.

<sup>2)</sup> Например, спин-орбитальное взаимодействие.

непараболичности спектра становится существенным и находит важное применение в физике. Благодаря неэквидистантности уровней Ландау, для релятивистского электрона был реализован мазер на циклотронном резонансе<sup>1)</sup>.

Кратко поясним принцип работы такого лазера. Рассмотрим систему уровней энергии электрона в магнитном поле (рис. 9.9). Для того, чтобы в такой системе возникли индуцированные переходы, необходимо на каком-то уровне создать инверсную заселенность. Создать для свободного электрона инверсную заселенность, например на уровне  $E_3$ , можно и без учета релятивистских эффектов, но при этом не будет генерации электромагнитного излучения. Дело в том, что если на уровне  $E_3$  создана инверсная заселенность, то электрон будет совершать вынужденный переход с уровня  $E_3$  на уровень  $E_2$ . Однако возможен также переход электрона с уровня  $E_3$  на уровень  $E_4$ . Такому переходу соответствует поглощение фотона. Теоретический расчет показывает, что при переходе с уровня  $E_3$  коэффициент поглощения больше, чем коэффициент усиления, поэтому генерация излучения невозможна. Но при учете релятивистских эффектов ситуация кардинально меняется. Расстояние между уровнями  $E_2$  и  $E_3$  будет немного большим, чем между уровнями  $E_3$  и  $E_4$  (поскольку на уровне  $E_4$  скорость электрона больше, чем на уровне  $E_2$ , а эффективная масса растет с ростом скорости электрона). Тогда свет, который индуцирует переходы с уровня  $E_3$  на уровень  $E_2$  (его частота  $\omega_{32}$ ), уже не поглощается между уровнями  $E_3$  и  $E_4$ , так как  $\omega_{34} \neq \omega_{32}$ . В результате коэффициент усиления становится больше коэффициента внутренних потерь излучения, а это есть необходимое условие для возникновения генерации электромагнитного излучения. На этом принципе и был создан в 1959 г. мазер на циклотронном резонансе (А. В. Гапоновым-Греховым).

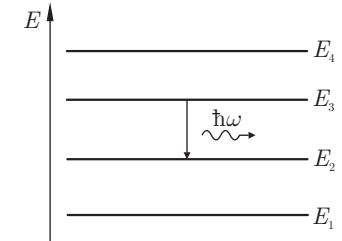


Рис. 9.9. Схематическое изображение неэквидистантных уровней Ландау электрона во внешнем магнитном поле. Отмечен вынужденный переход электрона с уровня  $E_3$  на уровень  $E_2$

## 9.2. Эффект Холла и магнетосопротивление

**9.2.1. Классический эффект Холла.** Если полупроводник, по которому течет ток, поместить в однородное магнитное поле  $\mathbf{H}$ , составляющее прямой угол с направлением электрического тока, то в полупроводнике возникает электрическое поле, направленное перпендикулярно

<sup>1)</sup> Мазер — это квантовый усилитель СВЧ-диапазона (см. § 8.5).



плоскости, содержащей векторы электрического тока и напряженности магнитного поля (рис. 9.10). Этот эффект называется эффектом Холла по имени автора, открывшего его в 1879 г. Эффект Холла стал одним из самых эффективных методов исследования физических свойств носителей заряда в полупроводниках.

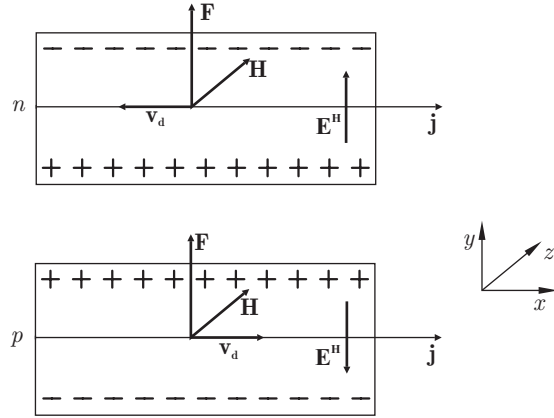


Рис. 9.10. Схема возникновения поля Холла  $\mathbf{E}^H$  в электронном и дырочном полупроводниках. Электроны, движущиеся со скоростью  $\mathbf{v}_d$  против направления  $Ox$  под действием магнитного поля  $H_z$ , наложенного параллельно оси  $Oz$ , отклоняются к верхней поверхности образца, которая в результате приобретает заряд, создающий электрическое поле  $\mathbf{E}^H$ . Для дырочного полупроводника — все аналогично

Итак, возьмем полупроводник в виде параллелепипеда (шириной  $d$ ), по которому в направлении оси  $x$  течет электрический ток с плотностью  $\mathbf{j}$  и который находится в магнитном поле, направленном по оси  $z$  (рис. 9.10). Носители заряда в электрическом поле приобретают скорость направленного движения  $\mathbf{v}_d$  — дрейфовую скорость, которая направлена по полю для дырок и против поля — для электронов. При включении магнитного поля на электроны и дырки действует сила Лоренца

$$\mathbf{F} = \frac{e}{c}[\mathbf{v}_d \mathbf{H}], \quad (9.38)$$

которая перпендикулярна  $\mathbf{v}_d$  и  $\mathbf{H}$ . С другой стороны,

$$\mathbf{v}_d = \mu_d \mathbf{E} = \frac{e\tau}{m} \mathbf{E}, \quad (9.39)$$

где  $\tau$  — время релаксации. В результате

$$\mathbf{F} = \frac{e^2 \tau}{m} [\mathbf{E} \mathbf{H}]. \quad (9.40)$$

Мы получили, что сила Лоренца не зависит от знака носителей заряда, а определяется только направлением полей  $\mathbf{E}$  и  $\mathbf{H}$ , или ориен-

тацией тока  $\mathbf{j}$  и магнитного поля  $\mathbf{H}$ . Для случая, представленного на рис. 9.10, сила  $\mathbf{F}$  направлена по оси  $y$ . Итак, носители заряда — электроны и дырки — отклоняются (в скрещенных магнитных и электрических полях) в одну и ту же сторону. Если полупроводник с носителями заряда одного знака имеет конечные размеры в направлении оси  $y$ , то в результате действия полей  $\mathbf{E}$  и  $\mathbf{H}$  произойдет накопление носителей заряда на верхней (см. рис. 9.10) стороне образца и возникнет их дефицит на нижней. Противоположные стороны образца заряжаются, и возникает поперечное по отношению к  $\mathbf{E}$  электрическое поле  $\mathbf{E}^H$ . Это поле носит название поля Холла, а явление возникновения поперечного поля под действием магнитного поля называют эффектом Холла.

Направление поля Холла  $\mathbf{E}^H$  зависит от знака носителей заряда. В нашем случае  $\mathbf{E}^H$  направлено вверх (по оси  $y$ ) в  $n$ -образце и вниз (по  $y$ ) в  $p$ -образце (рис. 9.10). Величина поля  $\mathbf{E}^H$  будет расти до тех пор, пока оно не скомпенсирует силу Лоренца (9.38). После этого носители заряда будут двигаться только под действием одного поля  $\mathbf{E}$ , и траектория носителей заряда будет представлять собой снова прямую линию, направленную вдоль оси  $x$ , при этом вектор  $\mathbf{j}$  будет направлен по полю  $\mathbf{E}$ . Суммарное электрическое поле, действующее на электрон,  $\mathbf{E}' = \mathbf{E} + \mathbf{E}^H$  будет повернуто на некоторый угол  $\varphi$  относительно оси  $x$  или  $\mathbf{j}$  (см. рис. 9.11, а, б). Таким образом, в ограниченном полупроводнике поворачивается вектор электрического поля  $\mathbf{E}'$ , и между током  $\mathbf{j}$  и полем  $\mathbf{E}'$  возникает угол  $\varphi$ , называемый углом Холла.

Итак, в направлении, перпендикулярном току, появляется разность потенциалов  $V^H = dE^H$ , где  $d$  — размер образца в направлении, перпендикулярном  $\mathbf{E}$  и  $\mathbf{H}$ , величина  $V^H$  носит название холловской разности потенциалов. Холл экспериментально нашел, что поле  $\mathbf{E}^H$  определяется плотностью тока  $\mathbf{j}$  и напряженностью магнитного поля  $\mathbf{H}$ , а также свойствами образца. Свойства образца описываются некоторой величиной  $R$ , называемой коэффициентом Холла. Таким образом, четыре величины:  $\mathbf{E}^H$ ,  $\mathbf{j}$ ,  $\mathbf{H}$  и  $R$  — связаны эмпирическим соотношением

$$\mathbf{E}^H = R[\mathbf{H} \mathbf{j}]. \quad (9.41)$$

Коэффициент Холла  $R$  можно найти, если учесть, что холловское поле  $\mathbf{E}^H$  должно компенсировать силу Лоренца:

$$e\mathbf{E}^H + \mathbf{F} = 0. \quad (9.42)$$

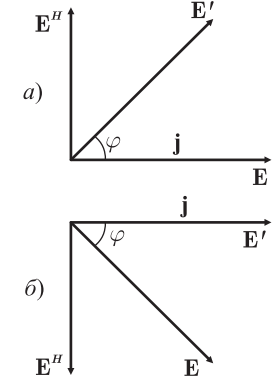


Рис. 9.11. Угол Холла между током  $\mathbf{j}$  и полем  $\mathbf{E}'$  для полупроводника  $n$ -типа (а) и полупроводника  $p$ -типа (б)

Из (9.42) с учетом (9.38) и (9.39) находим

$$\mathbf{E}^H = -\frac{1}{c}[\mathbf{v}_d \mathbf{H}] = -\frac{\mu_d}{c}[\mathbf{E} \mathbf{H}]. \quad (9.43)$$

Принимая во внимание, что плотность тока равна

$$\mathbf{j} = \sigma \mathbf{E}, \quad (9.44)$$

где  $\sigma = \frac{ne^2}{m}\tau$  — проводимость, (9.41) можно переписать в виде

$$\mathbf{E}^H = -R[\mathbf{j} \mathbf{H}] = -R\sigma[\mathbf{E} \mathbf{H}]. \quad (9.45)$$

Сравнивая (9.43) и (9.45), мы видим, что

$$R = \frac{\mu_d}{\sigma c} = \frac{1}{nec}. \quad (9.46)$$

Здесь через  $n$  обозначена концентрация носителей заряда (электронов или дырок). Мы видим, что коэффициент Холла  $R$  обратно пропорционален концентрации носителей заряда и его знак совпадает со знаком носителей заряда. (Для электронов  $R = -\frac{1}{nec}$ , для дырок  $R = \frac{1}{p ec}$ , где  $e > 0$ ).

Как будет видно из дальнейшего, формула (9.46) справедлива, строго говоря, лишь тогда, когда время релаксации  $\tau$  не зависит от энергии частицы. В более общем случае необходимо включить в (9.46) еще и численный множитель  $A$  порядка единицы, который определяется преобладающим механизмом рассеяния и, кроме того, зависит от степени вырождения электронов в зоне проводимости или дырок в валентной зоне. Таким образом, коэффициент Холла  $R$  следует записывать в виде

$$R = \frac{A}{nec}. \quad (9.47)$$

Множитель  $A$  называется холловским фактором. Его значение зависит от механизмов рассеяния носителей заряда, кроме того, величина  $A$  для электронов и дырок может быть разной.

С помощью равенства

$$|R|\sigma \cdot c = \mu_H \quad (9.48)$$

можно дать определение величины  $\mu_H$ , имеющей размерность подвижности. Ее обычно называют холловской подвижностью. Подвижность, определяемую из электропроводности, то есть дрейфовую подвижность, мы будем обозначать через  $\mu_d$ . Очевидно, что если время релаксации  $\tau = \text{const}$ , то  $\mu_H = \mu_d$ ; в противном случае  $\mu_H = A \cdot \mu_d$ .

Приведем более строгий вывод выражения для коэффициента Холла  $R$ .

Рассмотрим полупроводник в форме параллелепипеда (рис. 9.12). Пусть магнитное поле направлено по оси  $z$ ; а электрическое поле имеет две компоненты,  $E_x$  и  $E_y$ . Для нахождения проводимости воспользу-

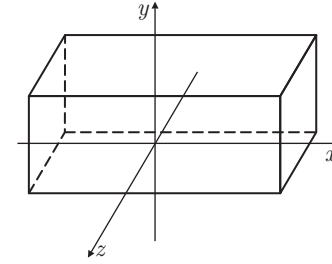


Рис. 9.12. К выводу выражения для коэффициента Холла (9.59): полупроводник конечных размеров (в форме параллелепипеда) во внешнем магнитном  $\mathbf{H}$  (вдоль  $z$ ) и электрическом  $\mathbf{E}$  (в плоскости  $x, y$ ) полях

емся уравнением Лоренца–Друде. Поскольку нас интересует стационарное состояние, то уравнение Лоренца–Друде в этом случае имеет вид

$$\frac{e}{m}\mathbf{E} + \frac{e}{mc}[\mathbf{v} \mathbf{H}] - \frac{\mathbf{v}}{\tau} = 0. \quad (9.49)$$

Вектор скорости имеет две компоненты:  $v_x$  и  $v_y$ . Тогда уравнение (9.49) можно переписать в виде

$$\begin{aligned} \frac{e}{m}E_x + \omega_H v_y - \frac{v_x}{\tau} &= 0, \\ \frac{e}{m}E_y - \omega_H v_x - \frac{v_y}{\tau} &= 0. \end{aligned} \quad (9.50)$$

Решение системы уравнений (9.50) можно представить как

$$\begin{aligned} v_x &= \frac{e\tau}{m} \frac{1}{1 + \tau^2 \omega_H^2} [E_x + \tau \omega_H E_y], \\ v_y &= \frac{e\tau}{m} \frac{1}{1 + \tau^2 \omega_H^2} [E_y - \tau \omega_H E_x]. \end{aligned} \quad (9.51)$$

Компоненты плотности тока принимают вид

$$\begin{aligned} j_x &= nev_x = \sigma_{xx} E_x + \sigma_{xy} E_y, \\ j_y &= \sigma_{yy} E_y + \sigma_{yx} E_x. \end{aligned} \quad (9.52)$$

Здесь

$$\sigma_{xx} = \frac{e^2 n}{m} \left\langle \frac{\tau}{1 + \tau^2 \omega_H^2} \right\rangle = \sigma_{yy}, \quad (9.53)$$

$$\sigma_{xy} = -\sigma_{yx} = \frac{e^2 n}{m} \left\langle \frac{\tau^2 \omega_H}{1 + \tau^2 \omega_H^2} \right\rangle. \quad (9.54)$$

Следует отметить, что диагональные компоненты тензора проводимости  $\sigma_{xx}(\sigma_{yy})$  принято называть омической проводимостью, а недиагональные компоненты  $\sigma_{xy}$  — холловской проводимостью.

**Замечание.** При выводе (9.52) мы исходили из уравнения Лоренца–Друде (9.49), которое было написано для одного электрона.

Время релаксации импульса  $\tau$ , которое входит в (9.49), зависит от энергии электронов, совершающих беспорядочное движение, поэтому мы должны в выражениях для проводимости усреднить  $\tau$  по энергии. Угловые скобки в (9.53) и (9.54) обозначают усреднение проводимости по энергии с функцией распределения носителей по энергии (это может быть, например, функция распределения Больцмана).

Вернемся к выражениям (9.52) для плотности тока. В стационарных условиях ток вдоль оси  $y$  ( $j_y$ ) равен нулю. (Напомним, что ось  $x$  — это направление приложенного поля, а  $y$  — направление поля Холла.) Таким образом, вдоль оси  $y$  есть только холловское поле, а  $j_y = 0$ . Тогда из второго уравнения (9.52) получаем, что

$$E_y = -\frac{\sigma_{yx}}{\sigma_{yy}} E_x. \quad (9.55)$$

В результате, плотность тока вдоль оси  $x$  с учетом (9.55) принимает вид

$$j_x = \left( \sigma_{xx} - \frac{\sigma_{xy}\sigma_{yx}}{\sigma_{yy}} \right) E_x. \quad (9.56)$$

Тогда холловское поле  $E_y$  можно выразить через плотность тока  $j_x$  согласно (9.56):

$$E_y = -\frac{\sigma_{yx}}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}\sigma_{yx}} j_x. \quad (9.57)$$

Мы получили общую формулу для холловского поля  $E_y$ . Коэффициент пропорциональности между полем и плотностью тока  $j_x$  является достаточно громоздким; его можно упростить для ряда частных случаев.

**А. Сильные магнитные поля:**  $\omega_H \tau \gg 1$ . Это случай, когда электрон много раз вращается по ларморовской орбите, прежде чем он столкнется с рассеивателем. Пренебрегая единицей по сравнению с  $(\omega_H \tau)^2$ , получим из (9.53) и (9.54)

$$\sigma_{xx} = \frac{e^2 n}{m \omega_H^2} \left\langle \frac{1}{\tau} \right\rangle, \quad \sigma_{xy} = -\sigma_{yx} = \frac{e^2 n}{m \omega_H}. \quad (9.58)$$

При этом  $\frac{\sigma_{xx}}{\sigma_{xy}} = \frac{1}{\omega_H} \left\langle \frac{1}{\tau} \right\rangle \ll 1$ , так что  $\sigma_{xx}\sigma_{yy} - \sigma_{xy}\sigma_{yx} \equiv \sigma_{xx}^2 + \sigma_{xy}^2 \approx \sigma_{xy}^2$ . Из (9.57) с учетом (9.58) и (9.41) получим

$$E_y = \frac{H}{nec} j_x \equiv RH j_x, \quad (9.59)$$

где  $R = \frac{1}{nec}$  — коэффициент Холла. Мы получили прежнее выражение для коэффициента Холла.

Отметим интересную особенность транспорта в сильных магнитных полях ( $\omega_H \tau \gg 1$ ). Из (9.58) видно, что  $\sigma_{xx}$  пропорциональна частоте столкновений  $\tau^{-1}$ , то есть в отсутствии рассеяния носителей ток вдоль поля не возникает. Этот результат легко понять, поскольку без рассеяния орбиты электронов замкнуты.

**В. Слабые магнитные поля:**  $\omega_H \tau \ll 1$ . В линейном приближении по магнитному полю, то есть когда мы сохраняем в разложении в ряд по  $H$  члены не выше первой степени, получим (см. (9.53) и (9.54))

$$\sigma_{xx} = \frac{e^2 n}{m} \langle \tau \rangle; \quad \sigma_{xy} = \frac{e^2 n}{m} \omega_H \langle \tau^2 \rangle. \quad (9.60)$$

В слабом магнитном поле недиагональные компоненты проводимости гораздо меньше диагональных ( $\sigma_{xy} \ll \sigma_{xx}$ ). Поэтому в (9.57) мы можем пренебречь недиагональными компонентами по сравнению с диагональными. Тогда для поля Холла имеем

$$E_y = \frac{1}{nec} \frac{\langle \tau^2 \rangle}{\langle \tau \rangle} H j_x \equiv RH j_x. \quad (9.61)$$

В результате коэффициент Холла в слабых магнитных полях оказывается равным

$$R = \frac{1}{nec} \frac{\langle \tau^2 \rangle}{\langle \tau \rangle} \equiv \frac{A}{nec}. \quad (9.62)$$

Следовательно, холловский фактор равен

$$A = \frac{\langle \tau^2 \rangle}{\langle \tau \rangle e^2}. \quad (9.63)$$

Из этих выражений непосредственно видно, что если  $\tau$  от энергии не зависит, то  $R$  соответствует выражению (9.46). Таким образом, величина холловского фактора зависит от механизма рассеяния.

**Замечание 1.** Исследования эффекта Холла являются важным методом определения типа проводимости. С другой стороны, зная коэффициент Холла  $R$ , можно определить концентрацию носителей:

$$n = \frac{A}{R ec}. \quad (9.64)$$

**Замечание 2.** На практике часто удобнее измерять не проводимость, а сопротивление. Из системы уравнений (9.52) можно выразить напряженность электрического поля через плотность тока:

$$E_x = \rho_{xx} j_x + \rho_{xy} j_y, \quad E_y = \rho_{yx} j_x + \rho_{yy} j_y. \quad (9.65)$$

Здесь

$$\rho_{xx} = \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{xy}^2} = \rho_{yy}, \quad (9.66)$$

$$\rho_{xy} = -\rho_{yx} = \frac{\sigma_{yx}}{\sigma_{xx}^2 + \sigma_{xy}^2}. \quad (9.67)$$

Равенство  $\rho_{xx} = \rho_{yy}$ , а также (9.53) следуют из кубической симметрии кристалла (например, Si, Ge, GaAs...). Равенство  $\rho_{xy} = -\rho_{yx}$  и (9.54)

следуют из соотношений Онсагера<sup>1)</sup>. Знак холловских компонент  $\sigma_{xy}$  и  $\rho_{xy}$  определяется знаком носителей заряда. Выражения (9.65) и (9.52) можно записать в более компактном виде:

$$\mathbf{j} = \hat{\sigma} \mathbf{E}, \quad \mathbf{E} = \hat{\rho} \mathbf{j}. \quad (9.68)$$

Здесь  $\hat{\sigma}$  и  $\hat{\rho}$  — тензоры удельной проводимости и сопротивления, удовлетворяющие равенству:

$$\hat{\sigma} \hat{\rho} = \hat{1}, \quad (9.69)$$

где  $\hat{1}$  — единичная матрица.

**9.2.2. Квантовый эффект Холла.** Квантовый эффект Холла — одно из наиболее неординарных явлений физики полупроводников, открытое в 1980 г. немецким физиком Клаусом фон Клитцингом<sup>2)</sup>. Уникальность эффекта связана с тем, что измеряемая величина — холловское сопротивление двумерного полупроводникового образца  $\rho_{xy}$  — принимает фиксированные, *квантованные* значения с точностью, превышающей  $10^{-8}$ . Квантовый эффект Холла наблюдается в сильных магнитных полях  $H \gtrsim 1$  Тл<sup>3)</sup> при низких температурах  $T \lesssim 1$  К. Экспериментальными образцами могут служить различные двумерные полупроводниковые гетероструктуры, например, квантовые ямы GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As или In<sub>x</sub>Ga<sub>1-x</sub>As/GaAs<sup>4)</sup>, структуры металл-диэлектрик-полупроводник. В последнее время также активно изучается квантовый эффект Холла в двумерных слоях графита (графене)<sup>5)</sup>. Принципиальной особенностью двумерных структур является наличие уровней размерного квантования, отделенных достаточно большими энергетическими зазорами по сравнению с температурой, при которых проводится эксперимент<sup>6)</sup>. Обычно все носители заряда в таких структурах находятся на нижнем уровне размерного квантования.

<sup>1)</sup> См. Ландау Л. Д., Лифшиц Е. М. Электродинамика сплошных сред. — М., 1982. 623 с. § 21.

<sup>2)</sup> Подробное описание целочисленного и дробного квантового эффекта Холла можно найти в книге: Квантовый эффект Холла / Под ред. Р. Пренджа и С. Гирвина. — М.: Мир, 1989.

<sup>3)</sup> 1 Тесла =  $10^4$  эрстед.

<sup>4)</sup> В обоих случаях квантовая яма образуется в результате формирования очень узких ( $L \sim 100$  Å), практически бездефектных слоев полупроводникового материала с меньшей шириной запрещенной зоны между слоями с большей шириной запрещенной зоны. В первом случае материалом квантовой ямы является полупроводник GaAs, во втором — твердый раствор In<sub>x</sub>Ga<sub>1-x</sub>As.

<sup>5)</sup> В графене квантовый эффект Холла может наблюдаться при комнатной температуре, см. Novoselov K. S. *et al.* Two-Dimensional Gas of Massless Dirac Fermions in Graphene // Nature. 2005. V. 438. P. 197.

<sup>6)</sup> Типичное расстояние между уровнями размерного квантования составляет десятки мэВ, температуре 1 К соответствует энергия  $k_B T = 0,1$  мэВ.

Прежде всего рассмотрим, к чему приводит решение задачи о вычислении тензора проводимости двумерного полупроводникового образца при низких температурах в рамках классической механики. Формулы (9.53)–(9.54) для компонент  $\sigma_{xx}$  и  $\sigma_{xy}$ , полученные в § 9.2.1, остаются в силе и в двумерном случае, если в них заменить трехмерную концентрацию  $n$  на двумерную (поверхностную)  $n_s$ :

$$\sigma_{xx} = \frac{n_s e^2 \tau}{m} \frac{1}{1 + (\omega_H \tau)^2}, \quad (9.70)$$

$$\sigma_{xy} = \frac{n_s e c}{H} \frac{(\omega_H \tau)^2}{1 + (\omega_H \tau)^2}. \quad (9.71)$$

Здесь  $\omega_H = eH/(mc)$  — циклотронная частота;  $\tau$  — время релаксации импульса на уровне Ферми<sup>1)</sup>. Используя формулы (9.57), (9.70) и (9.71), несложно получить выражение для холловского сопротивления

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2} = \frac{H}{n_s e c}. \quad (9.72)$$

Таким образом, согласно классической кинетической теории Друде–Больцмана холловское сопротивление двумерного образца прямо пропорционально магнитному полю  $H$ , обратно пропорционально концентрации носителей  $n_s$  и не зависит от времени релаксации импульса  $\tau$ .

Такое поведение холловского сопротивления действительно имеет место в относительно слабых магнитных полях, однако в сильных полях картина кардинально меняется (рис. 9.13). С увеличением магнитного поля линейная зависимость  $\rho_{xy}(B)$ , описываемая формулой (9.72), переходит в последовательность горизонтальных отрезков, называемых *плато*, и практически вертикальных переходных областей между ними. Одновременно в областях плато продольное сопротивление обращается в нуль  $\rho_{xx} = 0$ , а в переходных областях между соседними плато имеется последовательность пиков (максимумов)  $\rho_{xx}$ .

Однако наиболее примечательным является даже не сам факт наличия ступенек на зависимости  $\rho_{xy}(B)$ , а те значения, которые может принимать холловская проводимость в областях плато. Как показали измерения фон Клитцинга, эти значения являются квантованными, а именно

$$\rho_{xy} = \frac{1}{N} \frac{h}{e^2}, \quad (9.73)$$

<sup>1)</sup> В формулах (9.70)–(9.71) мы использовали тот факт, что при низких температурах двумерный электронный газ является вырожденным, и в этом пределе проводимость определяется временем релаксации импульса именно на уровне Ферми. Кроме того, при их выводе мы неявно подразумевали, что влияние имеющихся в образце примесей и дефектов на движение носителей не приводит к их локализации.

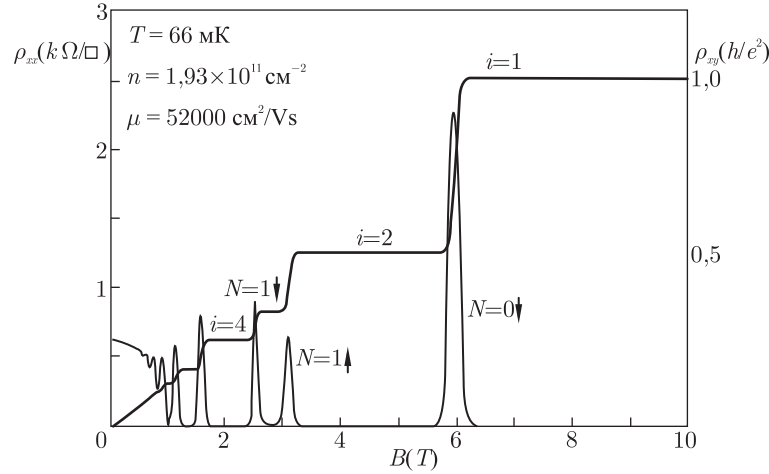


Рис. 9.13. Холловское сопротивление  $\rho_{xy}$  и продольное сопротивление  $\rho_{xx}$  в сильных магнитных полях, из работы Wei H.P. et al. // Phys. Rev. B. 2003. V. 45. P. 3926

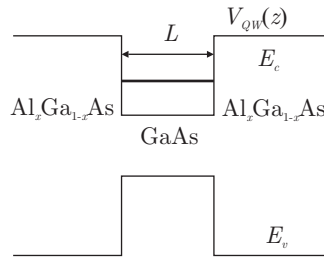


Рис. 9.14. Зонная диаграмма квантовой ямы GaAs/Al<sub>x</sub>Ga<sub>1-x</sub>As. Показан единственный уровень размерного квантования в зоне проводимости, на котором при низких температурах находятся все электроны (образец *n*-типа)

где  $N$  является с огромной точностью целым числом. В настоящее время надежно установлено, что отличие  $N$  от целого не превышает  $10^{-8}$  и ограничено лишь современными возможностями прецизионных измерений.

Объяснение столь необычного поведения компонент тензора проводимости в сильных магнитных полях требует последовательного решения квантово-механической задачи о движении носителей в магнитном поле в присутствии *хаотического потенциала*, образованного случайно расположенными в плоскости образца примесями и дефектами. Квантово-механическая динамика носителей в магнитном поле <sup>1)</sup>  $\mathbf{H}$

<sup>1)</sup> Здесь и ниже мы будем рассматривать магнитное поле, направленное строго перпендикулярно плоскости квантовой ямы,  $\mathbf{H} = (0, 0, H)$ .

описывается уравнением Шредингера

$$\left[ \frac{(-i\hbar\nabla - e/c\mathbf{A})^2}{2m} + V_{QW}(z) + V_{\text{imp}}(\mathbf{r}) \right] \Psi(\mathbf{r}) = E\Psi(\mathbf{r}). \quad (9.74)$$

Здесь  $\mathbf{A}(\mathbf{r})$  — векторный потенциал магнитного поля;  $\text{rot } \mathbf{A} = \mathbf{H}$ ,  $V_{QW}(z)$  — потенциал квантовой ямы, ограничивающей движение электронов в направлении оси  $z$ , см. рис. 9.14, а  $V_{\text{imp}}(\mathbf{r}) = \sum U_i(\mathbf{r} - \mathbf{r}_i)$  — суммарный хаотический потенциал различных несовершенств кристаллической решетки (примесей, дефектов, флуктуаций состава твердого раствора), расположенных в случайных точках  $\mathbf{r}_i$ . Обычно зависимость хаотического потенциала от координаты  $z$  пренебрегают, рассматривая  $V_{\text{imp}}(x, y)$ , см. рис. 9.15. При этом в зависимости от конкретного образца может реализовываться как случай плавного хаотического потенциала,  $\lambda \gg a_H$ , так и случай резкого хаотического потенциала,  $\lambda \ll a_H$ . Здесь  $\lambda$  — характерный масштаб изменения хаотического потенциала,  $a_H = \sqrt{\hbar c / (eH)}$  — магнитная длина, равная приблизительно 100 Å в магнитном поле  $B = 6,5$  Тл.

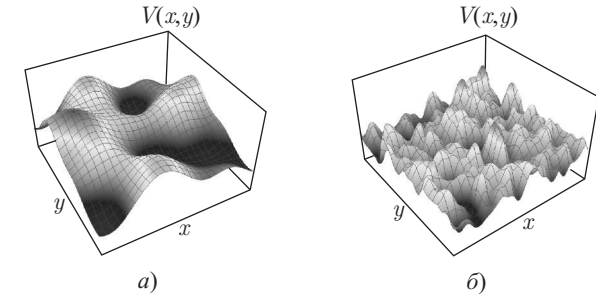


Рис. 9.15. Вид хаотического потенциала примесей дефектов  $V_{\text{imp}}(x, y)$ : а) плавный хаотический потенциал,  $\lambda \gg a_H$ ; б) резкий хаотический потенциал,  $\lambda \ll a_H$

В отсутствие хаотического потенциала,  $V_{\text{imp}} = 0$ , уравнение (9.74) решается точно, при этом спектр электронных состояний имеет вид лестницы уровней Ландау

$$E_n = \hbar\omega_H(n + 1/2), \quad (9.75)$$

где энергия  $E$  отсчитывается от уровня размерного квантования носителей в квантовой яме,  $n \geq 0$  — номер уровня Ландау (ср. с (9.37) в §9.1.2). Специфика двумерного случая состоит в том, что кинетическая энергия движения вдоль оси  $z$  заменяется на энергию размерного квантования, поэтому энергетический спектр носителей является полностью дискретным. Это приводит к дельта-функциональному виду

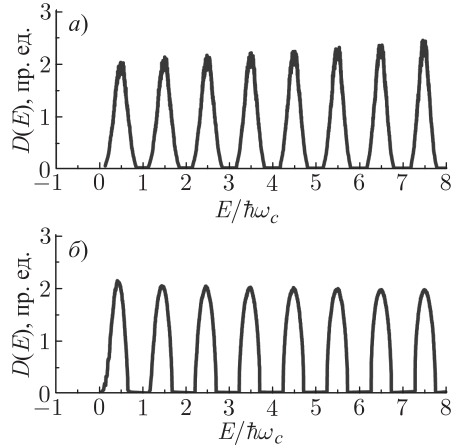


Рис. 9.16. Плотность состояний в магнитном поле: а) плавный хаотический потенциал,  $\lambda \gg a_H$ ; б) резкий хаотический потенциал,  $\lambda \ll a_H$

плотности состояний <sup>1)</sup>,

$$D(E) = n_L \sum_n \delta(E - \hbar\omega_H(n + 1/2)). \quad (9.76)$$

Здесь  $n_L = (2\pi a_H^2)^{-1}$  — число состояний, приходящихся на один уровень Ландау в расчете на единицу площади образца.

Формула (9.76) показывает, что стандартный подход к расчету проводимости, учитывающий хаотический потенциал  $V_{\text{imp}}$  в рамках теории возмущений, не применим для расчета проводимости в режиме квантового эффекта Холла, так как электронные состояния на каждом уровне Ландау являются многократно вырожденными. Если же вообще не учитывать хаотический потенциал  $V_{\text{imp}}$ , то мы получим формулу (9.71) при  $\tau = \infty$ , то есть вернемся к результату классической кинетической теории. Таким образом, влияние несовершенств кристаллической решетки на квантово-механическое движение электрона в магнитном поле необходимо учитывать точно, и это является основной сложностью при решении соответствующей задачи. Преодоление этой сложности возможно с использованием численных расчетов, и результаты этих расчетов заключаются в следующем. Прежде всего, благодаря наличию хаотического потенциала  $V_{\text{imp}}$  дельта-функциональный вид плотности состояний (9.76) размывается <sup>2)</sup>, см. рис. 9.16. Большинство электронных

<sup>1)</sup> Напомним, что в отсутствие магнитного поля плотность состояний в двумерном случае является константой.

<sup>2)</sup> Для простоты здесь и ниже мы рассматриваем случай *разделенных* уровней Ландау, соответствующий условию  $V \ll \hbar\omega_H$ , где  $V = \sqrt{\langle V_{\text{imp}}^2(\mathbf{r}) \rangle}$  — амплитуда хаотического потенциала.

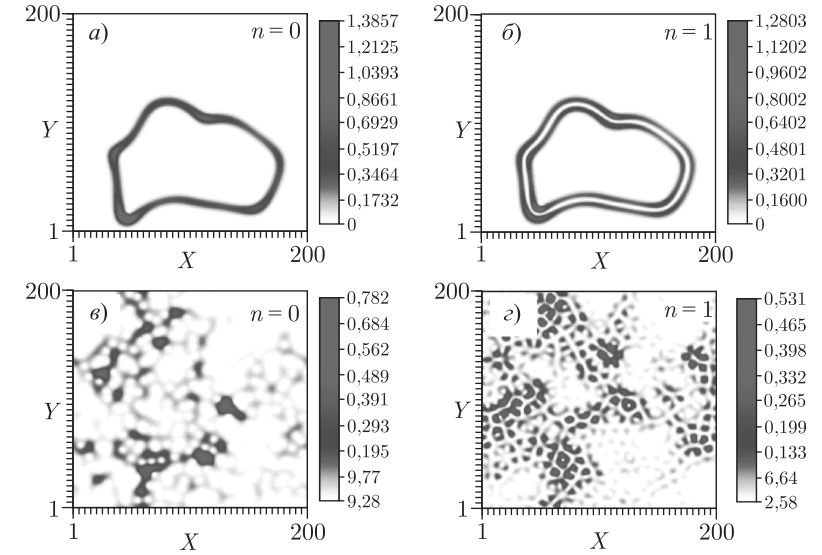


Рис. 9.17. Примеры собственных волновых функций локализованных и делокализованных состояний, принадлежащих уширенным уровням Ландау с номером  $n$ . а и б) плавный хаотический потенциал,  $\lambda \gg a_H$ ; в и г) резкий хаотический потенциал,  $\lambda \ll a_H$ . Амплитуда  $\psi^2(\mathbf{r})$  изображена при помощи разной интенсивности

состояний, принадлежащих уширенным (размытым) уровням Ландау, является локализованным, при этом характерная длина локализации  $\xi$  степенным образом зависит от энергии, отсчитанной от центра соответствующего уровня Ландау,

$$\xi(E) \propto \left| E - \hbar\omega_c \left( n + \frac{1}{2} \right) \right|^{-\nu}, \quad (9.77)$$

а показатель степени  $\nu \approx 2,3$  считается не зависящим от номера уровня Ландау  $n$  и особенностей хаотического потенциала <sup>1)</sup>. Примеры локализованных и делокализованных электронных состояний, полученные из численных расчетов, представлены на рис. 9.17. Отметим, что локализация электронных состояний в магнитном поле происходит не

<sup>1)</sup> Степенной вид зависимости  $\xi(E)$  свидетельствует об отсутствии характерного масштаба, в качестве которого служит магнитная длина  $a_H$  до тех пор, пока длина локализации  $\xi$  не становится много больше  $a_H$ . Наличие степенной зависимости (9.77) лежит в основе так называемой скейлинговой теории квантового эффекта Холла, см. обзор *Huckestein B. Scaling Theory of the Integer Quantum Hall Effect // Rev. Mod. Phys. 1995. V. 67. P. 357.*



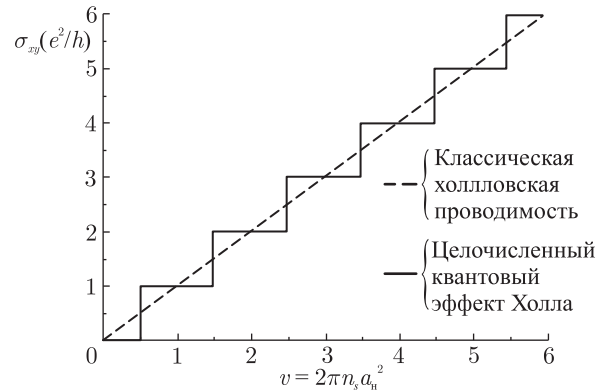


Рис. 9.18. Схематический вид холловской проводимости при нулевой температуре. Прямая линия соответствует классической холловской проводимости

только в потенциальных ямах (минимумах  $V_{\text{имр}}$ ), но и на потенциальных холмах (в максимумах  $V_{\text{имр}}$ )<sup>1)</sup>.

Наличие степенного закона (9.77) позволяет разом объяснить экспериментально наблюдаемое поведение как холловского, так и продольного сопротивления в зависимости от фактора заполнения  $\nu = n_s/n_L$ . При изменении концентрации носителей  $n_s$  уровень Ферми последовательно пересекает зоны локализованных состояний, при этом  $\rho_{xx} = 0$ <sup>2)</sup> и  $d\rho_{xy}/d\nu = 0$ , и делокализованных состояний, при этом  $\rho_{xx} \neq 0$  и  $d\rho_{xy}/d\nu \neq 0$ . В результате зависимости  $\rho_{xx}$  и  $\rho_{xy}$  от фактора заполнения  $\nu$  приобретают вид, показанный схематически на рис. 9.18<sup>3)</sup>. Таким образом, мы показали, что причиной возникновения плато на зависимости  $\rho_{xy}(\nu)$  является совместное действие магнитного поля  $\mathbf{H}$  и хаотического потенциала  $V_{\text{имр}}$ . При этом величины  $\rho_{xy}$  на плато остались неопределенными, в то время как именно эти значения и представляют основную интригу квантового эффекта Холла.

В заключение отметим, что за рамками нашего рассмотрения остался ряд факторов, оказывающих существенное влияние на проводимость в сильных магнитных полях. Наиболее важным из них является кулоновское взаимодействие носителей, приводящее к возможности суще-

<sup>1)</sup> В случае *резкого* хаотического потенциала, т. е. меняющегося на масштабах, меньших магнитной длины  $a_H$ , следует говорить о локализации в *эффективном* хаотическом потенциале, из которого исключена быстро меняющаяся составляющая, соответствующая волновым векторам  $q \gtrsim a_H^{-1}$ .

<sup>2)</sup> Используя соотношение  $\sigma_{xx} = \rho_{xx}/(\rho_{xx}^2 + \rho_{xy}^2)$ , нетрудно убедиться, что это эквивалентно  $\sigma_{xx} = 0$ .

<sup>3)</sup> Поскольку энергетическая ширина зон делокализованных состояний стремится к нулю в пределе бесконечного образца и нулевой температуры, наклон кривых  $\rho_{xy}$  в переходных областях стремится к вертикальному.

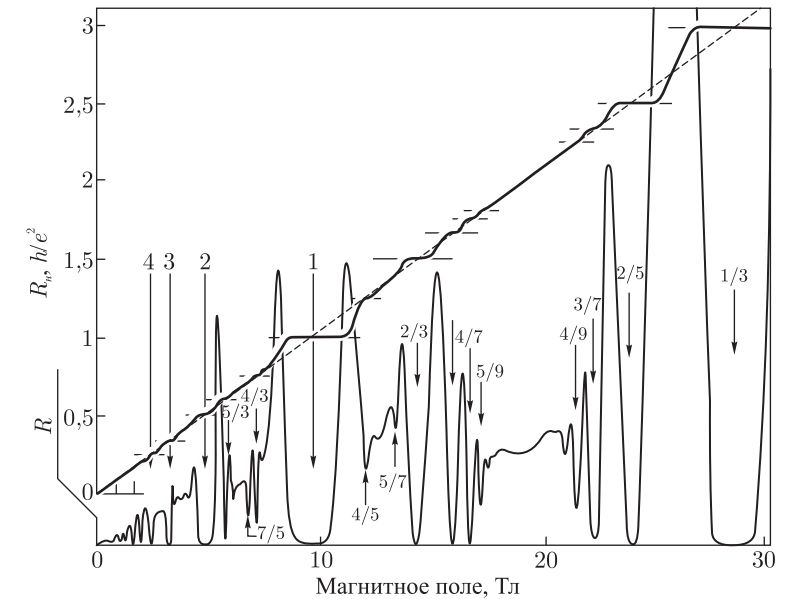


Рис. 9.19. Картина дробного квантового эффекта Холла из работы *Stormer H. L.* // *Physica B.* 1992. V. 177. P. 401

ствования плато с дробными  $N$ , на которых холловское сопротивление принимает значения

$$\rho_{xy} = \frac{p}{q} \frac{h}{e^2}, \quad (9.78)$$

где  $p$  и  $q$  — целые числа. Соответствующее явление, именуемое *дробным* квантовым эффектом Холла в отличие от *целочисленного* квантового эффекта Холла, рассмотренного до сих пор, было открыто в 1982 г. Тсуи, Штормером и Госсардом, см. рис. 9.19. Для наблюдения дробного квантового эффекта Холла необходимы более низкие температуры, чем в целочисленном эффекте, и образцы с более высокой подвижностью носителей. В этом случае электрон-электронное взаимодействие оказывается более сильным, чем хаотический потенциал примесей и дефектов, и именно его следует учитывать в первую очередь.

В последующих двух параграфах мы приводим доказательство квантования холловской проводимости, следуя работе Таулеса и соавторов<sup>1)</sup>. В 9.2.2.1 дан вывод квантово-механического выражения для холловской проводимости, а в 9.2.2.2 с его помощью доказан факт

<sup>1)</sup> См. *Niu Q., Thouless D. J., Wu Y. S.* Quantized Hall Conductance as a Topological Invariant // *Phys. Rev. B.* 1985. V. 31. P. 3372.

квантования. Отметим, что 9.2.2.1 и 9.2.2.2 содержат более сложный материал и могут быть пропущены при первом чтении.

**9.2.2.1. Вывод формулы линейного отклика для холловской проводимости**<sup>1)</sup>. Прежде всего заметим, что расчет тензора проводимости (и других кинетических коэффициентов) требует знания неравновесной функции распределения носителей, а точнее — ее квантово-механического аналога, матрицы плотности. Мы, однако, получим выражение для холловской проводимости более простым путем, используя обычную квантово-механическую теорию возмущений вместо формализма матрицы плотности. Итак, рассмотрим систему невзаимодействующих электронов при нулевой температуре  $T = 0$ . Квантово-механическое движение электронов в магнитном поле подчиняется уравнению (9.74), собственные энергии и волновые функции будем обозначать  $E_i$  и  $\Psi_i$  соответственно. Приложим к системе внешнее электрическое поле в направлении оси  $y$ ,  $\mathbf{E} = (0, \mathcal{E}, 0)$ , при этом в уравнении (9.74) появится дополнительный электростатический потенциал  $e\mathcal{E}y$ , с учетом которого уравнение Шредингера приобретает вид

$$\left[ \frac{(\hat{\mathbf{p}} - e/c\mathbf{A})^2}{2m^*} + V(\mathbf{r}) + e\mathcal{E}y \right] \Psi(\mathbf{r}) = E\Psi(\mathbf{r}). \quad (9.79)$$

Поскольку нас интересует линейный отклик на приложенное электрическое поле,  $j_x = \sigma_{xy}\mathcal{E}$ , учтем последнее слагаемое в уравнении (9.79) в первом порядке теории возмущений, при этом каждая из волновых функций  $\Psi_i$  приобретет добавку, пропорциональную напряженности приложенного электрического поля,

$$\tilde{\Psi}_i = \Psi_i + e\mathcal{E} \sum_{j \neq i} \frac{y_{ji}}{E_i - E_j} \Psi_j = \Psi_i + ie\hbar\mathcal{E} \sum_{j \neq i} \frac{(v_y)_{ji}}{(E_i - E_j)^2} \Psi_j, \quad (9.80)$$

где  $y_{ji}$  и  $(v_y)_{ji}$  — матричные элементы операторов координаты и скорости между собственными функциями уравнения (9.74),

$$y_{ij} = \int \Psi_i^*(\mathbf{r}) y \Psi_j(\mathbf{r}) d^3\mathbf{r}, \quad (9.81)$$

$$(v_y)_{ij} = \frac{\hbar}{m^*} \int \Psi_i^*(\mathbf{r}) (-i\nabla) \Psi_j(\mathbf{r}) d^3\mathbf{r}, \quad (9.82)$$

При получении второго равенства в (9.80) мы использовали связь между матричными элементами операторов координаты и скорости, непосредственно следующую из определения оператора скорости,  $\hat{\mathbf{v}} = i/\hbar[\hat{H}, \mathbf{r}]$ . Суммарный ток в направлении оси  $x$ , переносимый

<sup>1)</sup> Материал данного и следующего параграфов может быть интересен специалистам, интересующимся квантовым эффектом Холла, и при первом чтении может быть пропущен.

электронами, заполняющими при  $T = 0$  состояния с энергиями, меньшими энергии Ферми  $E_F$ , равен

$$J_x = \sum_{E_i < E_F} \langle \tilde{\Psi}_i | \hat{J}_x | \tilde{\Psi}_i \rangle = \frac{ie^2\hbar\mathcal{E}}{L_x} \sum_{E_i < E_F} \sum_{j \neq i} \frac{(v_x)_{ij}(v_y)_{ji} - (v_x)_{ji}(v_y)_{ij}}{(E_i - E_j)^2}. \quad (9.83)$$

Здесь мы использовали факт отсутствия тока в нулевом электрическом поле,  $\sum_i (J_x)_{ii} = 0$ . Поделив полный протекающий ток  $J_x$  на ширину образца  $L_y$ , получим для холловской проводимости

$$\sigma_{xy} = \frac{ie^2\hbar}{S} \sum_{E_i < E_F} \sum_j \frac{(v_x)_{ij}(v_y)_{ji} - (v_x)_{ji}(v_y)_{ij}}{(E_i - E_j)^2}. \quad (9.84)$$

**9.2.2.2. Квантование холловской проводимости.** Покажем теперь, как, используя выведенную выше формулу (9.84), можно получить квантование холловской проводимости в случае малой амплитуды хаотического потенциала  $V = \sqrt{\langle V^2(\mathbf{r}) \rangle} \ll \hbar\omega_c$ , когда отдельные уровни Ландау отделены энергетическими щелями (рис. 9.16).

Прежде всего обсудим вопрос о граничных условиях, которым удовлетворяют волновые функции  $\Psi_i$ . Для того, чтобы описать протекание электрического тока через образец, мы должны потребовать, чтобы электронная плотность и плотность тока с противоположных сторон образца совпадали. В отсутствие магнитного поля таким требованиям удовлетворяют периодические граничные условия. Однако в магнитном поле они должны быть видоизменены, чтобы удовлетворить требованиям калибровочной инвариантности<sup>1)</sup>. Воспользовавшись конкретным видом векторного потенциала (калибровкой Ландау,  $\mathbf{A} = (-By, 0, 0)$ ), мы можем записать граничные условия в следующем виде:

$$\Psi(x + L_x, y) = \exp(i\theta) \Psi(x, y), \quad (9.85)$$

$$\Psi(x, y + L_y) = \exp(i\varphi) \exp\left(-\frac{ixL_y}{a_H^2}\right) \Psi(x, y), \quad (9.86)$$

где  $\theta = \alpha L_x$  и  $\varphi = \beta L_y$  — фазы, набираемые волновой функцией на длине образца. В случае  $\alpha = \beta = 0$  граничные условия (9.85)–(9.86) соответствуют стандартным периодическим граничным условиям Борна–Кармана. Применим следующий прием: введем вспомогательные волновые функции  $\Phi_i(x, y) = \exp(-i\alpha x - i\beta y) \Psi_i(x, y)$ , удовлетворяющие граничным условиям  $\Phi(x + L_x, y) = \Phi(x, y)$ ,  $\Phi(x, y + L_y) = \exp(-ixL_y/a_H^2) \Phi(x, y)$ . Уравнение на волновые функции

<sup>1)</sup> Ландау Л.Д., Лифшиц Е.М. Квантовая механика (нерелятивистская теория). — М.: Наука, 1989. § 112.

$\Phi$ , полученное из (9.74), имеет следующий вид:

$$\left[ \frac{(\hat{\mathbf{p}} - e/c\mathbf{A} + \mathbf{\Lambda})^2}{2m} + V(\mathbf{r}) - E_i \right] \Phi_i(\mathbf{r}; \alpha, \beta) = 0, \quad \mathbf{\Lambda} = (\hbar\alpha, \hbar\beta). \quad (9.87)$$

При этом недиагональные ( $i \neq j$ ) матричные элементы оператора скорости в выражении (9.84) могут быть выражены с помощью волновых функций  $\Phi_i$  в следующем виде:

$$\begin{aligned} \mathbf{v}_{ij} &= \langle \Psi_i | \hat{\mathbf{v}} | \Psi_j \rangle = \langle \Psi_i | \frac{\hat{\mathbf{p}} - e/c\mathbf{A}}{m} | \Psi_j \rangle = \\ &= \langle \Phi_i | \frac{\hat{\mathbf{p}} - e/c\mathbf{A} + \mathbf{\Lambda}}{m} | \Phi_j \rangle = \langle \Phi_i | \hat{\mathbf{v}} | \Phi_j \rangle. \end{aligned} \quad (9.88)$$

Далее найдем вариацию волновых функций при изменении граничных условий. Пусть  $\Phi_i(\mathbf{r}; \alpha, \beta)$  — волновая функция, удовлетворяющая уравнению (9.87). Тогда  $\Phi_i(\mathbf{r}; \alpha + d\alpha, \beta + d\beta)$  удовлетворяет следующему уравнению:

$$\left[ \frac{(\hat{\mathbf{p}} - e/c\mathbf{A} + \mathbf{\Lambda})^2}{2m} + \hat{\mathbf{v}} d\mathbf{\Lambda} + \frac{(d\mathbf{\Lambda})^2}{2m} + V(\mathbf{r}) - E_i \right] \Phi_i(\mathbf{r}; \alpha, \beta) = 0, \quad (9.89)$$

где  $\hat{\mathbf{v}} = (\hat{\mathbf{p}} - e/c\mathbf{A} - \mathbf{\Lambda})/m$ ,  $d\mathbf{\Lambda} = \hbar(d\alpha, d\beta)$ . Учитывая второе слагаемое в уравнении (9.89) по теории возмущений, найдем вариацию волновых функций  $\Phi$  при изменении параметров  $\alpha$  и  $\beta$ :

$$\left| \frac{\partial \Phi_i}{\partial \alpha} \right\rangle = \sum_{k \neq i} \frac{\langle \Phi_k | \hat{v}_x | \Phi_i \rangle}{E_i - E_k} | \Phi_k \rangle, \quad (9.90)$$

$$\left| \frac{\partial \Phi_i}{\partial \beta} \right\rangle = \sum_{k \neq i} \frac{\langle \Phi_k | \hat{v}_y | \Phi_i \rangle}{E_i - E_k} | \Phi_k \rangle. \quad (9.91)$$

Отсюда получаем следующее соотношение:

$$\left\langle \frac{\partial \Phi_i}{\partial \alpha} \middle| \frac{\partial \Phi_i}{\partial \beta} \right\rangle = \hbar^2 \sum_{j, k \neq i} \frac{\langle \Phi_i | \hat{v}_x | \Phi_j \rangle \langle \Phi_k | \hat{v}_y | \Phi_i \rangle}{(E_i - E_j)(E_i - E_k)} \langle \Phi_i | \Phi_k \rangle = \hbar^2 \sum_{j \neq i} \frac{(v_x)_{ij} (v_y)_{ji}}{(E_i - E_j)^2}. \quad (9.92)$$

Формула (9.92) позволяет переписать выражение для холловской проводимости через вариации волновых функций при изменении параметров  $\alpha$  и  $\beta$ , определяющих вид граничных условий:

$$\sigma_{xy}(\alpha, \beta) = \frac{ie^2}{\hbar S} \sum_{E_i < E_F} \left[ \left\langle \frac{\partial \Phi_i}{\partial \alpha} \middle| \frac{\partial \Phi_i}{\partial \beta} \right\rangle - \left\langle \frac{\partial \Phi_i}{\partial \beta} \middle| \frac{\partial \Phi_i}{\partial \alpha} \right\rangle \right]. \quad (9.93)$$

Среди всевозможных граничных условий (9.85)–(9.86) неэквивалентными являются лишь те, для которых фазы, набираемые волновыми функциями на длине образца, удовлетворяют условиям  $0 \leq \theta = \alpha L_x < 2\pi$  и  $0 \leq \varphi = \beta L_y < 2\pi$ . Поскольку холловская проводимость макроскопической системы не должна зависеть от выбора фаз в граничных

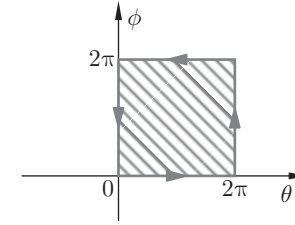


Рис. 9.20. Область интегрирования в формуле (9.94)

условиях (9.85)–(9.86), мы вправе отождествить  $\sigma_{xy}$  с его средним значением<sup>1)</sup>,

$$\begin{aligned} \bar{\sigma}_{xy} &= \frac{e^2}{h} \sum_{E_i < E_F} \frac{1}{2\pi i} \int_0^{2\pi} \int_0^{2\pi} d\theta d\varphi \left[ \left\langle \frac{\partial \Phi_i}{\partial \varphi} \middle| \frac{\partial \Phi_i}{\partial \theta} \right\rangle - \left\langle \frac{\partial \Phi_i}{\partial \theta} \middle| \frac{\partial \Phi_i}{\partial \varphi} \right\rangle \right] = \\ &= \frac{e^2}{h} \sum_{E_i < E_F} F[\Phi_i]. \end{aligned} \quad (9.94)$$

Здесь для удобства мы ввели обозначение  $F[\Phi_i]$  для усредненного по граничным условиям вклада в холловскую проводимость электронного состояния с собственной энергией  $E_i$ . Представим матричные элементы в явном виде и воспользуемся формулой Грина<sup>2)</sup>, чтобы преобразовать выражение (9.94) в интеграл по периметру квадрата, изображенного на рис. 9.20:

$$\begin{aligned} F[\Phi_i] &= \frac{1}{2\pi i} \int d^2\mathbf{r} \int_0^{2\pi} \int_0^{2\pi} d\theta d\varphi \left[ \frac{\partial}{\partial \varphi} \left( \Phi_i^* \frac{\partial \Phi_i}{\partial \theta} \right) - \frac{\partial}{\partial \theta} \left( \Phi_i^* \frac{\partial \Phi_i}{\partial \varphi} \right) \right] = \\ &= \frac{1}{2\pi i} \oint_C dl \langle \Phi_i | \nabla_l | \Phi_i \rangle. \end{aligned} \quad (9.95)$$

Поскольку значения 0 и  $2\pi$  параметров  $\theta$  и  $\varphi$  в формуле (9.85) эквивалентны, волновые функции  $\Phi_i(\mathbf{r}; \theta, \varphi)$  на противоположных сторонах квадрата, изображенного на рис. 9.20, могут различаться лишь не зависящим от  $\mathbf{r}$  фазовым множителем,

$$\Phi_i(\mathbf{r}; \theta, 2\pi) = \exp(iU(\theta)) \Phi_i(\mathbf{r}; \theta, 0), \quad (9.96)$$

$$\Phi_i(\mathbf{r}; 2\pi, \varphi) = \exp(iW(\varphi)) \Phi_i(\mathbf{r}; 0, \varphi). \quad (9.97)$$

<sup>1)</sup> Усреднение в формуле (9.94) производится по параметрам  $\theta$  и  $\varphi$ , характеризующим граничные условия (9.85)–(9.86).

<sup>2)</sup> Кудрявцев Л. Д. Математический анализ. Т. 2. С. 130.

Дифференцируя соотношения (9.96)–(9.97), имеем

$$\left. \frac{\partial \Phi_i}{\partial \theta} \right|_{(\theta, 2\pi)} = \left[ \left. \frac{\partial \Phi_i}{\partial \theta} \right|_{(\theta, 0)} + i \frac{dU}{d\theta} \Phi_i(\mathbf{r}; \theta, 0) \right] \exp(iU(\theta)), \quad (9.98)$$

$$\left. \frac{\partial \Phi_i}{\partial \varphi} \right|_{(2\pi, \varphi)} = \left[ \left. \frac{\partial \Phi_i}{\partial \varphi} \right|_{(0, \varphi)} + i \frac{dW}{d\varphi} \Phi_i(\mathbf{r}; 0, \varphi) \right] \exp(iW(\varphi)). \quad (9.99)$$

Подставляя (9.96)–(9.99) в формулу (9.95), получаем выражение для  $F[\Phi_i]$  через значения функций  $U(\theta)$  и  $W(\Phi_i)$  в точках  $0, 2\pi$ :

$$\begin{aligned} F[\Phi_i] &= \frac{1}{2\pi i} \int d^2 \mathbf{r} \left\{ \int_0^{2\pi} d\theta \left[ \Phi_i^*(\mathbf{r}; \theta, 0) \left. \frac{\partial \Phi_i}{\partial \theta} \right|_{(\theta, 0)} - \Phi_i^*(\mathbf{r}; \theta, 2\pi) \left. \frac{\partial \Phi_i}{\partial \theta} \right|_{(\theta, 2\pi)} \right] + \right. \\ &\quad \left. + \int_0^{2\pi} d\varphi \left[ \Phi_i^*(\mathbf{r}; 2\pi, \varphi) \left. \frac{\partial \Phi_i}{\partial \varphi} \right|_{(2\pi, \varphi)} - \Phi_i^*(\mathbf{r}; 0, \varphi) \left. \frac{\partial \Phi_i}{\partial \varphi} \right|_{(0, \varphi)} \right] \right\} = \\ &= \frac{1}{2\pi i} \int d^2 \mathbf{r} \left[ i \int_0^{2\pi} \frac{dW}{d\varphi} |\Phi_i^*(\mathbf{r}; 0, \varphi)|^2 d\varphi - i \int_0^{2\pi} d\theta \frac{dU}{d\theta} |\Phi_i^*(\mathbf{r}; \theta, 0)|^2 \right] = \\ &= \frac{1}{2\pi} \left[ \int_0^{2\pi} \frac{dW}{d\varphi} d\varphi - \int_0^{2\pi} d\theta \frac{dU}{d\theta} \right] = \frac{1}{2\pi} [W(2\pi) - W(0) - U(2\pi) + U(0)]. \end{aligned} \quad (9.100)$$

С другой стороны, значения функций  $U$  и  $W$  в точках  $0$  и  $2\pi$  не являются независимыми. Используем для нахождения связи между ними формулы (9.96)–(9.97),

$$\begin{aligned} \Phi_i(\mathbf{r}; 2\pi, 2\pi) &= \exp(iU(2\pi)) \Phi_i(\mathbf{r}; 2\pi, 0) = \\ &= \exp(iU(2\pi) + iW(0)) \Phi_i(\mathbf{r}; 0, 0), \end{aligned} \quad (9.101)$$

$$\begin{aligned} \Phi_i(\mathbf{r}; 2\pi, 2\pi) &= \exp(iW(2\pi)) \Phi_i(\mathbf{r}; 0, 2\pi) = \\ &= \exp(iW(2\pi) + iU(0)) \Phi_i(\mathbf{r}; 0, 0), \end{aligned} \quad (9.102)$$

откуда

$$W(2\pi) + U(0) - W(0) - U(2\pi) = 2\pi n_i, \quad n_i \in \mathbb{Z}. \quad (9.103)$$

Таким образом, мы показали, что  $F[\Phi_i] = n_i$ , где  $n_i$  — целое число, зависящее от номера электронного состояния  $i$ . Производя суммирование по заполненным электронным состояниям в формуле (9.94), получаем окончательный результат для холловской проводимости,

$$\bar{\sigma}_{xy} = N \frac{e^2}{h}, \quad N = \sum_{E_i < E_F} n_i \in \mathbb{Z}. \quad (9.104)$$

**9.2.3. Магнетосопротивление. Отрицательное магнетосопротивление.** Внешнее однородное магнитное поле  $\mathbf{H}$  не только приводит к появлению эффекта Холла, но и влияет на величину проводимости полупроводника. В отсутствие магнитного поля плотность тока равна

$$\mathbf{j} = \sigma \mathbf{E} = \mathbf{E}/\rho, \quad (9.105)$$

где  $\sigma$  и  $\rho$  — удельные электропроводность и сопротивление. Без магнитного поля электрон движется (в электрическом поле) прямолинейно и между двумя столкновениями проходит путь, равный длине свободного пробега  $l$ . Если включить магнитное поле, то траектория электрона будет представлять собой участок циклоиды. За время свободного пробега вдоль поля  $\mathbf{E}$  электрон пройдет путь, меньший, чем  $l$ . Поскольку за время свободного пробега  $\tau$  электрон проходит меньший путь вдоль поля  $\mathbf{E}$ , то это равносильно уменьшению дрейфовой скорости, или подвижности, а тем самым и проводимости, при этом сопротивление должно возрастать. Если магнитное поле  $\mathbf{H}$  направлено вдоль оси  $z$ , то направление тока  $\mathbf{j} \parallel x$  не совпадает с направлением полного электрического поля  $\mathbf{E}$ . Поэтому удельное сопротивление в магнитном поле можно определить как

$$\rho_H = E_x/j_x. \quad (9.106)$$

Тогда величина

$$\frac{\rho_H - \rho}{\rho} \equiv \frac{\Delta\rho}{\rho} \quad (9.107)$$

называется магнетосопротивлением<sup>1)</sup>. На опыте часто наблюдается, что при слабых магнитных полях  $\Delta\rho/\rho$  отрицательно. Прежде, чем в этом разобраться, получим основные выражения для магнетосопротивления на основе теории Лоренца–Друде. Воспользуемся выражениями для плотности тока из 9.2.1. Перепишем выражение (9.56) для плотности тока в виде

$$j_x = \frac{\sigma_{xx}\sigma_{yy} - \sigma_{xy}\sigma_{yx}}{\sigma_{yy}} E_x \equiv \frac{1}{\rho_H} E_x, \quad (9.108)$$

где

$$\rho_H = \frac{\sigma_{yy}}{\sigma_{xx}\sigma_{yy} - \sigma_{xy}\sigma_{yx}} = \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{xy}^2}. \quad (9.109)$$

Рассмотрим, как и при исследовании эффекта Холла, два случая.

**А. Сильные магнитные поля**  $\omega_H \tau \gg 1$ . В этом случае мы можем воспользоваться результатами 9.2.1. Подставив (9.58) в (9.109), получаем

$$\frac{\Delta\rho}{\rho} = \langle \tau \rangle \left\langle \frac{1}{\tau} \right\rangle - 1.$$

Мы видим, что в случае сильных магнитных полей магнетосопротивление  $\Delta\rho/\rho$  достигает насыщения.

<sup>1)</sup> Иногда в литературе магнетосопротивлением называют  $\rho_H$ .

**В. Слабые магнитные поля**  $\omega_H \tau \ll 1$ . Как уже было отмечено в 9.2.1, в этом случае диагональные компоненты проводимости  $\sigma_{xx} \gg \sigma_{xy}$ . Для получения выражения для  $\Delta\rho$  необходимо в  $\sigma_{xx}$  сохранить первые поправки по  $\omega_H \tau$  (см. (9.53)):

$$\sigma_{xx} = \sigma_{yy} \approx \frac{e^2 n}{\langle \tau \rangle} \langle \tau - \omega_H^2 \tau^3 \rangle = \sigma_0 \left( 1 - \frac{\langle \tau^3 \rangle}{\langle \tau \rangle} \omega_H^2 \right). \quad (9.110)$$

Здесь мы учли, что  $\sigma_0 = \frac{e^2 n}{m} \langle \tau \rangle$ . Для  $\sigma_{xy}$  мы можем воспользоваться выражением (9.60):

$$\sigma_{xy} = -\sigma_{yx} = \frac{e^2 n}{m} \omega_H \langle \tau^2 \rangle = \sigma_0 \frac{\langle \tau^2 \rangle}{\langle \tau \rangle} \omega_H. \quad (9.111)$$

Подставляя (9.110) и (9.111) в (9.109), получаем

$$\rho_H = \frac{1}{\sigma_0} + \frac{\omega_H^2}{\sigma_0} \left( \frac{\langle \tau^3 \rangle}{\langle \tau \rangle} - \frac{\langle \tau^2 \rangle e^2}{\langle \tau \rangle e^2} \right). \quad (9.112)$$

В результате для магнетосопротивления, согласно (9.107), получаем выражение

$$\frac{\Delta\rho}{\rho} = \omega_H^2 \left( \frac{\langle \tau^3 \rangle}{\langle \tau \rangle} - \frac{\langle \tau^2 \rangle e^2}{\langle \tau \rangle e^2} \right). \quad (9.113)$$

Отметим важное обстоятельство. Если время релаксации  $\tau$  не зависит от энергии, то магнетосопротивление в точности равно нулю.

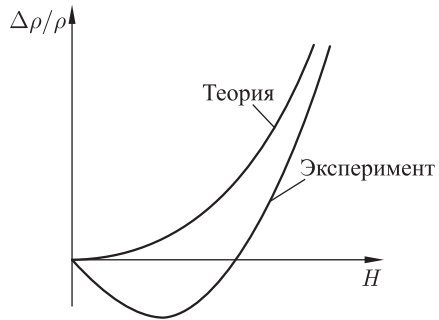


Рис. 9.21. Схематическая зависимость магнетосопротивления от магнитного поля (теория и эксперимент). Экспериментальное значение  $\Delta\rho/\rho$  при слабых магнитных полях отрицательно

Кратко поясним полученный результат. Рассмотрим ограниченный полупроводник. Если время релаксации не зависит от энергии, то под действием силы Лоренца электроны отклоняются одинаковым образом. В этом случае холловское поле компенсирует действие магнитного поля, в результате чего носители заряда движутся прямолинейно, поэтому магнетосопротивление равно нулю. Однако, если  $\tau$  зависит от энергии, или, что то же самое, если скорости электронов (дырок) различны, то на частицы, движущиеся со скоростями, большими средней скорости, магнитное поле действует сильнее, чем холловское. Тогда холловское поле компенсирует разность потенциалов не для всех носителей; и это отличие тем больше, чем сильнее  $\tau$  зависит от энергии.

Тогда холловское поле компенсирует разность потенциалов не для всех носителей; и это отличие тем больше, чем сильнее  $\tau$  зависит от энергии.

В результате зависимости  $\tau$  от энергии уменьшается вклад в проводимость быстрых и медленных носителей заряда, что и приводит к увеличению сопротивления  $\rho_H$ . Эффект магнетосопротивления оказывается чувствительным к форме образца. Далее, можно показать, что в рамках теории, изложенной выше, величина  $\Delta\rho/\rho$  (см. (9.113)) всегда положительна независимо от того, как  $\tau$  зависит от энергии. Теоретическая зависимость магнетосопротивления от магнитного поля представлена на рис. 9.21. На этом же рисунке представлена экспериментальная зависимость  $\Delta\rho/\rho$  от магнитного поля.

На опыте получается, что при слабых магнитных полях и при низких температурах  $\Delta\rho/\rho$  отрицательно. Это удивительное явление получило объяснение в 1980 г.<sup>1)</sup>

**9.2.3.1. Отрицательное магнетосопротивление.** Достоверно доказано, что эффект отрицательного магнетосопротивления не связан с типом примесей в полупроводнике. Причина этого явления имеет фундаментальный характер. Выше, при исследовании кинетических явлений, мы считали электрон классической частицей. Это означает, что длина свободного пробега электрона много больше длины волны де-Бройля:  $p_0 l \gg \hbar$ , где  $p_0$  — средний импульс электрона.

Однако, если считать, что электроны — это волны, то они могут интерферировать. Рассмотрим модель не взаимодействующих друг с другом электронов, которые рассеиваются, например, на примесях. Пусть электрон, испытывая рассеяния на примесях, переходит из точки  $a$  в точку  $b$  (рис. 9.22).

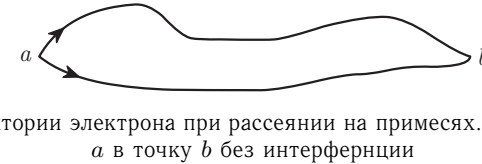


Рис. 9.22. Траектории электрона при рассеянии на примесях. Переход из точки  $a$  в точку  $b$  без интерференции

Он может при этом пройти по разным путям. Согласно квантовой механике должны складываться не вероятности этих путей, а соответствующие им амплитуды. Если мы обозначим эти амплитуды через  $A_i$ , то полная вероятность перехода из точки  $a$  в точку  $b$  будет равна квадрату модуля суммы всех амплитуд, то есть

$$W = \left| \sum_i A_i \right|^2 = \sum_i |A_i|^2 + \sum_{i \neq j} A_i A_j^*. \quad (9.114)$$

<sup>1)</sup> Альтшулер Б. Л., Хмельницкий Д. Е., Ларкин А. И., Ли П. А. 1980 (Altshuler B. L., Khmel'nitskii D. E., Larkin A. I., Lee P. A. Magnetoresistance and Hall Effect in a Disordered Two-Dimensional Electron Gas // Phys. Rev. B. 1980. V. 22. P. 5142).

Первое слагаемое в (9.114) описывает сумму вероятностей прохождения электроном каждого пути из точки  $a$  в точку  $b$ . Второе слагаемое в (9.114) соответствует интерференции разных амплитуд. Для большинства траекторий интерференция в (9.114) не важна, так как длины этих траекторий сильно отличаются, и поэтому сильно отличаются изменения фазы волновой функции электрона на этих траекториях

$$\Delta\varphi = \frac{1}{\hbar} \int_a^b \mathbf{p} d\mathbf{l}. \quad (9.115)$$

Здесь  $\mathbf{p}$  — импульс электрона;  $d\mathbf{l}$  — элемент траектории электрона. При суммировании по всем траекториям (в силу того, что  $p_0 l \gg \hbar$ )

интерференционное слагаемое, ввиду его осциллирующего характера, обращается в нуль.

Однако среди всех траекторий есть особые траектории, для которых нельзя пренебречь интерференцией, даже если  $p_0 l \gg \hbar$ . Это траектории с самопересечением (что соответствует интерференции при рассеянии назад) (рис. 9.23). Каждая такая траектория соответствует двум амплитудам, отличающимся направлением прохождения петли. При изменении направления движения вдоль петли происходит замена  $\mathbf{p}$

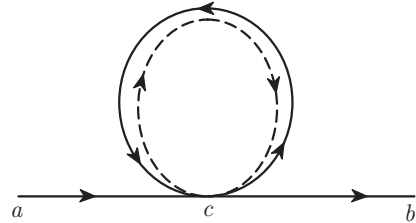


Рис. 9.23. Интерференционные замкнутые траектории (с самопересечением) электрона при рассеянии на примесях. Траектории, вносящие основной вклад в эффект слабой локализации

на  $-\mathbf{p}$  и  $d\mathbf{l}$  на  $-d\mathbf{l}$ . В результате,  $\Delta\varphi$ , согласно (9.115), для петли не меняется. Таким образом, две амплитуды оказываются когерентными, поскольку имеют одинаковые фазы (когда разность фаз между ними равна нулю). Тогда полная вероятность перехода из точки  $a$  в точку  $b$  равна

$$W = |A_1 + A_2|^2 = |A_1|^2 + |A_2|^2 + A_1 A_2^* + A_2 A_1^* = 4 |A_1|^2. \quad (9.116)$$

Мы получили величину, в два раза большую, чем при сложении вероятностей этих путей. Интерференция амплитуд рассеяния имеет квантовое происхождение, в то время как сложение вероятностей соответствует классическому описанию электрона с помощью кинетического уравнения. Таким образом, мы видим, что интерференционные эффекты важны для траекторий с самопересечением (то есть при рассеянии назад), и они приводят (согласно (9.116)) к увеличению полной вероятности рассеяния, а следовательно, к уменьшению проводимости, или, что то же самое, к увеличению сопротивления.

Весьма интересно поведение интерференционной картины в магнитном поле. В присутствии внешнего магнитного поля происходит замена  $\mathbf{p} \rightarrow \mathbf{p} - \frac{e}{c} \mathbf{A}$ , где  $\mathbf{A}$  — векторный потенциал магнитного поля. При изменении направления обхода электроном замкнутой петли импульс  $\mathbf{p}$  меняет знак на обратный, но векторный потенциал  $\mathbf{A}$  сохраняет свой знак. В результате у интерферирующих амплитуд появляется разность фаз

$$\Delta\varphi_H = \frac{2e}{c\hbar} \oint \mathbf{A} d\mathbf{l} = \frac{2e}{c\hbar} \int \text{rot } \mathbf{A} d\mathbf{S} = \frac{2e}{c\hbar} \Phi, \quad (9.117)$$

где  $\Phi$  — поток магнитного поля через петлю (рис. 9.23). При выводе (9.117) мы воспользовались теоремой Стокса. Введем квант магнитного потока

$$\Phi_0 = \pi \frac{c\hbar}{e}. \quad (9.118)$$

Тогда выражение для разности фаз принимает вид

$$\Delta\varphi_H = 2\pi \frac{\Phi}{\Phi_0}. \quad (9.119)$$

Итак, появление разности фаз приводит к разрушению интерференционной картины, а следовательно, к увеличению проводимости, или к уменьшению сопротивления. В результате интерференционная поправка приводит к отрицательному магнетосопротивлению.

Явление отрицательного магнетосопротивления — это проявление квантовых свойств электронов в кинетических явлениях.

### 9.3. Кинетическое уравнение

**Классическое кинетическое уравнение Больцмана.** Во многих главах мы рассмотрели физические свойства полупроводников, находящихся в состоянии термодинамического равновесия. Наряду с такими равновесными состояниями, большой теоретический и практический интерес представляет изучение электронов (дырок) проводимости в неравновесном состоянии, когда они движутся под действием приложенных внешних полей: электрического, магнитного или температурного. Такие процессы, связанные с перемещением электронов и дырок, называются явлениями переноса, или кинетическими эффектами.

Важной особенностью неравновесных процессов является то, что они существенно зависят от механизмов взаимодействия в системе. В нашем случае такими механизмами являются взаимодействия электрона проводимости с колебаниями решетки или дефектами кристалла.

В главе 5 мы видели, что в полупроводнике электроны, находящиеся в состоянии термодинамического равновесия, в классическом случае описываются равновесной функцией распределения Больцмана

$$f_0 = \exp\left(\frac{\mu - \varepsilon}{k_B T}\right), \quad (9.120)$$



где  $\varepsilon$  — полная энергия электрона;  $\mu$  — уровень химического потенциала. Аналогично для электронов в неравновесном состоянии можно ввести неравновесную функцию распределения в фазовом пространстве (6-мерном пространстве координат и импульсов)  $f(\mathbf{r}, \mathbf{p}, t)$ . Смысл функции распределения состоит в том, что

$$f(\mathbf{r}, \mathbf{p}, t) d^3r d^3p \quad (9.121)$$

есть число электронов в момент времени  $t$  в объеме  $d^3r = dx dy dz$  около точки  $\mathbf{r}$  и с импульсами, лежащими между  $p_x$  и  $p_x + dp_x$ ,  $p_y$  и  $p_y + dp_y$  и  $p_z$  и  $p_z + dp_z$ . Такое описание электронов посредством одновременного задания их координат и сопряженных импульсов возможно только в той мере, в какой их движение подчиняется законам классической механики.

Если известна функция  $f(\mathbf{r}, \mathbf{p}, t)$ , то можно определить любую физическую величину, зависящую от координат и импульса частиц, например, вычислить плотность тока электронов в точке  $\mathbf{r}$  в момент времени  $t$ . Чтобы вывести основное уравнение кинетической теории, то есть уравнение, определяющее функцию распределения электронов  $f(\mathbf{r}, \mathbf{p}, t)$ , мы пока пренебрежем процессами столкновения и будем считать, что электроны движутся во внешних полях (например, в электрическом поле). Тогда можно воспользоваться теоремой Лиувилля, которая гласит, что при движении электронов (дырок) вдоль фазовых траекторий функция распределения не зависит от времени. Следовательно<sup>1)</sup>,

$$\frac{df}{dt} = 0. \quad (9.122)$$

Полная производная в (9.122) означает дифференцирование вдоль фазовой траектории электрона, определяемой его уравнениями движения. Напомним, что теорема Лиувилля имеет место для функции распределения, определенной именно как плотность в фазовом пространстве (то есть в пространстве переменных  $\mathbf{p}$  и  $\mathbf{r}$ ). Распишем в явном виде полную производную в (9.122)

$$\frac{\partial f}{\partial t} + \mathbf{v} \frac{\partial f}{\partial \mathbf{r}} + \mathbf{F} \frac{df}{d\mathbf{p}} = 0. \quad (9.123)$$

Здесь мы учли, что

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}, \quad \frac{d\mathbf{p}}{dt} = \mathbf{F}. \quad (9.124)$$

С помощью уравнения (9.122) (то есть с помощью теоремы Лиувилля) можно решить многие задачи физической кинетики. Уравнение (9.123) имеет прозрачный физический смысл: функция распределения изменяется в силу того, что частицы перемещаются в координатном

пространстве (член с  $\mathbf{v} \frac{\partial f}{\partial \mathbf{r}}$ ), и в силу действия внешних сил, которые приводят к перемещению частиц в импульсном пространстве ( $\mathbf{F} \frac{df}{d\mathbf{p}}$ ).

Учет столкновений нарушает равенство (9.122), при этом функция распределения перестает быть постоянной вдоль фазовых траекторий. Поэтому вместо (9.122) надо писать

$$\frac{df}{dt} = I(f), \quad (9.125)$$

где  $I(f)$  — так называемый интеграл столкновений. Интеграл столкновений — это сложный нелинейный функционал от функций распределения, структура и конкретный вид которого определяются взаимодействием электронов с примесями, друг с другом или с другими квазичастицами (например фононами). В последнем случае система (9.125) должна быть дополнена кинетическими уравнениями для функций распределения соответствующих квазичастиц (фононов). Уравнения вида (9.125) называют кинетическими уравнениями.

**Замечание.** Запись кинетического уравнения (9.125) возможна только в тех случаях, когда движение частицы можно разделить на движение по фазовым траекториям и на столкновения, то есть на резкие изменения импульсов частиц без заметного (с макроскопической точки зрения) изменения координаты. Отсюда следует, что интеграл столкновений — это оператор, связанный с зависимостью функции распределения от импульса, но не от координаты или времени<sup>1)</sup>. Кинетическое уравнение приобретает смысл лишь после установления вида интеграла столкновений. К этому вопросу мы сейчас и перейдем. Рассмотрим, например, рассеяние электронов на примесях.

Пусть система электронов описывается функцией распределения  $f$ . Благодаря процессам рассеяния число электронов с импульсами в заданном элементе объема импульсного пространства в окрестности  $\mathbf{p}$  может возрасти (в результате переходов  $\mathbf{p}' \rightarrow \mathbf{p}$ ) или уменьшиться (при переходах  $\mathbf{p} \rightarrow \mathbf{p}'$ ). Каждый процесс может произойти лишь в том случае, если в начальном состоянии есть электрон, а конечное состояние не заполнено (поскольку принцип Паули не позволяет двум электронам находиться в одном состоянии). Поэтому если речь идет о системе электронов, то каждая индивидуальная вероятность рассеяния  $W(\mathbf{p}, \mathbf{p}')$  должна быть умножена на функцию распределения  $f$  для начального состояния и на  $1 - f$  для конечного состояния. Тогда

<sup>1)</sup> Здесь не рассматриваются эффекты слабой локализации (см. Гантмахер В. Ф. Электроны в неупорядоченных средах. — М.: ФИЗМАТЛИТ, 2003. 176 с.).

<sup>1)</sup> См. Ландау Л. Д., Лифшиц Е. М. Статистическая физика. — М.: Наука, 1976.

получаем

$$I(f) = -f(\mathbf{p}, \mathbf{r}, t) \sum_{\mathbf{p}'} W(\mathbf{p}, \mathbf{p}') [1 - f(\mathbf{p}', \mathbf{r}, t)] + \\ + \sum_{\mathbf{p}'} W(\mathbf{p}', \mathbf{p}) f(\mathbf{p}', \mathbf{r}, t) [1 - f(\mathbf{p}, \mathbf{r}, t)] = \\ = \sum_{\mathbf{p}'} W(\mathbf{p}, \mathbf{p}') \{f(\mathbf{p}') - f(\mathbf{p})\}. \quad (9.126)$$

Первое слагаемое в (9.126) соответствует уменьшению числа частиц за счет перехода из состояния  $\mathbf{p}$  в состояние  $\mathbf{p}'$ . Второе слагаемое описывает переход электрона из состояния с импульсом  $\mathbf{p}'$  в состояние с импульсом  $\mathbf{p}$ .

При выводе (9.126) мы воспользовались принципом детального равновесия, согласно которому вероятности прямого и обратного упругих процессов одинаковы,

$$W(\mathbf{p}, \mathbf{p}') = W(\mathbf{p}', \mathbf{p}). \quad (9.127)$$

С учетом (9.123), (9.126) и (9.127) мы можем написать кинетическое уравнение

$$\frac{\partial f}{\partial t} + \mathbf{v} \frac{\partial f}{\partial \mathbf{r}} + \mathbf{F} \frac{\partial f}{\partial \mathbf{p}} = \sum_{\mathbf{p}'} W(\mathbf{p}, \mathbf{p}') \{f(\mathbf{p}') - f(\mathbf{p})\}. \quad (9.128)$$

Это интегродифференциальное уравнение называют уравнением Больцмана. Оно было впервые установлено Людвигом Больцманом в 1872 г.

В общем случае кинетическое уравнение Больцмана (9.125) — это система сложных нелинейных уравнений, которые при точном задании граничных и начальных условий однозначно определяют состояние твердого тела. В общем случае эту систему решить невозможно, и требуется значительное число упрощений, определяемых физической постановкой задачи.

Если в уравнение (9.128) подставить равновесную функцию распределения  $f_0$ , зависящую лишь от энергии, то правая часть тождественно обращается в нуль. Если  $\mathbf{F} = 0$ , то и левая часть равна нулю.

Второе важное замечание состоит в том, что при выводе кинетического уравнения столкновения электронов рассматривались как мгновенные акты, происходящие в одной точке пространства. Ясно, что кинетическое уравнение позволяет, в принципе, следить за изменением функции распределения лишь за промежутки времени, большие по сравнению с длительностью столкновений, и на расстояниях, больших по сравнению с размерами области столкновения. Область столкновения — это величина порядка радиуса рассеяния электрона  $R_0$ , время же столкновения — порядка  $R_0/\bar{v}$ , где  $\bar{v}$  — средняя скорость электрона. Эти значения и устанавливают нижний предел расстояний и длительностей, рассмотрение которых допускается кинетическим уравнением.

Ниже, в качестве примера, рассмотрим невырожденный электронный газ во внешнем стационарном и однородном в пространстве электрическом поле  $\mathbf{E}$ . Так как внешнее электрическое поле, как правило, очень мало по сравнению с внутренними атомными полями, то отклонение системы электронов от состояния равновесия оказывается в большинстве случаев весьма малым. Это позволяет линеаризовать систему уравнений (9.128), подставляя вместо функции распределения  $f$  сумму

$$f = f_0 + f_1, \quad (9.129)$$

где  $f_0$  — равновесная функция Ферми или Больцмана (нулевое приближение), а малость функции  $f_1$  (первое приближение) обеспечивается малостью внешнего воздействия. Итак,  $|f_1| \ll f_0$ . Другими словами, функция  $f_1$  пропорциональна тем внешним силам, которые вывели систему из состояния равновесия (в данном частном случае функция  $f_1$  пропорциональна электрическому полю, поэтому по полупроводнику течет электрический ток). Направим внешнее электрическое поле вдоль оси  $z$ . В линейном приближении, с учетом стационарности и однородности в пространстве электрического поля кинетическое уравнение (9.128) в соотношениях (9.129) принимают вид

$$eE \frac{\partial f_0(\mathbf{p})}{\partial p_z} = \sum_{\mathbf{p}'} W(\mathbf{p}, \mathbf{p}') \{f_1(\mathbf{p}') - f_1(\mathbf{p})\}. \quad (9.130)$$

В правой части (9.130) слагаемые, пропорциональные  $f_0$ , в сумме дают нуль, так что остаются только слагаемые, пропорциональные неравновесной добавке  $f_1$ . В левой части уравнения (9.130), наоборот, мы сохраняем только слагаемое с равновесной функцией  $f_0$ , так как это слагаемое умножается на малую величину (на электрическое поле).

Дальнейшее упрощение интеграла столкновений в (9.130) может быть проведено для изотропной модели. Мы рассматриваем случай упругого рассеяния, в котором импульсы электронов до и после рассеяния  $\mathbf{p}$  и  $\mathbf{p}'$  равны по модулю:  $|\mathbf{p}'| = |\mathbf{p}|$ . Таким образом, при рассеянии электрона изменяется только направление его импульса, а величина импульса не изменяется (рассматривается параболический спектр).

Итак, при наличии электрического поля добавка к функции распределения  $f_1$  должна быть скаляром и в первом приближении должна линейно зависеть от  $\mathbf{E}$ . С учетом (9.130) представим  $f_1$  в виде

$$f_1 = eE \frac{\partial f_0}{\partial p_z} \eta(\varepsilon), \quad (9.131)$$

где  $\eta(\varepsilon)$  — некоторая функция энергии  $\varepsilon$ . Ниже мы будем считать, что энергетический спектр электронов квадратичный. Тогда для квадратичного спектра скорость электрона равна

$$v_z = \frac{p_z}{m}. \quad (9.132)$$

Следовательно,

$$\frac{\partial f_0}{\partial p_z} = \frac{\partial f_0}{\partial \varepsilon} \frac{\partial \varepsilon}{\partial p_z} = \frac{\partial f_0}{\partial \varepsilon} v_z = \frac{\partial f_0}{\partial \varepsilon} \frac{p_z}{m}. \quad (9.133)$$

Подставляя (9.131) в интеграл столкновений (9.130), считая столкновения упругими ( $\varepsilon = \varepsilon'$ ) и учитывая соотношение (9.133), получим

$$p_z = \eta(\varepsilon) \sum_{\mathbf{p}'} W(\mathbf{p}, \mathbf{p}') (p'_z - p_z). \quad (9.134)$$

Так как мы предполагали, что рассеяние электронов упругое, то вероятность  $W(\mathbf{p}, \mathbf{p}')$  зависит лишь от угла между  $\mathbf{p}'$  и  $\mathbf{p}$ , который мы обозначим через  $\theta$  и от  $|\mathbf{p}| = |\mathbf{p}'|$ . Проекция  $\mathbf{p}'$  и  $\mathbf{p}$  на ось  $z$  имеют вид

$$p_z = p \cos v, \quad p'_z = p \cos v'. \quad (9.135)$$

Здесь  $v$  — угол между импульсом  $\mathbf{p}$  и осью  $z$ ;  $v'$  — угол между  $\mathbf{p}'$  и осью  $z$ , а угол рассеяния, как отмечено выше, равен  $\theta$ . Из сферической геометрии известно, что угол между двумя векторами  $v'$  выражается через сферические углы этих векторов,

$$\cos v' = \cos v \cos \theta + \sin v \sin \theta \cos(\varphi' - \varphi), \quad (9.136)$$

где  $(\varphi' - \varphi)$  — угол между проекциями  $\mathbf{p}'$  и  $\mathbf{p}$  на плоскость, перпендикулярную  $\mathbf{p}$ .

Далее, нужно подставить (9.135) и (9.136) в (9.134). Поскольку относительно  $\varphi'$  имеется цилиндрическая симметрия, то при интегрировании в (9.134) по  $\varphi'$  слагаемое, пропорциональное  $\cos(\varphi' - \varphi)$ , дает нуль. В результате выражение (9.134) после всех этих преобразований принимает вид

$$p = \eta(\varepsilon) \sum_{\mathbf{p}'} p W(\theta) (\cos \theta - 1). \quad (9.137)$$

Множитель в правой части (9.137) при функции  $\eta(\varepsilon)$  есть не что иное, как транспортное сечение рассеяния (см. § 6.3). Таким образом, мы получили выражение для обратного времени рассеяния

$$\frac{1}{\tau} = \sum_{\mathbf{p}'} W(\theta) (1 - \cos \theta). \quad (9.138)$$

В результате введенная нами функция энергии  $\eta(\varepsilon)$  имеет вид

$$\eta(\varepsilon) = -\tau. \quad (9.139)$$

Таким образом, окончательное выражение для первой поправки к функции распределения получает следующий вид:

$$f_1 = -eE\tau(\varepsilon) \frac{\partial f_0}{\partial \varepsilon} \frac{p_z}{m}. \quad (9.140)$$

Электрический ток выражается через функцию распределения следующим образом:

$$\mathbf{j} = 2e \int \mathbf{v} f \frac{d^3 \rho}{(2\pi\hbar)^3}. \quad (9.141)$$

Здесь множитель «2» соответствует двум проекциям спина электрона. Таким образом, для вычисления плотности тока мы должны умножить скорость одной частицы на функцию распределения и проинтегрировать по всем разрешенным значениям импульса. Если подставить в (9.141) вместо  $f$  функцию  $f_0(\varepsilon)$ , то интеграл обратиться в нуль. Это связано с тем, что энергия  $\varepsilon$  — четная функция  $\mathbf{p}$ , а  $\mathbf{v}$  — нечетная функция  $\mathbf{p}$ . Поэтому в (9.141) в качестве  $f$  надо подставить  $f_1$ , согласно (9.140). При этом находим

$$j_z = -e^2 E \int \frac{p_z^2}{m^2} \tau(\varepsilon) \frac{\partial f_0}{\partial \varepsilon} \rho(\varepsilon) d\varepsilon \frac{d\Omega}{4\pi}. \quad (9.142)$$

Здесь  $\rho(\varepsilon)$  — плотность состояний электронов,  $d\Omega = \sin \theta d\theta d\varphi$  — элемент телесного угла. Если электроны невырождены, то

$$\frac{\partial f_0}{\partial \varepsilon} = -\frac{f_0}{k_B T}.$$

Кроме того, учтем, что  $\frac{p_z^2}{m} = \frac{2}{3}\varepsilon$ . Используя это обстоятельство и интегрируя по углам, получаем

$$j_z = E \frac{2}{3} \frac{e^2}{m} \int d\varepsilon \frac{\varepsilon}{k_B T} f_0(\varepsilon) \tau(\varepsilon) \rho(\varepsilon). \quad (9.143)$$

Проводимость  $\sigma$  определяется как коэффициент пропорциональности между  $\mathbf{j}$  и  $\mathbf{E}$ :  $\mathbf{j} = \sigma \mathbf{E}$ . Таким образом, находим

$$\sigma = \frac{2}{3} \frac{e^2}{m} \int d\varepsilon \frac{\varepsilon}{k_B T} f_0(\varepsilon) \tau(\varepsilon) \rho(\varepsilon) \equiv \frac{ne^2}{m} \langle \tau \rangle, \quad (9.144)$$

где  $n$  — концентрация электронов,

$$\langle \tau \rangle = \frac{2}{3n} \int d\varepsilon \frac{\varepsilon}{k_B T} f_0(\varepsilon) \tau(\varepsilon) \rho(\varepsilon). \quad (9.145)$$

Это выражение показывает, что для расчета проводимости следует усреднять именно время рассеяния  $\tau$ . Таким образом, мы получили общее правило, что время релаксации носителей  $\tau$ , если оно зависит от энергии, необходимо усреднять с весом

$$\frac{2}{3n} \frac{\varepsilon}{k_B T} f_0(\varepsilon) \rho(\varepsilon) = F(\varepsilon), \quad (9.146)$$

то есть

$$\langle \tau \rangle = \int d\varepsilon F(\varepsilon) \tau. \quad (9.147)$$

В предыдущих выражениях учтено, что функция распределения  $f_0$  нормирована на единицу. Если  $f_0$  не нормировать, то среднее значение времени релаксации следует вычислять по формуле

$$\langle \tau \rangle = \frac{\int d\varepsilon F(\varepsilon) \tau}{\int d\varepsilon f_0(\varepsilon) \rho(\varepsilon)}. \quad (9.148)$$

Подчеркнем, что формулы (9.146)–(9.148) верны для невырожденного электронного газа.

В заключение сделаем несколько замечаний относительно кинетического уравнения.

**Замечание 1.** Изложенный выше вывод кинетического уравнения Больцмана удовлетворителен с физической точки зрения, однако представляет значительный интерес проследить, каким образом это уравнение можно аналитически получить из уравнений движения частиц. Такой вывод впервые был дан Н. Н. Боголюбовым (1946 г.). Значение метода Боголюбова состоит также и в том, что он дает регулярную процедуру, позволяющую получить не только уравнение Больцмана, но и поправки к нему (в отсутствие магнитного поля)<sup>1)</sup>. Исходным пунктом метода Боголюбова является также теорема Лиувилля для функции распределения электронов в целом как системы  $N$  частиц.

**Замечание 2.** При выводе кинетического уравнения Больцмана для описания поведения электронного газа мы воспользовались классическим подходом. Например, для описания кинетических явлений в полупроводниках не всегда можно пренебрегать интерференцией между электронными волновыми пакетами (см. § 2.3).

Очевидно, что при исследовании поведения неравновесных квантовых систем следовало бы прибегнуть к более совершенной методике, в которой используется матрица плотности<sup>2)</sup>.

## Задачи

**Задача 9.1.** Установите правила отбора квантовых переходов в условиях циклотронного резонанса в случае параболического спектра.

**Ответ.** Разрешены переходы между соседними уровнями Ландау (то есть между уровнями с номерами  $n$  и  $n + 1$  при поглощении кванта и  $n$  и  $n - 1$  при излучении).

**Задача 9.2.** Получите выражения для тензора проводимости в магнитном поле из кинетического уравнения.

<sup>1)</sup> См. Лифшиц Е. М., Питаевский Л. П. Физическая кинетика. — М.: Наука, 1979. 528 с.

<sup>2)</sup> Подробнее об этих методах можно узнать там же.

**Указание.** Представьте функцию распределения в виде  $f_0 + f_1$ , где  $f_1$  — неравновесная добавка, обусловленная наличием электрического и магнитного полей. Эта поправка удовлетворяет уравнению

$$-e\mathbf{E} \frac{\partial f_0}{\partial \mathbf{p}} - \frac{e}{c} [\mathbf{vH}] \frac{\partial f_1}{\partial \mathbf{p}} = -\frac{f_1(\mathbf{p})}{\tau}.$$

Решение этого уравнения удобно искать в виде

$$f_1(\mathbf{p}) = \frac{\partial f_0}{\partial E} \mathbf{A}(E) \mathbf{p},$$

где  $\mathbf{A}$  — некоторый вектор, который зависит только от энергии электрона  $E$ . Направление этого вектора фиксировано.

**Задача 9.3.** Хорошо известно, что при протекании электрического тока поле совершает работу над электронами. Это должно приводить к постоянному увеличению их кинетической энергии. Почему этим явлением пренебрегают при рассмотрении кинетических явлений? Какие процессы ответственны за установление стационарного распределения носителей? Оцените величину разогрева электронного газа. Меняется ли этот баланс энергий в магнитном поле?

**Указание.** Энергия, получаемая электронами от поля, частично уходит в решетку за счет взаимодействия электронов с фононами.

**Задача 9.4.** Пусть электрон совершает диффузионное движение в неупорядоченной среде. Известно, что квантовая (интерференционная) поправка к проводимости пропорциональна вероятности возврата электрона в окрестность начальной точки размером порядка длины волны. Оцените квантовую поправку к проводимости для электрона в объемном полупроводнике.

**Решение.** Найдем вероятность того, что в интервале времени  $(t, t + dt)$ , где  $t$  — время, отсчитанное от начала движения, электрон вернется в исходную точку. Ясно, что за время  $t$  объем, охваченный траекторией, будет порядка

$$V(t) \sim (Dt)^{3/2}.$$

Здесь  $D$  — коэффициент диффузии,  $\sqrt{Dt}$  — характерное смещение электрона. За время  $dt$  электрон вернется в требуемую окрестность из объема

$$dV \sim \frac{v dt}{k^2},$$

где  $v$  — скорость электрона;  $k$  — его волновой вектор. Искомая вероятность возврата есть  $dV/V$ , а полная вероятность возврата, которой пропорциональна относительная поправка к проводимости, есть

$$\frac{\delta \sigma}{\sigma} \sim \int \frac{dV}{V} = \int \frac{v dt}{k^2 (Dt)^{3/2}}.$$

Этот интеграл расходится на нижнем пределе. Очевидно, однако, что диффузионное описание верно лишь на временах, больших чем время столкновений  $\tau$ , поэтому в качестве нижнего предела интеграла нужно выбрать  $\tau$ . Кроме того, слишком длинные траектории не могут давать вклада в интерференцию, поскольку неизбежны процессы неупругого рассеяния, которые меняют энергию электрона, а значит, и его фазу. Такие процессы нарушают интерференцию. Эти процессы называются процессами сбоя фазы и характеризуются временем  $\tau_\varphi$ . Окончательно,

$$\frac{\delta\sigma}{\sigma} \sim -\frac{1}{k^2 l} \left( \frac{1}{v\tau} - \frac{1}{\sqrt{D\tau_\varphi}} \right) < 0.$$