



САНКТ-ПЕТЕРБУРГСКИЙ  
ГОСУДАРСТВЕННЫЙ  
УНИВЕРСИТЕТ

ПРИОРИТЕТНЫЙ  
НАЦИОНАЛЬНЫЙ ПРОЕКТ  
"ОБРАЗОВАНИЕ"



**Проект «Инновационная образовательная среда в классическом университете»**

Пилотный проект № 22 «Разработка и внедрение инновационной образовательной программы «Прикладные математика и физика»»

Физический факультет

Кафедра фотоники

**А.А.Загрубский, А.П.Чернова**

**СТАТИСТИКА НОСИТЕЛЕЙ ЗАРЯДА  
И КИНЕТИЧЕСКИЕ ПРОЦЕССЫ**

Учебно-методическое пособие

Санкт Петербург

2007 г.

УДК 537.9, 538.931, 538.935

Рецензент: зав. кафедрой фотоники, к.ф.м.н. Чижов Ю.В.

Печатается по решению методической комиссии физического факультета. СПбГУ

Рекомендовано ученым советом ФУНЦ. СПбГУ

## **Структура кристаллов и электронные состояния. – СПб., 2007**

Изложены основные сведения, необходимые для начального ознакомления со строением и свойствами полупроводников, от структуры кристаллов до свойств электронного ансамбля в трех- и двумерных системах.

Пособие предназначено для студентов 4-5-го курсов кафедры фотоники.

## СОДЕРЖАНИЕ

<b>2.1. Электронные состояния в полупроводниках</b>	<b>6</b>
2.1.1. Собственные состояния. Законы дисперсии	6
2.1.2. Энергетическая диаграмма полупроводника	12
2.1.3. Дефекты кристалла. Локальные состояния	14
2.1.3.1. Дефекты объема	14
2.1.3.2. Поверхностные состояния	24
2.1.3.3. Функции Ванье	27
2.1.3.4. Метод эффективной массы. Огибающая волновая функция	29
2.1.4. Дефекты в низкоразмерных системах	35
2.1.5. Экситоны	37
2.1.5.1. Экситоны Ванье-Мотта	39
2.1.5.2. Экситоны в квантовых ямах	43
2.1.6. Заключительные замечания	44
<b>2.2. Статистика носителей заряда в полупроводниках</b>	<b>50</b>
2.2.1. Статистика электронов в собственном полупроводнике	53
2.2.1.1. Концентрация носителей	53
2.2.1.2. Средние энергия, скорость и волновой вектор носителей	55
2.2.1.3. Положение уровня Ферми в собственном полупроводнике	56
2.2.1.4. Закон действующих масс	57
2.2.1.5. Эффективная масса плотности состояний	58
2.2.2. Статистика электронов в примесном полупроводнике	59
2.2.2.1. Концентрация зарядов на простых локальных центрах	60
2.2.2.2. Полупроводник с одним типом примеси	61
2.2.2.3. Частично компенсированные полупроводники.	64
2.2.2.4. Полупроводник с многозарядными примесными центрами	67
2.2.3. Статистика электронов в квантовых ямах	70
<b>2.3. Кинетические процессы в полупроводниках</b>	<b>72</b>
2.3.1. Электропроводность	72
2.3.1.1. Слабая локализация	74
2.3.1.2. Отрицательное дифференциальное сопротивление в сильных полях. Диод Ганна	78

2.3.2. <i>Носители тока в магнитном поле</i>	81
2.3.2.1. <i>Циклотронный (диаманитный) резонанс</i>	81
2.3.2.2. <i>Экстремальные орбиты</i>	83
2.3.2.3. <i>Электроны в скрещенных полях</i>	84
2.3.2.4. <i>Эффект Холла</i>	87
2.3.2.5. <i>Магнетосопротивление</i>	90
2.3.2.6. <i>Эффект де Гааза-ван Альфена</i>	91
2.3.2.7. <i>Двумерные структуры. Квантовый эффект Холла</i>	94
<b>2.4. Явления на контактах</b>	<b>101</b>
2.4.1. <i>Контактная разность потенциалов, загиб зон</i>	101
2.4.1.1. <i>Измерение КРП</i>	102
2.4.1.2. <i>Приповерхностный загиб зон на контакте с металлом</i>	103
2.4.1.3. <i>Соотношение Эйнштейна</i>	105
2.4.1.4. <i>Условия равновесия контактирующих тел</i>	107
2.4.2. <i>Слой объемного заряда</i>	108
2.4.2.1. <i>Дебаевская длина экранирования</i>	111
2.4.2.2. <i>Перезарядка примесей в слое объемного заряда</i>	112
2.4.2.3. <i>Обогащенный слой</i>	114
2.4.2.4. <i>Обедненный (истоженный) слой</i>	116
2.4.2.5. <i>p-n-переход</i>	117
2.4.2.6. <i>Емкость p-n-перехода</i>	120
2.4.3. <i>Прохождение токов через контакт</i>	121
2.4.3.1. <i>Антизапорный слой. Токи, ограниченные пространственным зарядом</i>	121
2.4.3.2. <i>Ток через запорный контакт</i>	129
2.4.3.3. <i>Эффект Шоттки</i>	134
2.4.3.4. <i>Туннелирование, туннельный диод</i>	135
<b>2.5. Неравновесные носители</b>	<b>139</b>
2.5.1. <i>Время жизни носителей заряда</i>	140
2.5.2. <i>Уравнение непрерывности</i>	143
2.5.3. <i>Фотопроводимость</i>	145
2.5.4. <i>Квазиуровни Ферми</i>	150
2.5.5. <i>Инжекция в p-n-переходе, биполярный транзистор</i>	152
2.5.6. <i>Нейтральные пакеты неравновесных носителей</i>	154
2.5.6.1. <i>Амбиполярная диффузия</i>	154
2.5.6.2. <i>Амбиполярный дрейф</i>	157
2.5.6.3. <i>Общий случай</i>	158

2.5.6.4.	<i>Длины диффузии и дрейфа</i>	159
2.5.7.	<b>Рекомбинация</b>	162
2.5.7.1.	<i>Рекомбинация зона-зона</i>	164
2.5.7.2.	<i>Излучательные межзонные переходы</i>	167
2.5.7.3.	<i>Рекомбинация через локальные центры</i>	170
2.5.7.4.	<i>Время жизни при монополярном возбуждении</i>	173
2.5.7.5.	<i>Времена жизни при биполярном возбуждении</i>	175
2.5.7.6.	<i>Стационарные состояния</i>	177
2.6.	<b>Кинетическая теория Больцмана</b>	181
2.6.1.	<i>Кинетическое уравнение</i>	182
2.6.2.	<i>Состояние термодинамического равновесия.</i>	
	<i>Принцип детального равновесия.</i>	187
2.6.3.	<i>Малые отклонения от равновесия</i>	189
2.6.4.	<b>Кинетические эффекты</b>	191
2.6.4.1.	<i>Носители в слабом постоянном и однородном электрическом поле.</i>	191
2.6.4.2.	<i>Носители заряда в слабом постоянном и однородном температурном поле.</i>	193
2.6.4.3.	<i>Выделение тепла. Эффект Томсона</i>	195
2.6.4.4.	<i>Выделение тепла на контакте. Эффект Пельтье</i>	196
2.6.4.5.	<i>Полупроводник в электрическом и магнитном полях</i>	198
2.6.5.	<i>Общая форма неравновесной функции распределения</i>	199
2.6.6.	<i>Интеграл столкновений для упругого рассеяния в изотропном кристалле</i>	200
2.6.7.	<i>Стационарные решения кинетического уравнения для электронного ансамбля</i>	204
2.6.7.1.	<i>Статическая электропроводность</i>	204
2.6.7.2.	<i>Действие температурного поля</i>	208
2.6.7.3.	<i>Совместное действие электрического и магнитного полей</i>	213
2.6.7.4.	<i>Действие слабого переменного электрического поля</i>	218
	<b>Литература</b>	223

## 2.1. ЭЛЕКТРОННЫЕ СОСТОЯНИЯ В ПОЛУПРОВОДНИКАХ

### 2.1.1. Собственные состояния. Законы дисперсии

Проведенное рассмотрение, см. [1], диаграммы типа изображенных на рис. 1.6.4, и 1.6.5 и выражение (1.6.35) могут пояснить только основные черты энергетической структуры электронных состояний кристалла. А именно: наличие запрещенных энергетических интервалов, квазинепрерывность разрешенных зон, квадратичную зависимость энергии от квазиимпульса близ экстремумов зон и симметрию закона дисперсии,  $E(\mathbf{k}) = E(-\mathbf{k})$ . Реальные законы дисперсии (см., например, рис. 1.6.7) оказываются много сложнее изображенных на рис. 1.6.5 из-за наличия нескольких (возможно, разных) атомов в элементарной ячейке, вырождения исходных атомных термов, взаимодействия с атомами 2-й, 3-й и так далее координационных сфер, анизотропии кристалла...

Всегда следует учитывать две существенные особенности законов дисперсии, принципиальные для определения свойств электронного ансамбля и условий возбуждения электронов:

– Относительное расположение экстремумов  $v$ - и  $c$ -зон в  $\mathbf{k}$ -пространстве. По этому признаку полупроводники разделяются на "прямозонные" и "многодолинные".

– Форма поверхностей постоянной энергии (в  $\mathbf{k}$ -пространстве), их анизотропия близ экстремумов зон даже в кубических изотропных кристаллах. Анизотропия – следствие различий кривизны дисперсионных кривых для разных направлений, т.е. и различий эффективных масс, см. (1.6.11).

Поясним суть этих особенностей.

Такие распространенные полупроводники как Ge и Si имеют решетку алмаза. Многие соединения  $A^{II}B^{VI}$  и  $A^{III}B^V$  имеют подобные ей структуры, называемые структурой цинковой обманки, рис. 1.2.14, 1.2.15. И ту, и другую можно представить себе как

две вложенные друг в друга гранцентрированные кубические (ГЦК) решетки. Только у алмаза, Ge и Si обе решетки одинаковы, а в соединениях<sup>1</sup>  $A^{II}B^{VI}$  и  $A^{III}B^V$  они состоят из разных атомов.

На рис. 2.1.1 изображены пространственная структура алмаза, форма зоны Бриллюэна кристаллов со структурой алмаза и цинковой обманки и дисперсионные кривые для Si в направлениях  $X$  (на рисунке – вправо) и  $L$ . В зоне Бриллюэна (рис. 2.1.1б) показаны также изоэнергетические поверхности для электронов близ минимума  $c$ -зоны в Si. В Si они находятся в точках  $\Delta$ , на  $x$ -оси. На энергетической диаграмме рис. 2.1.1в они отмечены символом  $\Delta'_2$ .

Таких точек в зоне Бриллюэна шесть: три идентичных оси, по два экстремума на каждой, при положительных и отрицательных<sup>2</sup> значениях  $k$ . Поскольку вероятность заселения состояний (при заданных химическом потенциале и температуре) определяется только их энергией, то вероятности заселения всех этих шести минимумов одинаковы.

Области близ таких экстремумов будем называть *долинами*. Если долина находится не в центре зоны Бриллюэна, то идентичных долин обязательно будет несколько, по крайней мере – две. Следовательно, полупроводник с такой электронной структурой логично назвать **МНОГОДОЛИННЫМ**.

Изоэнергетическая поверхность электронов близ экстремумов (рис. 2.1.1б) представляет собой 6 эллипсоидов вращения. Их большие оси направлены вдоль главных осей кристалла, а в пер-

---

<sup>1</sup> Запись  $A^{II}B^{VI}$  или  $A^{III}B^V$  применяют для общего обозначения двухкомпонентных соединений из элементов соответственно 2-го и 6-го или 3-го и 5-го столбцов таблицы Менделеева.

<sup>2</sup> Не нужно забывать, что от знака квазиволнового вектора энергия не зависит, вся картина зоны Бриллюэна симметрична относительно  $k = 0$ .



в том, какой из этих минимумов окажется глобальным. В Si глобальным оказался минимум в точке  $\Delta$ , в Ge – в точке  $L$ . Это различие, в общем, не принципиально. Принципиальным является тот факт, что в обоих кристаллах близ порога возбуждения из  $v$ - в  $c$ -зону электронные переходы должны происходить с существенным изменением  $\mathbf{k}$ -вектора. Для термического возбуждения это не существенно, а для оптического означает, что фотон не может обеспечить такой переход, поскольку его импульс мал. Для получения необходимого импульса нужна третья частица, фотон, который несет незначительную энергию, но при этом может иметь любой импульс.

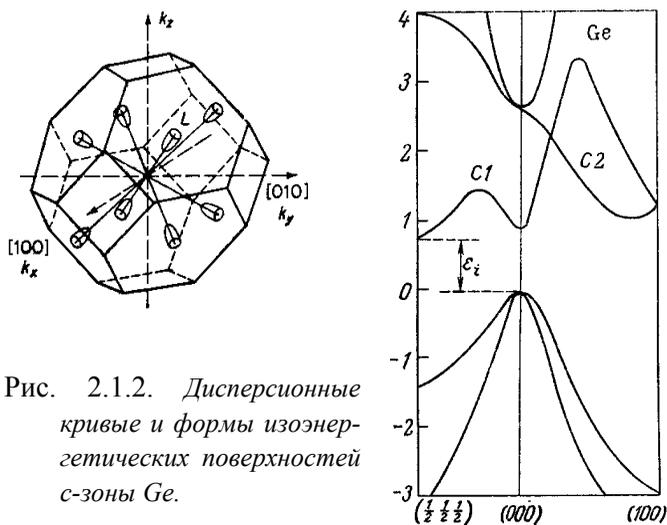


Рис. 2.1.2. Дисперсионные кривые и формы изоэнергетических поверхностей  $c$ -зоны Ge.

Такие оптические переходы называются **непрямыми**. Электрон не может "прямо" перейти из  $v$ - в  $c$ -зону, нужно участие третьей частицы для согласования импульсов.

**В многодолинных полупроводниках длинноволновый край поглощения обусловлен непрямыми переходами.**

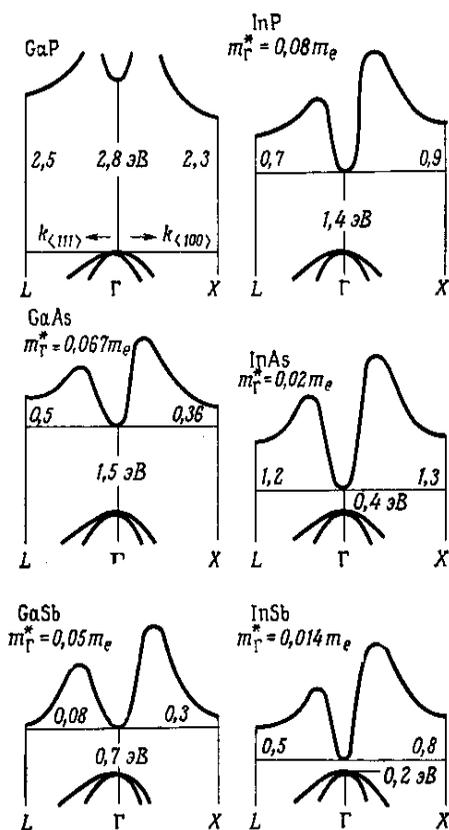


Рис. 2.1.3. Дисперсионные кривые ряда соединений типа  $A^{III}B^V$ .

Законы дисперсии с-зоны в структурах цинковой обманки имеют те же основные черты, рис. 2.1.3. Так же имеется по три локальных минимума. Но у них в большинстве случаев, за исключением GaP, глобальным является минимум в точке  $\Gamma$ , при  $k=0$ , причем минимум довольно резкий. Эффективные массы оказываются очень малы, в  $16 \div 70$  раз меньше массы свободного электрона.

Такие полупроводники будем называть **прямозонными**. Поскольку максимум v-зоны находится в точке  $\Gamma$ , то в них оптические переходы на краю поглощения — **прямые**. Переход не требует изменения импульса и происходит без участия фононов. Соответственно, вероятность таких переходов выше и существенно отличается форма спектра поглощения близ длинноволнового порога.

Во всех представленных здесь веществах при некотором превышении пороговой энергии поглощения "включаются" переходы разных типов. Если первый переход не прямой, как в Ge и

Si, то начало прямого поглощения, на расстоянии примерно  $0,2 \div 0,3$  эВ от порога, существенно сказывается на форме спектра. Если же первый переход прямой, как в GaAs, то добавление непрямых, менее интенсивных, сказывается уже не столь существенно.

Итак, возможны два принципиально различных типа зонной структуры полупроводников. Положения экстремумов зон либо совпадают в  $\mathbf{k}$ -пространстве, либо нет. Первые называются *прямозонными*, вторые *многодолинными*.

В многодолинных полупроводниках поверхности постоянной энергии с-зоны – эллипсоиды, т.е. эффективные массы для движения электронов близ дна с-зоны в различных направлениях различаются. В Si – примерно в пять раз, в Ge – почти в двадцать!

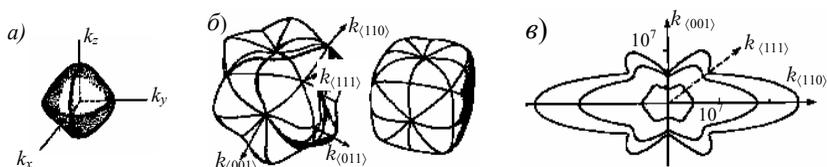


Рис. 2.1.4. *Изоэнергетические поверхности в  $v$ -зоне Si.*

*a* –  $v$ -зона легких дырок; *б* –  $v$ -зона тяжелых дырок, слева — для большей, справа — для меньшей энергии; *в* – сечения поверхностей тяжелых дырок плоскостью, содержащей ось  $z$  и диагональ  $xy$ .

Экстремумы валентных зон расположены в центре зоны Бриллюэна. Но они здесь вырождены и при отклонении от центра, при увеличении энергии дырок (на рисунках отсчитывается вниз от экстремума) расщепляются на две, существенно различающихся по крутизне зависимости  $E(\mathbf{p})$ , т.е. и по эффективной массе, и по плотности состояний. В более "плоской" зоне больше

плотность состояний и концентрация дырок, но у них больше эффективная масса и меньше подвижность.

Следует раздельно рассматривать зоны легких (с более крутым законом дисперсии) и тяжелых дырок. Их эффективные массы различаются в Si – втрое, в Ge – почти в семь раз. Формы изоэнергетических поверхностей Si представлены на рис. 2.1.1.

### **2.1.2. Энергетическая диаграмма полупроводника**

Определившись в параметрах собственных состояний, мы можем в дальнейшем не возвращаться лишней раз к законам дисперсии, а использовать усредненные характеристики при расчетах энергетического распределения и параметров носителей. В первую очередь нам нужно будет знать величины приведенных масс (и/или подвижность) и плотность состояний. Но, кроме того, при наличии любого внешнего воздействия обычно важно знать и отображать его распределение по глубине образца<sup>1</sup>. Поэтому энергетическую структуру будет удобнее изображать так, как показано на рис. 2.1.5.

Здесь показаны 3 варианта отображения основных параметров структуры. Во всех случаях для нас будут существенны только состояния близ потолка валентной зоны и дна зоны проводимости. Отображение зависимости  $E(k)$  показывает относительное расположение экстремумов  $v$ - и  $c$ -зон, а также – структуры зон. В данном случае  $v$ -зона состоит из двух подзон с разными эффективными массами. Для обеих зон указаны только состояния близ экстремумов.

В большинстве случаев придется рассматривать тепловое возбуждение носителей или оптические переходы близ края поглощения, так что зависимость  $E(k)$  будет не принципиальна, но

---

<sup>1</sup> В первую очередь будет интересоваться глубина проникновения электрического поля.

принципиален ход потенциала в пространстве. Тогда удобна схема рис. 2.1.5б, где на зависимости энергии состояний от координаты показана поверхность кристалла и основные энергетические параметры:

– потенциальная энергия электронов в вакууме принята за ноль,

–  $E_c$  и  $E_v$  – энергии краев зон,

–  $F$  – уровень Ферми,

–  $E_g$  – ширина запрещенной зоны,

–  $A$  – величина сродства кристалла к электрону,

–  $\Phi_{ph}$  и  $\Phi_T$  – фото- и термоэлектрическая работы выхода.

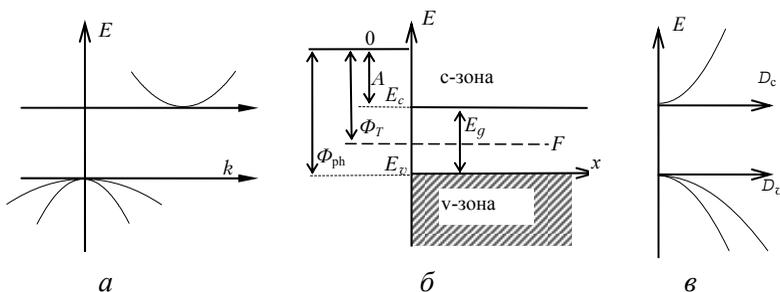


Рис. 2.1.5. Упрощенное изображение энергетической диаграммы полупроводника: а – в  $k$ -пространстве; б – в координатном пространстве; в – зависимость плотности состояний в зонах  $D_c$  и  $D_v$  от энергии.

В общем случае схему рис. 2.1.5б следует дополнить уровнями дефектных состояний в запрещенной зоне.

На рис. 2.1.5в по оси ординат отображена та же шкала энергий электронов, но по оси абсцисс – плотности состояний.

### 2.1.3. Дефекты кристалла. Локальные состояния

#### 2.1.3.1. Дефекты объема

Кристаллов без дефектов не бывает и не может быть. Представим себе идеальный кристалл (рис. 2.1.6а), находящийся в термодинамическом равновесии с внешней средой при температуре  $T$ . Его внутренняя энергия равна  $U_0$ . Энтропия идеального кристалла  $S = k \ln \Omega = 0$ , так как кратность вырождения такого состояния  $\Omega = 1$ . Реализовать идеальный кристалл можно только одним единственным способом. Получается, что в этом случае свободная энергия  $\mathcal{F} = U_0 - TS = U_0$  равна внутренней.

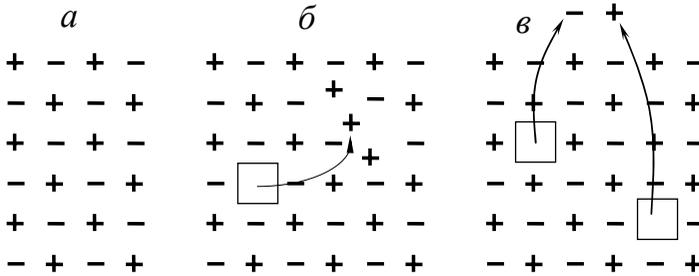


Рис. 2.1.6. Типы собственных дефектов структуры в ионной решетке:

*а* – регулярный кристалл;

*б* – дефект Френкеля, ионная вакансия и ион в междуузлии;

*в* – дефект Шоттки, пара разноименных ионных вакансий, ионы смещены на поверхность и достраивают кристалл.

При конечной температуре всегда есть конечная вероятность смещения атома (иона) из узла в междуузлие. Такой дефект называется дефектом Френкеля, рис. 2.1.6б. При этом внутренняя энергия несколько увеличится (на  $W$ ), но возрастет и энтропия. Если кристалл содержит  $N$  идентичных ионов и  $N_i$  идентичных

междуузельных положений, то свободная энергия кристалла с одним дефектом Френкеля будет равна:

$$\mathcal{F} = U_0 + W - kT \ln(NN_i),$$

а с  $n$  энергетически одинаковыми дефектами:

$$\mathcal{F} = U_0 + nW - kT \ln \left\{ \binom{N}{n} \binom{N_i}{n} \right\}, \quad (2.1.1)$$

где  $\binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{N(N-1)(N-2)\dots(N-n+1)}{1 \cdot 2 \cdot 3 \cdot \dots \cdot n}$ .

Предположив, что количество дефектов  $n$  в кристалле макроскопических размеров может быть много больше единицы, но много меньше  $N$  и  $N_i$ , т.е.  $N, N_i \gg n \gg 1$ , мы сможем для вычисления факториалов в (2.1.1) использовать формулу Стирлинга:

$$n! \cong n^n \exp(-n) \sqrt{2\pi n}$$

и найти равновесную концентрацию дефектов из условия минимума свободной энергии. Приравняв нулю производную от  $\mathcal{F}$  по  $n$ , получим:

$$n = \sqrt{NN_i} \exp\left(-\frac{W}{2kT}\right), \quad (2.1.2)$$

т.е. реально в кристалле при конечной температуре дефекты структуры будут всегда. Величина  $W$  обычно составляет несколько электронвольт, так что при комнатной температуре это количество невелико. Но быстро растет с температурой.

Типы собственных дефектов могут быть весьма разнообразны, все очень зависит от структуры кристалла и доминирующего типа связи. Иногда образуются парные точечные дефекты, например, анионная и катионная вакансии. Практически всегда имеются химические примеси, т.е. одиночные точечные дефекты. В рассматриваемых нами атомных кристаллах это – инородные

атомы, либо нарушения *стехиометрии*, правильного соотношения атомных концентраций, требуемого химической формулой вещества. Примесные атомы могут быть по-разному встроены в решетку и, соответственно, иметь разные свойства, разные спектры. Следует различать *примеси замещения* и *примеси внедрения* (1 и 3 на рис. 2.1.7). Обычно наличие химической примеси вызывает и деформацию основного кристалла близ инородного атома, см. область 2 на рис. 2.1.7.

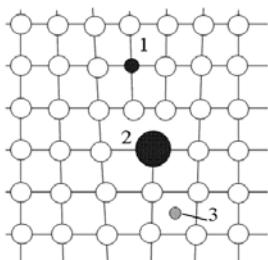


Рис. 2.1.7. Три типа точечных дефектов.

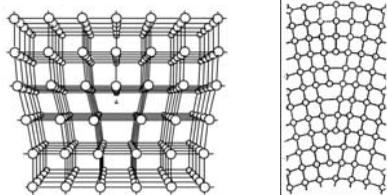


Рис. 2.1.8. Краевая дислокация и плоскость излома.

Иногда нарушается закон *расположения* атомов в пространстве, возникают протяженные одно- или двумерные дефекты упаковки, называемые *дислокациями*, *плоскостями излома*, *спайности* и т.д. На рис. 2.1.8 изображены разрезы, перпендикулярные оси краевой дислокации (слева, в кристалл "вставлена лишняя плоскость") и плоскости излома. Подробнее с этим вопросом можно ознакомиться по учебникам (например, [2]) и специальным монографиям.

Эксперимент показывает, что наличие дефектов самым существенным образом сказывается и на концентрациях носителей заряда, и на их подвижности, и на эффективности генерации неравновесных носителей при внешнем возбуждении (например, фотовозбуждении), и на времени установления равновесия при изменении воздействия. Словом, на всех параметрах электронно-

го ансамбля. Поэтому введение дефектов – основной технологический прием, используемый в производстве полупроводниковых кристаллов с заданными свойствами. Точнее – введение химических примесей, поскольку, во-первых, их концентрацией проще управлять, чем концентрацией структурных дефектов, а во-вторых, шире возможности вариации свойств дефектов за счет выбора нужных добавок. Процесс введения примесей называется *легированием*, а сами примеси – *лигандами*.

При слишком большой концентрации, при которой атомы примеси будут взаимодействовать друг с другом, они могут образовать собственную, *ПРИМЕСНУЮ ЗОНУ*, могут повлиять на пространственную структуру основного вещества. Но мы ограничимся случаями достаточно малых концентраций. Будем считать к тому же, что места расположения примесей достаточно далеки друг от друга. Достаточно для того, чтобы можно было считать основной объем кристалла не деформированным и не возмущенным, а примеси – не взаимодействующими.

Не вдаваясь в полное описание природы и энергетики дефектов, не пытаясь как-то увязывать их реальную природу и свойства, мы будем рассматривать только общий характер их взаимодействия с электронным ансамблем кристалла и влияния на его свойства. Это влияние достаточно полно описывается такими параметрами, как энергия (потенциал) ионизации дефекта и сечения рассеяния и захвата носителей.

Не будем останавливаться и на различиях между химическими примесями и дефектами упаковки, а термин "примесь", которым будем пользоваться далее в этом описании, может означать что угодно, любой точечный дефект. При необходимости его можно и конкретизировать.

Искажения, которые привносятся примесью в картину идеального кристалла, можно свести к трем основным факторам.

**Во-первых**, в результате внедрения инородного атома в его окрестности возникают искажения в ходе потенциала, вызванные и деформацией решетки, и различием собственных полей разных атомов. Разность между реальным полем в окрестности дефекта и полем в идеальном кристалле назовем дополнительным, возмущающим потенциалом. В этом дополнительном поле могут существовать собственные электронные состояния. Их энергии и типы симметрии орбиталей определяются параметрами этого поля и в общем случае могут быть самыми различными, но главное, что это – некая особая система термов, орбитали, существующие в центральном, а не периодическом потенциале. Если их энергии не соответствуют разрешенным зонам кристалла, то захваченные на них носители оказываются *локализованными* в пространстве.

**Во-вторых**, как правило, нарушается исходный баланс между количеством обобществленных, зонных электронных состояний и полным количеством валентных электронов. Если у замещающего атома валентных электронов больше, чем у атомов основного вещества, то "лишним" нет места в  $v$ -зоне, они остаются локализованными на примеси, но могут быть возбуждены в  $v$ -зону с меньшей затратой энергии, чем электроны валентной зоны. Если электронов недостаточно, то облегчается возможность образования дырки в  $v$ -зоне. При изоэлектронном замещении, если примесь к тому же не деформирует кристалл, это не так, но и здесь оказываются существенными различия в форме атомных потенциалов. Например, примесь Al в GaAs существенно влияет на ширину запрещенной зоны. У GaAs она равна  $\sim 1.5$  эВ (зависит от температуры), у AlAs –  $\sim 2.2$  эВ. Варируя соотношение Ga и Al, можно изменять ширину зоны смешанного кристалла  $\text{Ga}_x\text{Al}_{1-x}\text{As}$  в этом диапазоне.

**В третьих**, в поле примеси возмущены и собственные зонные состояния, на ее потенциале носители заряда должны рассеиваться. Тем самым, наличие примесей уменьшает время релаксации импульса  $\tau_p$ , а с ним и подвижность носителей  $\mu$ .

Последний фактор мы можем пока не рассматривать. При малых концентрациях примеси и высоких температурах, порядка комнатной, рассеяние носителей более эффективно происходит на *фононах*, колебательных возбуждениях кристалла.

Первые два сводятся к тому, что примеси должны рассматриваться как самостоятельная электронная система, построенная на взаимодействиях, принципиально отличных от взаимодействий в ансамбле собственных электронов. К картине обобществленных электронных состояний кристалла нужно добавить теперь описание локализованных состояний примесей и описание законов взаимодействия этих двух систем, т.е. их обмена зарядом и энергией между ними.

Представим себе, что в идеальном кристалле 4-валентного Ge один из атомов мы заменили на элемент 5-й группы, например, As. Получили примесь замещения. На "встраивание" в решетку и восстановление валентных связей в структуре Ge, ему понадобится 4 валентных электрона, а один останется "не у дел". Для этого электрона нет места в валентной зоне, а в зону проводимости он может попасть только при получении дополнительной энергии. Ведь около "своего" атома As он находится в поле единичного избыточного положительного заряда ядра. Примесь такого типа называется *донорной*. Она "поставляет" электроны в собственную зону полупроводника и может находиться в зарядовом состоянии (0) или (+1).

Оценить энергию, необходимую для ионизации такого центра просто, если мы знаем, как ведет себя электрон в центральном поле (задача об атоме водорода). Электрон в поле протона имеет стационарные состояния с энергиями связи:

$$E_n \approx \frac{1}{n^2} \frac{m_0 \cdot e^4}{2\hbar^2} = \frac{Ry}{n^2}, \quad (2.1.3)$$

$n = 1, 2, 3, \dots$ ;  $Ry = 13,6$  эВ.

Средний радиус орбиты в основном состоянии,  $n = 1$ , равен:

$$a_B = \frac{\hbar^2}{m_0 \cdot e^2} = 0,0529 \text{ нм.} \quad (2.1.4)$$

В нашем случае заряды находятся не в свободном пространстве, а в среде, поляризация которой экранирует заряды и уменьшает энергию их притяжения. Если заранее предположить, что радиус орбиты электрона много больше межатомного расстояния, то можно не рассматривать реальной пространственной структуры кристалла. Можно считать его *квазинепрерывной* средой и описать эти взаимодействия просто диэлектрической постоянной  $\epsilon$ . Тогда выражения (2.1.3), (2.1.4) сохраняются, но  $e^2$  в них следует заменить на  $e^2/\epsilon$ . Кроме того, движение электрона в периодическом поле кристалла характеризуется эффективной массой, не равной массе свободного электрона  $m_0$ . Тогда для потенциала ионизации основного состояния примесного центра получаем:

$$I_d \cong Ry \frac{m^*}{m_0} \frac{1}{\epsilon^2} \quad (2.1.5)$$

У Ge  $\epsilon \cong 16$  и  $m^* \cong 0,2m_0$ , так что:

$$I_d \cong 13,6 \cdot 0,2 \cdot 16^{-2} \cong 0,01 \text{ эВ,}$$

а радиус первой орбитали:

$$r \cong 16 a_B / 0.2 \cong 4.2 \text{ нм.} \quad (2.1.6)$$

Это много больше межатомного расстояния в Ge (0,245 нм) и можно считать оправданным наше предположение о возможности описания диэлектрических взаимодействий усредненной характеристикой квазинепрерывной среды –  $\epsilon$ .

Посмотрим на эту ситуацию несколько иначе. Был кристалл, со своей структурой, своей постоянной решетки и своим внутренним полем. Внедрением малого количества примеси мы ниче-

го этого не нарушили, т.е. все основные параметры энергетического строения остались неизменными, в том числе и количество состояний, см. рис. 2.1.9. Но появились "лишние" электроны, с очень малой энергией связи<sup>1</sup>, всего порядка 0,01 эВ. Приобретения такой энергии достаточно для того, чтобы электрон *ушел* от примеси на бесконечность. Правда, с минимально возможной для этого энергией, т.е. с энергией, соответствующей дну с-зоны. На энергетической диаграмме типа рис. 2.1.9 факт появления в кристалле таких дополнительных примесных состояний может быть отражен узким донорным уровнем, на расстоянии от с-зоны, равном  $I_d$ , см. рис. 2.1.9а. Он изображается именно *узким уровнем*, так как во-первых, кристалл считаем *однородным*, т.е. и примеси распределены в пространстве *равномерно*, и все точки расположения атомов примеси идентичны. Во-вторых, концентрацию  $N_d$  считаем достаточно малой для того, чтобы взаимодействием дефектов между собой можно было пренебречь.

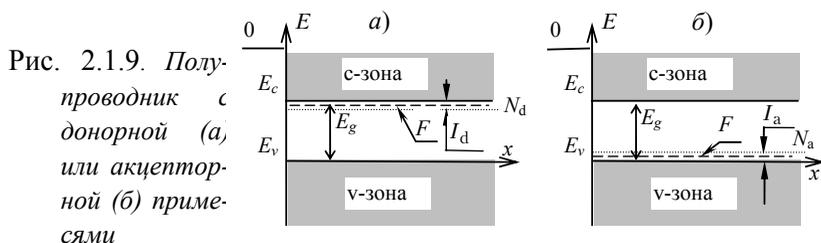


Рис. 2.1.9. Полупроводник с донорной (а), или акцепторной (б) примесями

Полностью аналогичное рассмотрение можно было бы провести для случая легирования Ge элементом 3-го столбца таблицы Менделеева, например, In. У него 3 валентных электрона, одного не хватает для образования 4-х связей и поэтому он очень легко присоединяет дополнительный электрон, захватывая его из

<sup>1</sup> Доноры и акцепторы с малой энергией связи называют "мелкими".

валентной зоны. Если немного изменим терминологию, то – "отдает дырку" в валентную зону. Дальше – все так же, как и для донорной примеси. Но в этом случае примесь называется **акцепторной**. Её зарядовые состояния – (0) или (–1). Соответствующий примесный уровень находится на малом расстоянии от валентной зоны, см. рис. 2.1.9б.

Примесь элементов других столбцов таблицы Менделеева дает *многозарядные центры*, например, медь в Ge (акцептор) может иметь заряд до (–3) и соответственно образует в запрещенной зоне три акцепторных уровня, с зарядами (–1), (–2) и (–3). Их суммарная концентрация равна концентрации атомов Cu, а реальное распределение атомов Cu по этим зарядовым состояниям зависит от наличия иных примесей и от температуры. Если атом захватил 2 электрона, то в энергетической диаграмме он должен быть отражен как донорный "(–2)" и как акцепторный "(–3)".

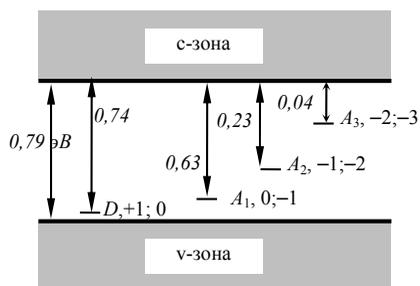


Рис. 2.1.10. Энергии электронов, локализованных на примеси Au в Ge, при  $T = 77$  К.

Отмечены: тип центра (донорный, акцепторный), величины зарядов и энергии связи электронов (в электронвольтах).

Подобный набор акцепторных уровней дают в Ge и Si все одновалентные атомы (энергии связи, естественно, различаются). Золото, кроме того, может отщеплять единственный электрон, для чего нужна энергия всего на 0.05 эВ меньше ширины запрещенной зоны в Ge (0.79 эВ при 0 К). Соответствующий донорный уровень должен изображаться на энергетической диаграмме на расстоянии 0.05 эВ от v-зоны. Таким образом, при легировании Ge золотом в запрещенной зоне возникают сразу четыре примес-

ных уровня, один донорный, с энергией связи 0.74 эВ и три акцепторных, с энергиями 0.16, 0.56 и 0.75 эВ относительно  $v$ -зоны.

На рис. 2.1.10 изображена возникающая схема уровней. Для удобства дальнейшего описания обозначены возможные зарядовые состояния и все энергии отсчитаны от дна зоны проводимости. Нужно только четко представлять себе, что реальная картина может оказаться беднее изображенной, если температура невелика и  $kT$  существенно меньше энергетического расстояния между уровнями. Каждому из возможных зарядовых состояний Au соответствуют 1–2 уровня. Ионы  $Au^+$  и  $Au^{3-}$  могут изменить свой заряд только в одну сторону, первый – принять, второй – отдать электрон. Им соответствуют уровни  $D$  (пустой) и  $A_3$  (занят). Центры Au,  $Au^-$  и  $Au^{2-}$  могут перезарядиться в обе стороны, им соответствуют один занятый и один свободный уровень.

Все энергетические характеристики примесных центров, о которых мы говорили, определяют энергии обмена электроном между центром и кристаллом, потенциалы их ионизации. В оптике полупроводников может оказаться существенным также спектр внутренних возбуждений примеси, проявляющийся и в поглощении, и в люминесценции. Нас интересует влияние примесей на проводимость, так что наличие возбужденных состояний представляет интерес только в том плане, что они должны учитываться в статистике электронов. Здесь нужно будет учесть полное количество различных квантовых состояний, в которых может существовать интересующий нас центр в рассматриваемом зарядовом состоянии.

В разделе 2.2.2.1 мы введем некоторый эффективный статистический вес, учитывающий и количество возможных возбужденных состояний, и их энергии. Причем, всех состояний, независимо от того, возможен ли между ними оптический переход. Этого будет достаточно для описания *термически равновесных*

распределений электронов, в том числе – концентраций носителей и заряженных примесных центров.

### 2.1.3.2. Поверхностные состояния

Поверхность полупроводника – очень сложная система, особенно если она не защищена от воздействия внешней атмосферы. Многие полупроводники являются эффективными катализаторами химических реакций. Это означает, что на их поверхности при адсорбции относительно инертных молекул могут образовываться химически активные комплексы, т.е. ионы, радикалы, полярные молекулы. Если учесть к тому же, что каждый поверхностный атом (их концентрация  $\sim 10^{15} \text{ см}^{-2}$ ) может стать центром адсорбции и что состав активных примесей в комнатной атмосфере изменяется очень сильно, то напрашивается простая, но неприятный вывод: поверхность может представлять собой просто *иную* электронную систему с плохо предсказуемыми свойствами.

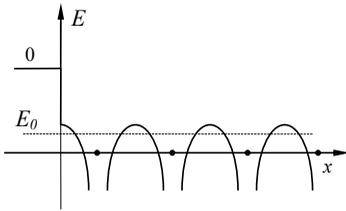


Рис. 2.1.11. *Ход потенциала близ поверхности раздела фаз вакуум/кристалл.*

На поверхности легко могут быть образованы и локальные центры с почти любыми донорно-акцепторными свойствами, и двумерные, поверхностные энергетические зоны. Накапливая избыточный заряд, поверхность может иметь потенциал, отличный от потенциала объема. В этом случае в объеме полупроводника возникает электрическое поле, влияющее в том числе – на объемную проводимость в приповерхностном слое.

Так что, работая с полупроводниками всегда нужно тщательно контролировать состояние поверхности. Мы же будем по-

лагать, что кристалл находится в идеальном вакууме и никаких активных адсорбционных комплексов нет. Тем не менее, специфические состояния есть и на чистой поверхности, причем не только на особых точках, неровностях, но и на атомно гладкой поверхности. Они возникают как следствие скачка потенциала на границе вакуум/кристалл.

Рассмотрим состояния электронов, движущихся по нормали к атомно гладкой поверхности. На рис. 2.1.11 ось  $x$  направлена вглубь кристалла. Ноль – на поверхности. Энергию будем отсчитывать от уровня потенциальной энергии электрона в вакууме. Нас будут интересовать только связанные состояния, с энергией  $E_0 < 0$ .

В левом полупространстве – вакуум, здесь кинетическая энергия отрицательна и волновой вектор электрона  $k = i \frac{\sqrt{2m|E_0|}}{\hbar}$  – величина чисто мнимая. В правом полупространстве волновой вектор либо действительный, если энергии  $E_0$  соответствует состояние в какой-либо разрешенной зоне, либо комплексный,  $k = k_1 + ik_2$ , если уровень  $E_0$  лежит в запрещенной зоне.

Волновую функцию поверхностного состояния мы можем построить, сшив при  $x = 0$  волновые функции падающих и отраженных волн. В вакууме, при  $x < 0$ , это плоские волны де Бройля:

$$\psi_1(x) = \alpha \exp\left(\frac{\sqrt{2m|E_0|}}{\hbar} x\right) \quad (2.1.7)$$

а в кристалле – блоховские функции:

$$\psi_2(x) = \beta_1 u_k \exp(ikx) + \beta_2 u_{-k} \exp(-ikx). \quad (2.1.8)$$

В (2.1.7) мы оставили только один член, так как второй (со знаком минус в показателе степени) устремится к бесконечности при  $x \rightarrow -\infty$ .

Коэффициенты  $\alpha$ ,  $\beta_1$  и  $\beta_2$  должны быть определены из условия неразрывности при  $x=0$  и самой волновой функции, и ее производной по координате:

$$\psi_1(0) = \psi_2(0) ; \quad \left. \frac{\partial \psi_1(x)}{\partial x} \right|_{x=0} = \left. \frac{\partial \psi_2(x)}{\partial x} \right|_{x=0} \quad (2.1.9)$$

В (2.1.8) при вещественном  $k$ , т.е. если энергия  $E_0$  оказывается в области разрешенных состояний, мы должны оставить оба члена. Оба они реальны. Получилось, что в двух уравнениях (2.1.9) имеется три неизвестных  $\alpha$ ,  $\beta_1$  и  $\beta_2$ , т.е. решения есть всегда. Следовательно, **на поверхности разрешены все электронные состояния, разрешенные в объеме.**

Если же  $k = k_1 + ik_2$ , то в (2.1.8) мы должны положить  $\beta_2 = 0$ , чтобы волновая функция не обратилась в бесконечность при  $x \rightarrow +\infty$  (считаем, что  $k_2 > 0$ ). Остается два уравнения с двумя неизвестными. Их решение единственно:

$$E = -\frac{\hbar^2}{2m} \left[ \frac{u'_k(0)}{u_k(0)} + ik \right]^2, \quad k = k_1 + ik_2. \quad (2.1.10)$$

Волновая функция (и пропорциональная ее квадрату вероятность обнаружения электрона с такой энергией) отлична от нуля при  $x=0$ , но экспоненциально затухает с расстоянием в обе стороны. Т.е. состояние действительно локализовано на поверхности, см. рис. 2.1.12. И.Е.Тамм показал, что такие состояния (их называют **уровни Тамма**) имеются в каждой запрещенной зоне, на каждой поверхности. Их плотность – порядка концентрации поверхностных атомов. Только не всегда они попадают в область,

которая запрещена для всех направлений в кристалле (см. разд. 1.6.1.7 и рис. 1.6.7 в [1]).

На неровностях поверхности – в углах и на ребрах ступенек, в местах выхода дислокаций и т.д., возникают свои локальные центры и обычно их спектр достаточно богат. В

некоторых случаях экспериментально наблюдаемые свойства удастся описать только в предположении, что энергетический спектр поверхностных состояний непрерывен, электроны на поверхности могут иметь любые энергии, остается только определить распределение плотности их состояний. При этом каталитическая активность поверхности обеспечивается обычно центрами, концентрация которых составляет  $10^{-4}$  монослоя и меньше. Остальные центры хоть и не проявляют химической активности, но играют очень существенную роль в перераспределении носителей заряда по состояниям.

### 2.1.3.3. Функции Ванье

Вблизи дефектов кристалла или близ поверхности на электрон действует дополнительный потенциал, который может привести к образованию локального состояния. Например, поле мелкой примеси, дополнительное поле вблизи поверхности. Таким же дополнительным полем является кулоновское поле дырки, действующее на электрон. Взаимодействие электрона и дырки может привести к возникновению связанных состояний, называемых *экситоном*. Наконец, дополнительное поле, создаваемое

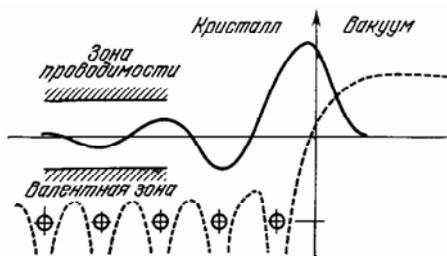


Рис. 2.1.12. Волновая функция поверхностных состояний (уровней Тамма) на границе кристалл/вакуум.

поляризацией ионного кристалла, которая вызвана действием самого электрона на решетку, приводит к так называемым *поляронным состояниям*.

Поскольку эти дополнительные поля локальны, часто оказывается целесообразно использовать в качестве нулевого приближения не функции Блоха

$$\psi_{nk}(\mathbf{r}) = u_{nk}(\mathbf{r}) \cdot \exp(i\mathbf{k}\mathbf{r}), \quad (2.1.11)$$

а функции Ванье, локализованные не в  $k$ -пространстве, а в реальном пространстве.

Поскольку функции Блоха (2.1.11) периодичны в  $k$ -пространстве с периодом векторов обратной решетки, они могут быть разложены в ряд Фурье по узлам решетки:

$$\psi_{nk}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_l \varphi_n(\mathbf{a}_l, \mathbf{r}) \cdot \exp(i\mathbf{k}\mathbf{a}_l). \quad (2.1.12)$$

Коэффициенты разложения  $\varphi_n(\mathbf{a}_l, \mathbf{r})$  ( $\mathbf{a}_l$  — вектор решетки  $l$ -го узла) и есть функции Ванье. Они осуществляют "узельное представление" электрона в идеальной решетке.

Обратное разложение имеет вид:

$$\varphi_n(\mathbf{a}_m, \mathbf{r}) = \frac{1}{\sqrt{N}} \sum_k u_{nk}(\mathbf{r}) \cdot \exp[i\mathbf{k}(\mathbf{r} - \mathbf{a}_m)]. \quad (2.1.13)$$

Поскольку  $u_{nk}(\mathbf{r}) = u_{nk}(\mathbf{r} - \mathbf{a}_m)$ , то  $\varphi_n(\mathbf{a}_m, \mathbf{r}) = \varphi_n(\mathbf{r} - \mathbf{a}_m)$ . Функции Ванье зависят не от координат взаимодействующих узлов, а только от расстояния между ними.

Можно показать (см. [3], глава V), что функции Ванье (2.1.13):

- локализованы близ своих узлов  $\mathbf{a}_m$  и с удалением от них быстро убывают, осциллируя;

- ортонормированы:

$$\int_V \varphi_n^*(\mathbf{r} - \mathbf{a}_m) \varphi_n(\mathbf{r} - \mathbf{a}_m) d^3r = \delta_{nn} \delta_{mm}. \quad (2.1.14)$$

#### 2.1.3.4. Метод эффективной массы. Огибающая волновая функция

*Методом эффективной массы* называют метод расчета влияния, оказываемого на собственные электронные состояния (на их волновые функции) возмущающим потенциалом  $V(x)$ . Он применим для рассмотрения собственных состояний близ экстремумов зон, если этот потенциал плавно изменяется в пространстве, когда можно пренебречь изменением этого потенциала в пределах элементарной ячейки. Критерий понятия "слабо изменяется":

$$a_i \frac{dV/dx}{V} \ll 1,$$

где  $a_i$  – постоянная решетки.

Критерий понятия "близ экстремумов зон":

$$E_k \ll E_g,$$

Т.е. кинетическая энергия электронов или дырок, отсчитанная от края соответствующей зоны,  $E_k$ , много меньше  $E_g$ , расстояния от рассматриваемого экстремума до ближайшей другой зоны, например – ширина запрещенной зоны.

Действие этого потенциала можно описать так же плавно изменяющимся в пространстве множителем к волновой функции собственного электронного состояния. Этот множитель, являющийся функцией координат, называют *оггибающей волновой функцией*. Такое название можно считать удачным хотя бы потому, что уравнение для нее – то же уравнение Шредингера, что и

для волновой функции собственного электронного состояния, только не с периодическим, а с некоторым, медленно изменяющимся в пространстве потенциалом. Действие периодического потенциала, приводящее изменению меры инерции носителя, может быть при этом учтено и описано заменой в уравнении Шредингера массы свободного электрона на эффективную массу электрона в невозмущенном собственном состоянии.

Мы рассмотрим этот метод применительно к описанию электронных состояний функциями Ванье. Этот подход представляется более продуктивным, т.к. поле примеси или поверхностных состояний локализовано в прямом, а не в обратном пространстве.

Итак, предположим, что кроме периодического потенциала решетки  $U(\mathbf{r})$  в некоторой области кристалла действует поле  $V(\mathbf{r})$ . Уравнение Шредингера в этом случае:

$$\left\{ -\frac{\hbar^2}{2m} \nabla^2 + [U(\mathbf{r}) + V(\mathbf{r})] \right\} \psi_i(\mathbf{r}) = E_i \psi_i(\mathbf{r}), \quad (2.1.15)$$

где  $\psi_i(\mathbf{r})$  и  $E_i$  —  $i$ -я собственная функция и соответствующая ей собственная энергия.

Разложим собственную функцию  $\psi_i(\mathbf{r})$  по полной системе ортонормированных функций Ванье (2.1.13):

$$\psi_i(\mathbf{r}) = \sum_{n'l'} f_{n'l'}^i(\mathbf{a}_{l'}) \phi_{n'l'}(\mathbf{r} - \mathbf{a}_{l'}). \quad (2.1.16)$$

Здесь  $n'$  — номер зоны,  $\mathbf{a}_{l'}$  — вектор решетки узла  $l'$ ,  $f_{n'l'}^i$  — коэффициенты разложения, которые нужно определить из (2.1.15).

Подставляя это выражение в (2.1.15), умножая слева на  $\phi_n^*(\mathbf{r} - \mathbf{a}_l)$  и интегрируя по пространству, получим:

$$\sum_{n'l'} f_n^i(\mathbf{a}_{l'}) \int \varphi_n^*(\mathbf{r} - \mathbf{a}_l) \left\{ -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) + V(\mathbf{r}) - E_i \right\} \times \quad (2.1.17)$$

$$\times \varphi_n(\mathbf{r} - \mathbf{a}_{l'}) d^3r = 0.$$

Здесь под интегралом – сумма четырех членов. Если  $V(\mathbf{r})$  изменяется в пространстве достаточно медленно,  $V(\mathbf{r}) \approx V(\mathbf{a}_{l'})$ , то интеграл от двух последних, в силу ортонормированности функций Ванье, равен:

$$\begin{aligned} & \int \varphi_n^*(\mathbf{r} - \mathbf{a}_l) [V(\mathbf{r}) - E_i] \varphi_n(\mathbf{r} - \mathbf{a}_{l'}) d^3r = \\ & = [V(\mathbf{a}_{l'}) - E_i] \int \varphi_n^*(\mathbf{r} - \mathbf{a}_l) \varphi_n(\mathbf{r} - \mathbf{a}_{l'}) d^3r = \\ & = [V(\mathbf{a}_{l'}) - E_i] \delta_{nn'} \delta_{ll'}. \end{aligned} \quad (2.1.18)$$

Тогда эта часть суммы в (2.1.17) равна:

$$\sum_{n'l'} [V(\mathbf{a}_{l'}) - E_i] f_n^i(\mathbf{a}_{l'}) \delta_{nn'} \delta_{ll'} = [V(\mathbf{a}_{l'}) - E_i] f_n^i(\mathbf{a}_{l'}). \quad (2.1.19)$$

Для вычисления первых двух членов в фигурных скобках в (2.1.17) удобно перейти к функциям Блоха (2.1.13):

$$\begin{aligned} & \sum_{n'l'} f_n^i(\mathbf{a}_{l'}) \frac{1}{N} \int \sum_k e^{i\mathbf{k}\mathbf{a}_l} \psi_{nk}^*(\mathbf{r}) \left[ -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) \right] \times \\ & \quad \times \sum_{k'} e^{-i\mathbf{k}'\mathbf{a}_{l'}} \psi_{n'k'}^*(\mathbf{r}) d^3r = \\ & = \frac{1}{N} \sum_{n'l'} f_n^i(\mathbf{a}_{l'}) \sum_{k,k'} e^{i\mathbf{k}\mathbf{a}_l - i\mathbf{k}'\mathbf{a}_{l'}} E_n(\mathbf{k}') \int \psi_{nk}^*(\mathbf{r}) \psi_{n'k'}(\mathbf{r}) d^3r = \\ & = \frac{1}{N} \sum_{n'l'} \sum_{kk'} f_n^i(\mathbf{a}_{l'}) e^{i\mathbf{k}\mathbf{a}_l - i\mathbf{k}'\mathbf{a}_{l'}} E_n(\mathbf{k}') \delta_{nn'} \delta_{kk'} = \\ & = \frac{1}{N} \sum_{l'} \sum_k f_n^i(\mathbf{a}_{l'}) e^{i\mathbf{k}(\mathbf{a}_l - \mathbf{a}_{l'})} E_n(\mathbf{k}) \end{aligned} \quad (2.1.20)$$

Здесь  $E_n(\mathbf{k}')$  - уже известное нам собственное значение гамильтониана для периодического поля кристалла, без возмущения:  $\left[ -\frac{\hbar^2}{2m} \nabla^2 + U(\mathbf{r}) \right]$ . Мы также воспользовались ортонормированностью функций Блоха.

Если заменить:  $\mathbf{a}_l - \mathbf{a}_l' = \mathbf{a}_m$ , то (2.1.20) примет вид:

$$\frac{1}{N} \sum_{mk} f_n^i(\mathbf{a}_l - \mathbf{a}_m) e^{i\mathbf{k}\mathbf{a}_m} E_n(\mathbf{k}) \quad (2.1.21)$$

Разложим  $f_n^i(\mathbf{a}_l - \mathbf{r})$  в ряд Тейлора по  $r$  близ точки  $\mathbf{a}_l$ :

$$\begin{aligned} f_n^i(\mathbf{a}_l - \mathbf{r}) &= f_n^i(\mathbf{a}_l) - \mathbf{r} \nabla f_n^i(\mathbf{a}_l) + \frac{1}{2} \mathbf{r} \mathbf{r} \nabla \nabla f_n^i(\mathbf{a}_l) - \dots = \\ &= \left[ 1 - \mathbf{r} \nabla + \frac{1}{2} \mathbf{r} \mathbf{r} \nabla \nabla - \dots \right] f_n^i(\mathbf{a}_l) = \exp(-\mathbf{r} \nabla) f_n^i(\mathbf{a}_l), \end{aligned} \quad (2.1.22)$$

где введен оператор  $\exp(-\mathbf{r} \nabla)$ . Итого, при  $\mathbf{r} = \mathbf{a}_m$  имеем:

$$f_n^i(\mathbf{a}_l - \mathbf{a}_m) = \exp(-\mathbf{a}_m \nabla) f_n^i(\mathbf{a}_l) \quad (2.1.23)$$

Энергия  $E_n(\mathbf{k}) = E_n(\mathbf{k} + \mathbf{b}_l)$ , она периодична в обратном пространстве с шагами по осям, равными элементарным векторам обратной решетки. Поэтому ее можно разложить в трехмерный ряд Фурье:

$$E_n(\mathbf{k}) = \sum_l c_l e^{i\mathbf{k}\mathbf{a}_l}. \quad (2.1.24)$$

Подставляя два последних выражения в (2.1.21), получим:

$$\frac{1}{N} \sum_{mkm'} e^{-\mathbf{a}_m \nabla} f_n^i(\mathbf{a}_l) c_m e^{i\mathbf{k}(\mathbf{a}_m - \mathbf{a}_m')} = \sum_{mm'} e^{-\mathbf{a}_m \nabla} f_n^i(\mathbf{a}_l) c_m \cdot \delta_{mm'}, =$$

$$= \left[ \sum_m c_m e^{-i\mathbf{a}_m \nabla} \right] f_n^i(\mathbf{a}_l). \quad (2.1.25)$$

Сравнивая этот результат с (2.1.24), видим, что (2.1.21) приводится к виду:

$$\frac{1}{N} \sum_{mk} f_n^i(\mathbf{a}_l - \mathbf{a}_m) e^{i\mathbf{k}\mathbf{a}_m} E_n(\mathbf{k}) = E(-i\nabla) f_n^i(\mathbf{a}_l). \quad (2.1.26)$$

Объединяя (2.1.26) и (2.1.19), получаем из уравнения Шредингера (2.1.15) дифференциальное уравнение для коэффициентов  $f_n^i(\mathbf{a}_l)$ :

$$E_n(-i\nabla) f_n^i(\mathbf{a}_l) + V(\mathbf{a}_l) f_n^i(\mathbf{a}_l) = E_i f_n^i(\mathbf{a}_l). \quad (2.1.27)$$

Можно, используя это уравнение, ввести "непрерывную" функцию, имеющую в узлах решетки значения, определяемые соотношением (2.1.27):

$$E(-i\nabla) f_n^i(\mathbf{r}) + V(\mathbf{r}) f_n^i(\mathbf{r}) = E_i f_n^i(\mathbf{r}). \quad (2.1.28)$$

Если мы рассматривали действие возмущающего потенциала на волновую функцию, соответствующую экстремуму зоны и если закон дисперсии в зоне параболический и сферически симметричен, то

$$E_n(-i\nabla) = -\frac{\hbar^2}{2m^*} \nabla^2$$

и (2.1.28) превращается в обычное уравнение Шредингера для волновой функции электрона в плавно изменяющемся потенциале  $V(\mathbf{r})$ :

$$-\frac{\hbar^2}{2m^*} \nabla^2 f_n^i(\mathbf{r}) + V(\mathbf{r}) f_n^i(\mathbf{r}) = E_i f_n^i(\mathbf{r}). \quad (2.1.29)$$

Таким образом, задача о состоянии электрона в потенциале сложной формы, периодический потенциал идеального кристалла плюс плавно изменяющийся в пространстве возмущающий потенциал  $V(\mathbf{r})$ , сведена к более простой. В ней можно учитывать только этот возмущающий потенциал, а эффект от воздействующего периодического потенциала уже учтен тем, что вместо массы свободного электрона в уравнении Шредингера фигурирует эффективная масса.

Функция  $f_n^i(\mathbf{r})$  – плавная, она модулирует функции Ванье для избранного электронного состояния. Поэтому ее называют *огibaющей или сглаженной волновой функцией*.

Этот метод сведения сложной задачи к более простой, в которой расчет производится именно для электрона, отличающегося от свободного законом дисперсии, иногда вполне адекватно описываемым эффективной массой, называется *метод эффективной массы*.

В частности, он позволяет достаточно просто описать состояния электрона в низкоразмерных структурах. Они строятся на базе обычных блоховских функций или функций Ванье, промодулированных огibaющей функцией, характеризующей пространственные ограничения.

Изложенный здесь вывод огibaющей функции и метода эффективной массы выполнен в применении к функциям Ванье (см. [3], глава V, или [13], глава 4).

Можно провести аналогичное рассмотрение применительно к функциям Блоха (см. [15], глава 4, §4). В этом случае волновая функция представляется произведением блоховской  $\psi(\mathbf{r})$  и огibaющей  $\chi(\mathbf{r})$  функций:

$$\Psi(\mathbf{r}) = \chi(\mathbf{r})\psi(\mathbf{r}). \quad (2.1.30)$$

Уравнение для огибающей  $\chi(\mathbf{r})$  получается то же, что и для огибающей функций Ванье, см. (2.1.28), (2.1.29).

#### 2.1.4. Дефекты в низкоразмерных системах

Потенциальная яма, ограничивающая движение носителей в пространстве, изменяет свойства не только свободных, но и связанных электронов. Рассмотрим влияние размерного квантования на мелкий водородоподобный примесный центр.

Пусть притягивающий центр располагается внутри квантовой ямы шириной  $a$ , в точке с координатой  $0 < z_0 < a$ . Энергии уровней  $E_i$ , создаваемых этим центром, определяются из уравнения Шредингера (в системе СГСЭ):

$$-\frac{\hbar^2}{2m} \nabla^2 \psi - \frac{e^2}{\varepsilon \sqrt{x^2 + y^2 + (z - z_0)^2}} \psi = E_i \psi. \quad (2.1.31)$$

В неограниченном пространстве это уравнение дает известное решение  $E_i = -\frac{Ry^*}{i^2} = -\frac{1}{i^2} \frac{m^* e^4}{2\varepsilon^2 \hbar^2}$ , где величину

$Ry^* = \frac{m^* e^4}{2\varepsilon^2 \hbar^2} = Ry \frac{m^*}{\varepsilon^2}$  обычно называют эффективным ридбергом.

Но при граничных условиях

$$\psi(0) = \psi(a) = 0 \quad (2.1.32)$$

оно общего решения не имеет.

Все упрощается, если ширина ямы достаточно мала, много меньше эффективного боровского радиуса:

$$a \ll a_B = \frac{\varepsilon \hbar^2}{m^* e^2}. \quad (2.1.33)$$

В этом случае в (2.1.31) потенциал можно считать двумерным, т.к.  $(x^2 + y^2) \gg (z - z_0)^2$ , и волновую функцию искать в виде  $\psi = \varphi(\rho)\chi(z)$ .

Для  $\chi(\mathbf{r})$  уравнение упрощается и получаются уже знакомые решения для одномерного квантования:  $\chi_n = \sqrt{\frac{2}{a}} \sin \frac{n\pi}{a} x$ , с собственными энергиями  $E_n = \frac{\hbar^2 \pi^2}{2m^* a^2} n^2$ , см. [1], разд. 1,7.

А собственные значения для "двумерного атома водорода" равны:

$$E_i = -\frac{Ry^*}{(i-1/2)^2}. \quad (2.1.34)$$

Итого, в области энергий ниже каждого  $n$ -го уровня размерного квантования имеется "почти ридберговская" серия состояний с полной энергией:

$$E_i = \frac{\hbar^2 \pi^2}{2m^* a^2} n^2 - \frac{m^* e^4}{2\varepsilon^2 \hbar^2} \frac{1}{(i-1/2)^2}. \quad (2.1.35)$$

Требование (2.1.32) по сути, означает, что должно выполняться требование

$$\frac{m^* e^4}{2\varepsilon^2 \hbar^2} \frac{1}{(i-1/2)^2} \ll \frac{\hbar^2 \pi^2}{2m^* a^2} n^2, \quad (2.1.36)$$

так что в (2.1.35) энергия наверняка больше нуля. Все эти состояния лежат в зоне проводимости, причем все они, за исключением нижнего квантоворазмерного состояния, с  $n = 1$ , перекрываются с подзонами нижних состояний. В результате возможен переход из локализованного состояния дефекта в делокализованное состоя-

ние в подзоне и, соответственно, их время жизни оказывается малым.

Отметим также, что энергия связи этой серии в четыре раза больше, чем в трехмерной системе. Однако, реально эта энергия зависит от ширины ямы  $a$ . При малом  $a$  велик шаг квантования по  $z$  – координате. Скачок потенциала на границе становится мал и не выполняется сделанное предположение о граничных условиях (2.1.32). С удалением от мелкой ямы хвосты волновой функции спадают медленно и может оказаться, что они превышают Боровский радиус  $a_B$ . Тогда задача опять сводится к трехмерной.

При увеличении  $a$  существенно становится разница эффективных масс и диэлектрических проницаемостей в яме и в барьере. При  $a \rightarrow 0$  оказываются существенны параметры барьера, при  $a \rightarrow \infty$  – ямы. Правда, относительно часто встречаются системы, состоящие из трех различных материалов. Например, в МДП-системах, где проводящий слой заключен между полупроводником и изолятором затвора.

В квантовых нитях задача еще более усложняется. Существенное отличие от случая квантовых ям заключается в невозможности рассмотрения предельного перехода к нулевой толщине нити. В пленках при  $a \rightarrow 0$  энергия связи имеет конечное значение  $4Ry^*$ , а при стремлении к нулю радиуса нити энергия связи стремится к бесконечности. Это означает, что в достаточно тонких нитях примесное состояние может быть в принципе сколь угодно глубоким.

### **2.1.5. Экситоны**

Экситонные состояния – нейтральные элементарные возбуждения в кристалле. Они не могут быть отнесены к одноэлектронным, т.к. нейтральное состояние подразумевает наличие хотя

бы пары связанных зарядов. В первую очередь нас будет интересовать пара зарядов, электрон и дырка.

Формы существования такого нейтрального возбуждения в общем случае, могут иметь два предела: расстояние между электроном и дыркой либо очень мало, меньше размера элементарной ячейки, либо очень велико, много больше его.

В первом случае мы, очевидно, не вправе рассматривать экситон как пару носителей, описываемых блоховскими функциями, делокализованными в пространстве. Скорее их можно рассматривать как внутриатомные или внутримолекулярные возбуждения. Реально они оказываются существенны и могут быть сравнительно просто описаны в молекулярных кристаллах, в которых межмолекулярное взаимодействие слабо и им можно пренебречь при описании энергетике возбужденных состояний или учесть влияние среды какими-то усредненными параметрами. Например, поляризуемостью. Такие экситоны впервые были описаны Я.И.Френкелем и называются *экситоны Френкеля* или *экситоны малого радиуса*.

Второй случай, *экситонов большого радиуса* или *экситонов Ванье-Мотта*, предполагает, что взаимное притяжение пары свободных носителей – слабое возмущение, накладываемое на уже известные блоховские состояния. Они наблюдаются и существенно сказываются на электронных свойствах классических, ковалентных и ионных полупроводников, в первую очередь – на оптических свойствах.

Экситоны Френкеля подвижны. Состояние кристалла с одной возбужденной молекулой  $N$ -кратно вырождено и возбуждение может перемещаться вследствие индукционно-резонансного взаимодействия идентичных молекул. Экситоны Ванье-Мотта подвижны, поскольку образованы подвижными квазичастицами. Но экситонами часто называют и вообще любые нейтральные возбуждения, не обязательно подвижные. Например, возбужде-

ние, образованное переводом основного электрона на свободный уровень атома или молекулы можно рассматривать как неподвижное, поскольку оно имеет большую энергию и быстро деградирует, отдавая ее на более мелкие возбуждения.

Мы ограничимся рассмотрением связанных состояний одноэлектронных носителей, электрона и дырки.

### 2.1.5.1. Экситоны Ванье-Мотта

Пусть  $\mathbf{r}_n, \mathbf{r}_p$  – радиус-векторы электрона и дырки,

$\mathbf{k}_n, \mathbf{k}_p$  – их волновые векторы,

$E_c, E_v$  – их энергии, соответственно, в зоне проводимости и в валентной.

Обобщая уравнение Ванье (2.1.28) на систему, состоящую из электрона и дырки, взаимодействующих друг с другом по закону Кулона, получим:

$$\left[ E_c(-i\nabla_n) - E_v(-i\nabla_p) - \frac{e^2}{\varepsilon|\mathbf{r}_n - \mathbf{r}_p|} \right] \psi(\mathbf{r}_n, \mathbf{r}_p) = E\psi(\mathbf{r}_n, \mathbf{r}_p). \quad (2.1.37)$$

Здесь  $\nabla_n, \nabla_p$  - операторы в координатах, соответственно, электрона и дырки.

Потенциальную энергию взаимодействия мы записали так, как было бы справедливо в однородной непрерывной среде, в которой экранирование адекватно описывается диэлектрической проницаемостью  $\varepsilon$ . Реально же такое приближение не точно, но его можно принять при больших расстояниях между зарядами, существенно превышающих межатомные расстояния. В конце рассмотрения энергетики экситона Ванье-Мотта мы должны будем оценить справедливость такого описания.

Если полупроводник – прямозонный, т.е. экстремумы обеих зон находятся в центре зоны Бриллюэна, и если законы диспер-

сии можно считать параболическими, с определенными эффективными массами, то энергии носителей вблизи экстремумов:

$$E_c(k_n) = E_c(0) + \frac{\hbar^2 k_n^2}{2m_n}, \quad E_v(k_p) = E_v(0) - \frac{\hbar^2 k_p^2}{2m_p} \quad (2.1.38)$$

и тогда (2.1.37) упростится:

$$\left[ -\frac{\hbar^2}{2m_n} \nabla_n^2 - \frac{\hbar^2}{2m_p} \nabla_p^2 - \frac{e^2}{\varepsilon |\mathbf{r}_n - \mathbf{r}_p|} \right] \psi(\mathbf{r}_n, \mathbf{r}_p) = (E - E_g) \psi(\mathbf{r}_n, \mathbf{r}_p), \quad (2.1.39)$$

здесь  $E_g$  – ширина запрещенной зоны.

Введем радиус-векторы  $\mathbf{R}$  – центра тяжести электрона и дырки и  $\mathbf{r}$  – положения электрона относительно дырки:

$$\mathbf{R} = \frac{m_n \mathbf{r}_n + m_p \mathbf{r}_p}{m_n + m_p}, \quad \mathbf{r} = \mathbf{r}_n - \mathbf{r}_p, \quad (2.1.40)$$

приведенную массу электрона и дырки,  $\mu$ :  $\frac{1}{\mu} = \frac{1}{m_n} + \frac{1}{m_p}$ , и их суммарную массу  $M = m_n + m_p$ . В этих новых обозначениях (2.1.39) запишется как:

$$\left[ -\frac{\hbar^2}{2M} \nabla_R^2 - \frac{\hbar^2}{2\mu} \nabla_r^2 - \frac{e^2}{\varepsilon r} \right] \psi(\mathbf{R}, \mathbf{r}) = (E - E_g) \psi(\mathbf{R}, \mathbf{r}) \quad (2.1.41)$$

и переменные разделяются:

$$\psi(\mathbf{R}, \mathbf{r}) = \chi(\mathbf{R})\varphi(\mathbf{r}), \quad (2.1.42)$$

$$-\frac{\hbar^2}{2M} \frac{\nabla_R^2 \chi(\mathbf{R})}{\chi(\mathbf{R})} - \frac{1}{\varphi(\mathbf{r})} \left[ \frac{\hbar^2}{2\mu} \nabla_r^2 \varphi(\mathbf{r}) + \frac{e^2}{\varepsilon r} \varphi(\mathbf{r}) \right] = E - E_g. \quad (2.1.43)$$

Поскольку справа стоит величина постоянная, то оба слагаемых слева – тоже постоянные величины. Обозначим их  $W$  и  $E_e$ , причем, естественно,  $W + E_e = E - E_g$ .

Получили два уравнения:

$$-\frac{\hbar^2}{2M} \nabla_R^2 \chi(\mathbf{R}) = W \chi(\mathbf{R}) \quad (2.1.44a)$$

и 
$$-\frac{\hbar^2}{2\mu} \nabla_r^2 \varphi(\mathbf{r}) - \frac{e^2}{\varepsilon r} \varphi(\mathbf{r}) = E_e \varphi(\mathbf{r}). \quad (2.1.44б)$$

Первое из них – уравнение движения частицы массой  $M$ . его энергия равна:

$$W = \frac{\hbar^2 k_M^2}{2M}, \quad (2.1.45)$$

где  $k_M$  – волновой вектор плоской волны, характеризующей движение центра масс экситона.

Второе – хорошо известное уравнение, дающее связанные водородоподобные состояния с энергиями, отличающимися от ридберговских из-за отличия приведенной массы экситона от массы свободного электрона и из-за наличия диэлектрического экранирования поля:

$$E_{en} = -\frac{1}{n^2} R y_\infty \cdot \frac{\mu}{m_0} \frac{1}{\varepsilon^2}, \quad (2.1.46)$$

где  $R y_\infty = \frac{m_0 e^4}{2\hbar^2} \approx 13,6 \text{ эВ}$  – постоянная Ридберга для водородоподобного атома с бесконечно тяжелым ядром.

Радиусы орбиталей:

$$r_n = \frac{\hbar^2 \epsilon n^2}{\mu e^2} = a_B \frac{m_0 \epsilon}{\mu} n^2, \quad (2.1.47)$$

где  $a_B = \frac{\hbar^2}{m_0 e^2} \approx 0,0529$  нм – боровский радиус.

Оценки величин энергии связи и радиусов экситонов можно сделать для некоторого гипотетического вещества со средними для классических полупроводников параметрами:  $\mu = 0,1m_0$  и  $\epsilon = 16$ . Получим  $E_{el} \approx 0,005$  эВ и  $r_1 \approx 8,5$  нм. Это расстояние более, чем на порядок превышает межатомные. Поэтому можно полагать, что наше начальное предположение о допустимости учета вклада среды в притяжение электрона и дырки ее усредненной по пространству характеристикой,  $\epsilon$ , удовлетворительно.

Таким образом, мы получили, что пара носителей в кристалле может образовывать связанные состояния с энергией, не намного меньше ширины запрещенной зоны  $E_g$ , см. (2.1.46). Они могут быть образованы при оптическом возбуждении квантами

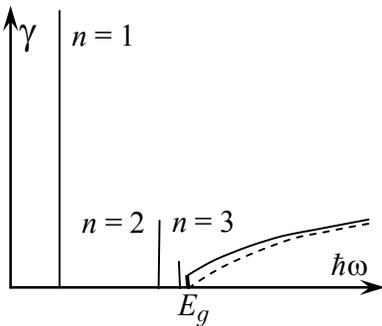


Рис. 2.1.13. Идеализированный вид спектра поглощения в припороговой области.

света с энергией, на  $E_{en}$  меньше  $E_g$  и, соответственно, давать дополнительные полосы поглощения. Но для их наблюдения нужны низкие температуры. Наша оценка энергии связи, 5 мэВ, впятеро меньше  $kT$  при комнатной температуре.

Наличие экситонного взаимодействия сказывается

и на форме спектра собственного поглощения в области  $\hbar\omega > E_g$ , увеличивая плотность конечных состояний оптического перехода по сравнению с плотностью состояний истинно свободных электронов, описываемых функциями Блоха.

На рис. 2.1.13 изображен идеализированный спектр поглощения в экситонной области. Изображены три разрешенных перехода в связанные состояния при  $\hbar\omega < E_g$  и область переходов в свободные состояния, при  $\hbar\omega \geq E_g$ . Сплошной линией – с учетом экситонных эффектов, штриховой – теоретический спектр, переходы между блоховскими состояниями.

### 2.1.5.2. Экситоны в квантовых ямах

Здесь рассмотрение – то же, что было проведено в разд. 2.1.4, 2.1.5.1. Можно отправляться от уравнения (2.1.39). Затем, приняв допущение (2.1.33), разделить задачу на две независимые, о движении носителей в плоскости ямы и перпендикулярно к ней. Здесь несколько усложнится вид волновой функции, поскольку мы рассматриваем два носителя, а не один:

$$\Psi = \varphi(\rho)\chi_n(z_n)\chi_p(z_p).$$

Далее, для решения задачи о движении носителей в плоскости ямы введем, как и прежде, обозначения:

$$\mathbf{R} = \left( \frac{m_n x_n + m_p x_p}{m_n + m_p}, \frac{m_n y_n + m_p y_p}{m_n + m_p} \right), \quad \mathbf{r} = (x_n - x_p, y_n - y_p),$$

$$\frac{1}{\mu} = \frac{1}{m_n} + \frac{1}{m_p}, \quad M = m_n + m_p.$$

Здесь мы снова приходим к уравнению типа (2.1.44б), но в двухкоординатном пространстве. Поэтому решение получаем в

виде (2.1.46) с той лишь разницей, что изменяется закон квантования:

$$E_{ei} = -\frac{1}{(i-1/2)^2} Ry_{\infty} \cdot \frac{\mu}{m_0} \frac{1}{\varepsilon^2}. \quad (2.1.48)$$

Как и в случае мелких примесей (разд. 2.1.4) здесь энергия связи основного состояния в 4 раза больше, чем в массивном образце. Поэтому в квантовых ямах экситонные переходы иногда можно наблюдать и при комнатной температуре.

В наиболее распространенных системах, на основе Si, Ge, GaAs, в области занятых состояний перекрываются зоны тяжелых и легких дырок. Соответственно, приведенная масса в (2.1.48) может иметь два значения, и в этих веществах наблюдаются так называемые "легкие" и "тяжелые" экситоны.

### 2.1.6. Заключительные замечания

Собственные электронные состояния в кристаллах "классических"<sup>1</sup> полупроводников – бегущие волны, описываемые бловскими волновыми функциями

$$\psi_k(\mathbf{r}) = u_k(\mathbf{r}) \cdot \exp(i\mathbf{k}\mathbf{r}).$$

Перенос заряда, т.е. движение электронов в пространстве, описывается в них практически так же, как токи в вакууме. Различия заключаются в сложном законе дисперсии, в наличии двух типов носителей, и в наличии рассеяния электронов.

Поскольку законы дисперсии сложны, то и в тех случаях, когда они могут быть приближены параболой, коэффициент пропорциональности между энергией и квадратом импульса не равен

---

<sup>1</sup> В данном случае "классическими" мы считаем полупроводники, давно используемые в электронной технике. В первую очередь к ним относятся Ge, Si, GaAs, вещества с высокой подвижностью носителей.

обратной массе свободного электрона. Приходится вводить понятие *эффективной массы*, которая может быть и больше, и меньше массы свободного электрона. Более того, во-первых, эта эффективная масса обычно зависит от направления движения, а во-вторых, возможно одновременное существование в кристалле одноименных носителей с одинаковыми импульсами, но разными эффективными массами. В валентной зоне как правило, приходится вводить понятия "тяжелых" и "легких" дырок.

Впрочем, и в таких сложных случаях для описания проводимости обычно достаточно усреднить их, ввести понятие средней эффективной массы.

Наличие двух типов носителей оказывается очень принципиальным свойством кристаллов. На различии энергий электронов и дырок основаны действия всех полупроводниковых приборов.

Рассеяние движущихся носителей происходит на флуктуациях внутрикристаллического потенциала. Они могут возникнуть при наличии дефектов структуры или фононов, колебаний решетки. Эффективность первого типа рассеяния почти не зависит от температуры<sup>1</sup>, а рассеяние на фононах, естественно, зависит очень сильно. В чистых кристаллах при изменении температуры от 4 до 290 К вероятность рассеяния может возрасти на 5 ÷ 6 порядков.

Рассеяние приводит к стабилизации средней скорости дрейфа носителей, которая линейно зависит от напряженности приложенного поля. Оно же приводит к возникновению таких явлений, как наличие конечной проводимости (не зависящей от приложенного поля), и выделение джоулева тепла.

Но для технического использования, пожалуй, самое принципиальное свойство полупроводников – легко изменяемая кон-

---

<sup>1</sup> Но оно зависит от энергии носителей.

центрация носителей. Тип (электроны или дырки) и равновесную концентрацию носителей можно сформировать, введя нужные примеси при выращивании кристалла. Далее, эту исходную концентрацию (и даже тип) носителей можно в очень широких пределах изменять температурой, облучением, внешним электрическим полем... В технике для управления составом и, соответственно, свойствами приповерхностных слоев пользуются ионной имплантацией или диффузией примеси при прогреве кристаллов в ее парах.

Если привести в контакт два разнородных кристалла, то между ними возникает переходная область (*p-n*-переход) с очень интересными и полезными свойствами. Мы их рассмотрим в разделе 2.4.2.5.

Все разнообразие свойств кристаллов, которое необходимо для создания современных электронных устройств обеспечивает управление процессами образования в кристалле необходимых дефектов. В основном – вводимых химических примесей, обеспечивающих образование в кристалле дополнительных электронных состояний, не свойственных чистому кристаллу. Любой дефект, любое нарушение идеальной структуры кристалла приведет к отклонениям от идеального хода потенциала и, следовательно, к образованию некоторого специфического электронного состояния, энергия которого зачастую соответствует запрещенной зоне. Перенос же заряда происходит, как правило, по собственным состояниям кристалла, делокализованным в пространстве. Электроны и дырки, занимающие примесные состояния, оказываются локализованы в пространстве. Они не принимают участия в проводимости, а являются чем-то вроде резервуара, поставщика носителей в зоны. Поэтому рассмотрение дефектов и связанных с ними состояний не менее важно, чем рассмотрение собственных состояний кристалла.

Дефекты могут иметь различную природу. В основном, следует рассматривать как существенно различные:

- точечные дефекты структуры – дефекты Френкеля и Шоттки, атомы в междузлиях, (см. рис. 2.1.6 и 2.1.7),
- точечные дефекты – химические примеси;
- линейные (одномерные) дефекты, – дислокации, см. рис. 2.1.8;
- двумерные дефекты, – плоскости излома, скольжения и, конечно, поверхность, свойства которой можно существенно изменять адсорбцией<sup>1</sup> или наложением электрического поля.

В разделе 2.2.2 показано, что даже при небольших концентрациях они могут существенно повлиять на концентрацию носителей тока.

В линейных и двумерных дефектах возможно образование состояний, делокализованных в пространстве. Соответственно, одно- или двумерных. Их рассмотрение затруднительно и мы ограничимся предположением, что имеются только точечные дефекты, тем более, что человечество уже научилось выращивать достаточно большие кристаллы, в которых нет или почти нет пространственно протяженных дефектных состояний.

Не вдаваясь в детали строения возможных дефектов кристалла, будем характеризовать связанные с ними локальные состояния параметрами, существенными для нашего рассмотрения.

---

<sup>1</sup> Адсорбция чужеродных атомов или молекул – процесс, аналогичный объемному легированию, введению примесей. Различия между нами очевидны: адсорбированными слоями легче управлять и существуют они в среде, существенно отличной от объема. Кроме того, что кристалл у них занимает лишь полпространства, на границе раздела возникают и дополнительные состояния (уровни Тамма) и дополнительные пространственные фигуры (изломы, ступеньки и т.д.).

В первую очередь это:

- тип дефекта (донорный, акцепторный);
- концентрация;
- перечень возможных зарядовых состояний примеси и энергии соответствующих им основного и возбужденных уровней;
- сечения захвата носителей дефектом в различных зарядовых и - возбужденных состояниях.

Резюмировать сказанное можно следующим образом, добавив к тому же выводы, которые мы здесь не пытались обосновывать:

1. Примеси и иные дефекты кристаллической решетки дают в энергетическом спектре полупроводника *локальные* уровни в запрещенной зоне:

Термин *локальные* означает, что локальны реальные точки кристалла, в которых электрон может иметь данную энергию. Его состояние в этом случае не описывается бегущей волной и, за исключением случаев очень высокой концентрации примесей, перенос электрона по локальным уровням невозможен. Для перемещения с одного центра на другой, электрон (дырку) необходимо возбудить в зону, там он может переместиться куда угодно и в конце концов “захватиться” на другой (не обязательно такой же) локальный центр.

2. При наличии примесей может существенно измениться энергия Ферми в полупроводнике. Введение донорной примеси смещает уровень Ферми вверх (уменьшается энергия связи электронов в кристалле); введение акцепторной - вниз (увеличивается энергия связи).

3. При наличии примесей разных типов возможна их взаимная компенсация, – электроны доноров захватываются акцепторами. Если концентрации доноров и акцепторов примерно равны и их уровни расположены близко к с- и v-зонам соответственно,

то эта компенсация приводит к тому, что по проводимости полупроводник получается практически “собственным”, высокоомным. Такой полупроводник называют *компенсированным*.

4. Заряженные примеси или иные дефекты могут увеличить вероятность рассеяния, понизить подвижность носителей и тем самым понизить проводимость. Однако, этот эффект при малых концентрациях примесей и относительно высоких температурах (комнатных) существенно слабее, чем увеличение проводимости за счет увеличения концентрации носителей.

5. При увеличении концентрации точечных дефектов уменьшается расстояние между соседними центрами и возникает их взаимодействие, которое так же, как и в основном веществе, приводит к расщеплению их уровней и, возможно, достаточно эффективному переносу заряда по образовавшейся "дефектной" зоне. Если расщепление становится велико, то "дефектная" зона может перекрыться с собственной. Например, зона донорных примесей с с-зоной. В этом случае возникает ситуация, при которой уровень Ферми, разделяющий области занятых и свободных состояний, оказывается в с-зоне. Полупроводник вырождается в металл. Такие полупроводники и называются *ВЫРОЖДЕННЫМИ*. Мы будем рассматривать, в основном, эффекты изменения концентрации носителей в невырожденных полупроводниках и их взаимодействие с полями. Эффекты вырождения нас будут интересовать при рассмотрении квантовых ям, двумерных полупроводниковых структур. Если создается поверхностный слой 2D-состояний, то степень вырождения, концентрация носителей и толщина слоя легко могут быть изменены внешним электрическим полем.

## 2.2. СТАТИСТИКА НОСИТЕЛЕЙ ЗАРЯДА В ПОЛУПРОВОДНИКАХ

Зная типы и свойства электронных состояний в кристалле, можно рассчитать и концентрации носителей в зонах. Сначала – термически равновесные концентрации.

Для электронов, подчиняющихся принципу Паули, справедлива статистика Ферми-Дирака, согласно которой вероятность заполнения электроном состояния с энергией  $E$  равна:

$$f(E, F) = \frac{1}{\exp\left(\frac{E - F}{kT}\right) + 1}. \quad (2.2.1)$$

Здесь  $F$  – *электрохимический потенциал* или *энергия Ферми*, т.е. средняя энергия связи, проявляющаяся в термодинамически равновесных реакциях обмена электроном. В (2.2.1)  $F$  входит как параметр распределения. На энергетических диаграммах типа изображенной на рис. 2.1.9 ее обычно отмечают, рисуя в запрещенной зоне (для *невырожденного полупроводника*<sup>1</sup>) уровень с энергией  $F$ . Поэтому термин "*уровень Ферми*" столь же обычен, как и "*энергия Ферми*".

Вероятность того, что состояние свободно (т.е. "занято дыркой") равна, очевидно,  $1 - f(E, F)$  :

$$f_p(E, F) = 1 - f(E, F) = \frac{1}{\exp\left(-\frac{E - F}{kT}\right) + 1}. \quad (2.2.2)$$

Вдали от уровня Ферми, при  $|E - F| \gg kT$ , вероятности заполнения  $f(E, F)$  стремятся к предельным значениям, – к нулю при  $E > F$  и к единице при  $E < F$ . Вероятность  $f_p(E, F)$ , соответствен-

---

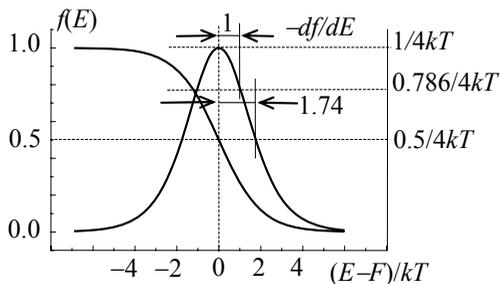
<sup>1</sup> См. Термины и определения в [1].

но, наоборот – к единице при  $E > F$  и к нулю при  $E < F$ . В этих случаях можно считать, что принцип Паули уже не ограничивает распределения электронов по состояниям и выражения (2.2.1), (2.2.2) упрощаются. При  $|E - F| \gg kT$  получаем обычные больцмановские факторы:

$$f_c(E, F) \approx \exp\left(-\frac{E - F}{kT}\right), \quad f_p(E, F) \approx \exp\left(\frac{E - F}{kT}\right) \quad (2.2.3)$$

Если расстояния от уровня Ферми до краев зон,  $E_c$  и  $E_v$ , больше нескольких  $kT$ , то выражения (2.2.3) справедливы для всех зонных состояний. Эту ситуацию мы будем называть **далекой от вырождения**. Она представляет для нас основной интерес, поскольку является как бы "чисто полупроводниковой". Когда приближение (2.2.3) не справедливо, получим некоторые переходные электронные свойства кристалла, нечто среднее между металлом и полупроводником. Ошибка, возникающая при замене (2.2.2) на (2.2.3), не превышает 5% уже при  $|E - F| > 3kT$ .

Рис. 2.2.1. Фермиевская функция распределения  $f(E)$  и ее производная по энергии. Ширина функции  $df/dE$  на полувысоте  $\approx 3.5 kT$



На рис. 2.2.1 изображена функция (2.2.1) и производная от нее по энергии (со знаком минус):

$$\frac{df(E)}{dE} = \frac{1}{kT} \frac{\exp\left(\frac{E-F}{kT}\right)}{\left[\exp\left(\frac{E-F}{kT}\right) + 1\right]^2} \quad (2.2.4)$$

Термически равновесное распределение электронов по всем возможным состояниям может быть определено, если известны ширина запрещенной зоны  $\Delta E_g$ , распределение плотности состояний в зонах  $D(E)$  (см. раздел 1.6.1.1 в [1]), концентрации и энергии локальных состояний в запрещенной зоне, температура.

Фигурирующая в (2.2.1), (2.2.2) величина энергии Ферми  $F$  не является какой-то точно заданной величиной, как, например, энергии краев зон или локальных состояний. Это – некоторая средняя энергия электронного ансамбля и она, конечно, может зависеть от условий, в которые мы поместили кристалл. В том числе – от температуры. Но вот полный заряд кристалла от температуры и иных внешних воздействий не зависит. Будем считать его равным нулю<sup>1</sup>. Тогда при любой температуре сумма положительных зарядов, свободных и локализованных, должна быть равна сумме отрицательных:

$$n + N_a^- = p + N_d^+ . \quad (2.2.5)$$

Здесь  $n$  и  $p$  – концентрации электронов в с-зоне и дырок в v-зоне,  $N_a^-$  и  $N_d^+$  – концентрации заряженных акцепторов и доноров.

---

<sup>1</sup> Кристалл можно зарядить, но даже при потенциале  $10^6$  В количество электронов в кристалле объемом  $1 \text{ мм}^3$  изменится меньше, чем на  $10^{-6} \%$  !

### **2.2.1. Статистика электронов в собственном полупроводнике**

**СОБСТВЕННЫМИ** называют полупроводники, в которых концентрация дефектов пренебрежимо мала.

В практическом плане собственные полупроводники малоинтересны, так как они имеют очень низкую проводимость. Но рассмотрение их свойств весьма полезно, так как, во-первых, это – одно из предельных состояний полупроводника, во-вторых, перенос заряда в любом случае<sup>1</sup> осуществляется именно по собственным состояниям. И, наконец, в практически важных полупроводниковых устройствах, диодах, транзисторах и т.д., обязательно где-то между электродами присутствует слой с собственной проводимостью, запиорный слой. И энергетическое распределение, и подвижности носителей в зонах, в первом приближении можно считать одинаковыми и в собственном, и примесном полупроводниках.

#### **2.2.1.1. Концентрация носителей**

Для определенности мы будем говорить об электронах в  $s$ -зоне. Все соображения целиком переносимы на дырки в  $v$ -зоне, отличие только в направлении отсчета энергии и в законах дисперсии. Мы будем считать их параболическими, т.е. характеризуемыми **эффективными массами**, см. разд. 1.6.1.2 в [1].

Количество электронов с энергиями от  $E$  до  $E + dE$ ,  $N(E)dE$ , определяется плотностью состояний  $D(E - E_c)$  и функцией заполнения (2.2.1). Объем считаем единичным, так что

---

<sup>1</sup> Проводимость по примесным состояниям возможна, но лишь тогда, когда расстояния между примесными атомами становятся достаточно малы, т.е. при очень большой их концентрации. Мы такие ситуации рассматривать не будем.

$$N(E)dE = D(E - E_c)f(E, F)dE \quad (2.2.6)$$

Концентрация электронов – интеграл от (2.2.6) по всей зоне. Верхний предел интегрирования можно положить равным бесконечности, поскольку функция Ферми (2.2.1) достаточно быстро стремится к нулю при возрастании энергии  $E$ .

$$n = \int_{E_c}^{\infty} D(E - E_c)f(E)dE = \frac{(2m^*)^{3/2}}{2\pi^2\hbar^3} \int_{E_c}^{\infty} \frac{\sqrt{E - E_c}}{\exp\left(\frac{E - F}{kT}\right) + 1} dE. \quad (2.2.7)$$

Интеграл в (2.2.7) в явном виде не берется, он сводится к так называемому **интегралу Ферми-Дирака с индексом  $1/2$** ,  $\Phi_{1/2}(\zeta^*)$ :

$$\Phi_{1/2}(\zeta^*) \equiv \frac{2}{\sqrt{\pi}} \int_0^{\infty} \frac{x^{1/2}}{\exp(x - \zeta^*) + 1} dx, \quad (2.2.8)$$

В нашем случае:

$$\zeta \equiv F - E_c, \quad \zeta^* \equiv \frac{F - E_c}{kT}, \quad x \equiv \frac{E - E_c}{kT}, \quad (2.2.9a)$$

где  $\zeta$  - химический потенциал для электронов,  $\zeta^*$  - безразмерный химический потенциал.

При такой замене переменных в (2.2.7) получим

$$n = 2 \left( \frac{2\pi m_n kT}{(2\pi\hbar)^2} \right)^{3/2} \Phi_{1/2}(\zeta^*) \equiv N_c \Phi_{1/2}(\zeta^*). \quad (2.2.10)$$

Здесь введено новое обозначение. Величина

$$N_c = 2 \left( \frac{2\pi m_n kT}{(2\pi\hbar)^2} \right)^{3/2}, \quad (2.2.11a)$$

зависящая только от температуры и кривизны закона дисперсии  $E(k)$ , т.е. эффективной массы  $m_n$ , называется **эффективной плотностью состояний** электронов в с-зоне. При  $m_n = m_0$  и  $T = 300$  К  $N_c \approx 2,51 \cdot 10^{19}$  см<sup>-3</sup>, а в общем случае:

$$N_c \approx 2,51 \cdot 10^{19} \left( \frac{m_n}{m_0} \frac{T}{300} \right)^{3/2} \text{ см}^{-3}. \quad (2.2.12)$$

Далее мы всегда будем считать, что полупроводник далек от вырождения. Тогда единицей в знаменателе (2.2.1), (2.2.8) можно пренебречь (принцип Паули не существен) и функция заполнения (2.2.1) превращается в больцмановскую, интеграл в (2.2.7) берется. Получаем:

$$n = N_c \exp\left(-\frac{E_c - F}{kT}\right). \quad (2.2.13a)$$

Аналогичные выражения справедливы и для дырок:

$$p = N_v \exp\left(\frac{E_v - F}{kT}\right) = N_v \exp\left(\frac{\eta}{kT}\right) = N_v \exp(\eta^*), \quad (2.2.13b)$$

где 
$$N_v = 2 \left( \frac{2\pi m_p kT}{(2\pi\hbar)^2} \right)^{3/2} \quad (2.2.11b)$$

и 
$$\eta \equiv E_v - F, \quad \eta^* \equiv \frac{E_v - F}{kT}. \quad (2.2.9b)$$

### 2.2.1.2. Средние энергия, скорость и волновой вектор носителей

Распределение (2.2.6) для *параболической зоны в невырожденном случае* не отличается от максвелловского. Если концентрация электронов во всех случаях достаточно мала для того, чтобы запрет Паули не влиял на распределение носителей по состояниям (такой ансамбль называют **Ферми-газом**) и закон дис-

персии в зоне можно считать параболическим, то распределение в Ферми-газе – то же, что в идеальном газе молекул. В нормированном на единицу виде получим:

$$\frac{N(E)}{n} = \frac{2}{\sqrt{\pi}} \exp\left(-\frac{E_k}{kT}\right) \sqrt{E_k} \quad (2.2.14)$$

Здесь  $E_k \equiv E - E_c > 0$ .

Как известно (см., например, [4]), средние параметры движения частиц при максвелловском распределении равны:

$$\langle E_k \rangle = \frac{3}{2} kT; \quad \langle v \rangle = \sqrt{\frac{8kT}{\pi m^*}}; \quad \langle k \rangle = \frac{1}{\hbar} \sqrt{\frac{8m^* kT}{\pi}}. \quad (2.2.15)$$

Заметьте, что от положения уровня Ферми они не зависят.

При  $T = 300$  К и  $m^* = 0.1m_0$  получим средние значения:

$$\langle E_k \rangle = 0,039 \text{ эВ}, \quad \langle v \rangle \cong 3.4 \cdot 10^7 \text{ см} \cdot \text{с}^{-1}; \quad \langle k \rangle \cong 2.9 \cdot 10^6 \text{ см}^{-1} \quad (2.2.16)$$

При постоянной решетки порядка 0,3 нм размер зоны Бриллюэна равен  $\pm 10^8 \text{ см}^{-1}$ , так что термически равновесные электроны заполняют лишь малую область близ ее минимума.

Средняя длина волны  $\lambda = 2\pi/k \approx 22$  нм, много больше постоянной решетки.

### 2.2.1.3. Положение уровня Ферми в собственном полупроводнике

Энергию Ферми  $F$  можно определить из условия электронейтральности. В собственном, беспримесном полупроводнике, оно выглядит просто  $n = p \equiv n_i$ . Символом  $n_i$  всегда обозначают концентрацию носителей в собственном полупроводнике.

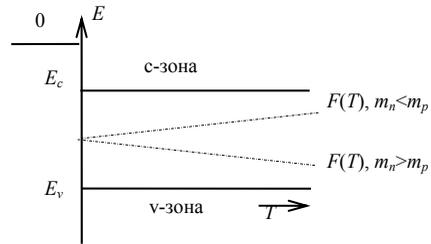
$$N_c \exp\left(-\frac{E_c - F}{kT}\right) = N_v \exp\left(-\frac{F - E_v}{kT}\right). \quad (2.2.17)$$

Прологарифмировав, получим:

$$\begin{aligned} F &= \frac{1}{2}(E_v + E_c) + \frac{1}{2}kT \ln(N_v/N_c) = \\ &= \frac{1}{2}(E_v + E_c) + \frac{3}{4}kT \ln(m_p/m_n). \end{aligned} \quad (2.2.18)$$

Рис. 2.2.2. Зависимость положения уровня Ферми в собственном полупроводнике от температуры.

При нагревании уровень Ферми линейно смещается от середины запрещенной зоны к зоне более легких носителей.



При нулевой температуре уровень Ферми расположен в середине запрещенной зоны. С повышением температуры он приближается к зоне более легких носителей. В ней меньше плотность состояний, а условие электронейтральности выполняться должно. Следовательно, в (2.2.17) при меньшем множителе (\$N\_c\$ или \$N\_v\$) должен стоять больший больцмановский фактор.

#### 2.2.1.4. Закон действующих масс

**Произведение концентраций электронов и дырок в невырожденном полупроводнике не зависит от положения уровня Ферми.** Действительно, из (2.2.13), (2.2.13б) получим:

$$\begin{aligned} np = n_i^2 &= N_c N_v \exp\left(-\frac{E_c - E_v}{kT}\right) \\ &= 4 \frac{(2\pi kT)^3}{(2\pi\hbar)^6} (m_n m_p)^{3/2} \exp\left(-\frac{E_g}{kT}\right). \end{aligned} \quad (2.2.19)$$

В дальнейшем мы увидим, что легированием можно почти произвольно изменять положение уровня Ферми, но произведение концентраций носителей зависит *только* от ширины запрещенной зоны и температуры. Увеличивая концентрацию одних носителей, мы обязательно уменьшим концентрацию других во столько же раз.

Закон действующих масс равно справедлив и для собственного, и для примесного полупроводников. В частности, из него следует, что собственный полупроводник имеет максимальное сопротивление из возможных для данного вещества. Легированием его сопротивление можно только уменьшить, так как проводимость пропорциональна первой степени концентраций:

$$\sigma = e(n\mu_n + p\mu_p) = 2e\langle\mu\rangle n_i \quad (2.2.20)$$

(здесь  $\mu_n$ ,  $\mu_p$  и  $\langle\mu\rangle$  – подвижности электронов, дырок и средняя подвижность). Проводимость собственного полупроводника  $\sigma_i \approx 2e\langle\mu\rangle n_i$ . Сделав легированием  $n = 10n_i$ , получим  $p = 0.1n_i$  и проводимость  $\sigma \approx 10,1 \cdot e\langle\mu\rangle n_i$ , 5 раз больше, чем в собственном!

### 2.2.1.5. Эффективная масса плотности состояний

В тех случаях, когда поверхности постоянной энергии не-сферичны, но тензор обратной эффективной массы приводится к осям, как в с-зоне Si или Ge (см. рис. 2.1.1, 2.4.2), для определения концентрации носителей можно все же пользоваться выражением (2.2.13), подставив в  $N_c$  эффективную массу, усредненную по всем направлениям и откорректированную на количество долин  $\nu$ :

$$m_d = \nu^{2/3} (m_x m_y m_z)^{1/3}. \quad (2.2.21)$$

В Si поверхности постоянной энергии в с-зоне – эллипсоиды вращения, характеризуемые массами  $m_{xx} = 0,9163m_0$  и

$m_{\perp} = 0,1905m_0$ . Количество долин  $\nu = 6$ . Эффективная масса плотности состояний равна

$$m_d = 6^{2/3} (m_{xx} m_{\perp}^2)^{1/3} \approx 1,06m_0.$$

### 2.2.2. Статистика электронов в примесном полупроводнике

**ПРИМЕСНЫМ** называют полупроводник, в котором концентрация дефектов (структурных или химических) достаточно велика для того, чтобы повлиять на концентрацию носителей и положение уровня Ферми.

Все приведенные выше соотношения относительно плотностей состояний в зонах, вероятности возбуждения электронов при конечной температуре, параметров носителей заряда и другие остаются и в этом случае в силе, но изменяется условие электронной нейтральности кристалла в целом. Теперь его надо записывать в виде (2.2.5), причем под  $N_a^-$  и  $N_d^+$  надо понимать полные суммы концентраций всех типов заряженных доноров и акцепторов с учетом величин заряда (может быть 1, 2 и более). В общем случае задача расчета параметров такого полупроводника сложна, но для примера и оценки того, что можно ожидать, мы рассмотрим три случая, – полупроводник с одним типом дефектов, способным захватить только единичный заряд, полупроводник с двумя однозарядными разноименными и полупроводник с одним типом дефектов, способных захватить (отдать) несколько электронов<sup>1</sup>.

Проблема расчета концентрации носителей сводится к определению положения уровня Ферми в общем случае.

---

<sup>1</sup> Как, например, примесь Au в Ge или Si, см. рис. 2.1.10.

### 2.2.2.1. Концентрация зарядов на простых локальных центрах

Этот вопрос следует рассмотреть прежде, чем, воспользовавшись условием электронейтральности (2.2.5), определим положение уровня Ферми. Нужно определить концентрации локализованных зарядов,  $N_a^-$  и  $N_d^+$ .

Пусть  $N_t$  – концентрация простых центров, которые могут либо содержать 1 электрон с энергией  $E_1$ , либо не содержать. Их статвеса в этих зарядовых состояниях обозначим  $g_1$  и  $g_0$ .

Вероятности того, что некоторый уровень будет заполнен или свободен, равны, соответственно,

$$N_1 = N_t f \text{ и } N_0 = N_t(1 - f),$$

где  $f$  – функция распределения Ферми (2.2.1).

С учетом статвесов отношение концентраций занятых и свободных центров составит:

$$\frac{N_1}{N_0} = \frac{g_1}{g_0} \exp\left(\frac{F - E_1}{kT}\right), \quad (2.2.22)$$

а сумма, естественно,

$$N_1 + N_0 = N_t. \quad (2.2.23)$$

Отсюда получаем  $f^{(1)}$  и  $f^{(0)}$ , вероятности заполнения и освобождения центров:

$$f^{(1)} = \frac{N_1}{N_t} = \left[ 1 + \frac{g_0}{g_1} \exp\left(\frac{E_1 - F}{kT}\right) \right]^{-1} \text{ и} \quad (2.2.24a)$$

$$f^{(0)} = \frac{N_0}{N_t} = \left[ 1 + \frac{g_1}{g_0} \exp\left(\frac{F - E_1}{kT}\right) \right]^{-1}. \quad (2.2.24б)$$

Введем удобную величину:

$$n_1 = \frac{g_0}{g_1} N_c \exp\left(-\frac{J}{kT}\right), \quad (2.2.25)$$

где  $J \equiv E_c - E_1$  – энергия переброса электрона с локального центра в  $s$ -зону. В этих обозначениях запись (2.2.24) существенно упрощается:

$$f^{(1)} = \frac{N_1}{N_t} = \frac{n}{n + n_1}, \quad f^{(0)} = \frac{N_0}{N_t} = \frac{n_1}{n + n_1}. \quad (2.2.26)$$

В принципе, центр, захвативший электрон, может иметь и возбужденные состояния с энергией в запрещенной зоне. Это приведет к эффективному увеличению статвеса занятого состояния и может быть учтено соответствующим увеличением  $g_1$ .

Пусть у занятого центра имеется  $k$  различных энергетических состояний, каждое – с энергией  $E_k$  и статвесом  $\beta_k$ . Обозначим энергии возбуждения состояний  $\varepsilon_k \equiv E_k - E_1$ . Естественно,  $\varepsilon_1 = 0$ . Тогда вероятность заселения каждого из возможных состояний пропорциональна  $\beta_k \exp\left(-\frac{\varepsilon_k}{kT}\right)$  и суммарный эффективный статвес  $g_1$  теперь будет равен сумме:

$$g_1 = \sum_{k=1,2,\dots} \beta_k \exp\left(-\frac{\varepsilon_k}{kT}\right). \quad (2.2.27)$$

Выражения (2.2.24) ÷ (2.2.26) при этом остаются в силе.

### 2.2.2.2. Полупроводник с одним типом примеси

Энергетическая схема такого полупроводника представлена на рис. 2.1.9, *а* или *б*. Для определенности, рассмотрим случай донорной примеси. Для акцепторной все соображения останутся теми же и получится тот же результат.

Простые доноры могут находиться в зарядовых состояниях 0 или +1. Предположим также, что при наличии доноров концентрация электронов будет существенно больше, чем в собственном полупроводнике и, соответственно, концентрация дырок пренебрежимо мала, см. закон действующих масс (2.2.19). Тогда уравнение электронейтральности кристалла запишется в виде:

$$n = p_t, \quad (2.2.28)$$

где  $p_t = N_d f^{(0)}$  – концентрация “захваченных” дырок есть концентрация опустошенных доноров.

Используя (2.2.26), сразу получим квадратное уравнение:

$$n^2 + n_1 n - n_1 N_d = 0,$$

решение которого дает:

$$n = \frac{n_1}{2} \left[ \sqrt{1 + \frac{4N_d}{n_1}} - 1 \right]. \quad (2.2.29)$$

Положение уровня Ферми для невырожденного полупроводника можно определить отсюда, подставив вместо  $n$  его значение из (2.2.13) и, прологарифмировав. Получим:

$$F = E_c + kT \ln \left( \frac{n_1}{2N_c} \left[ \sqrt{1 + \frac{4N_d}{n_1}} - 1 \right] \right) \quad (2.2.30)$$

При нагреве до достаточно высоких температур, когда  $4N_d \ll n_1$ , получим  $n = N_d$ . Все доноры отдали электроны в  $s$ -зону и концентрация электронов далее изменяться не может. А положение уровня Ферми:

$$F_{T\uparrow} = E_c + kT \ln \left( \frac{N_d}{N_c} \right) \quad (2.2.31)$$

и в этом случае зависит от температуры. Это естественно, т.к. если концентрация не изменяется, то из (2.2.13) следует, что для этого должно выполняться  $\frac{E_c - F}{kT} = Const$ . Это и написано в (2.2.31). Уровень Ферми находится ниже с-зоны ( $N_d < N_c$ ), с ростом температуры удаляется все дальше, а положение уровня доноров здесь почти не играет роли. Все они ионизованы, так что их перезарядка не возможна. Но (2.2.31) справедливо только пока нет межзонных перебросов. Здесь мы предполагали, что электроны в с-зоне появились только от доноров.

При низких температурах  $n_1 \rightarrow 0$ ,  $4N_d/n_1 \gg 1$  и тогда:

$$n_{T\downarrow} = \left( \frac{g_0}{g_1} N_d N_c \right)^{1/2} \exp\left(-\frac{J}{2kT}\right). \quad (2.2.32)$$

Концентрация экспоненциально зависит от обратной температуры. На графике  $\ln\left(nT^{-3/4}\right)$  от  $1/kT$  получим прямую линию с наклоном  $-J/2$ .

Воспользовавшись (2.2.13), перепишем соотношение (2.2.32) в виде:

$$\exp\left(\frac{F}{kT} - \frac{E_c + E_d}{2kT}\right) = \left(\frac{g_0}{g_1} \frac{N_d}{N_c}\right)^{1/2}$$

и, прологарифмировав его, найдем положение уровня Ферми:

$$F_{T\downarrow} = \frac{E_c + E_d}{2} + \frac{kT}{2} \ln\left(\frac{g_0}{g_1} \frac{N_d}{N_c}\right). \quad (2.2.33)$$

При нуле температуры он находится посередине между зоной и уровнем доноров. При нагревании удаляется от зоны со скоростью, вдвое меньшей (при  $g_0 = g_1$ ), чем при высоких  $T$ .

В широком температурном диапазоне получим сначала рост концентрации носителей по экспоненциальному закону с наклоном, определяемым половиной интервала между уровнем донора и дном зоны проводимости. Затем все доноры ионизируются и рост концентрации прекратится до тех пор, пока не станет возможным межзонное возбуждение, с энергией активации  $\frac{\Delta E_g}{2}$ , см. рис.

### 2.2.3.

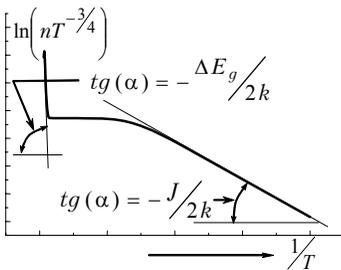


Рис. 2.2.3 Температурная зависимость концентрации носителей в полупроводнике с одним типом примеси.

#### 2.2.2.3. Частично компенсированные полупроводники.

Обычно все-таки в полупроводнике присутствуют и донорные, и акцепторные примеси. И обычно в существенно разных концентрациях. Одна из них доминирует, вторая – частично компенсирует эффект от первой.

Рассмотрим ситуацию, в которой присутствует два типа простых центров, с концентрациями  $N_d > N_a$ . В этом случае можно считать, что все акцепторы заряжены. Это справедливо при нуле температуры. Освободиться акцепторы могут лишь при наличии дырок или при конечной вероятности переброса электронов на-

верх, в с-зону или на доноры<sup>1</sup>. Эти процессы бывают возможны в узкозонных полупроводниках, при довольно высоких температурах, так что мы будем считать их маловероятными. Тогда условие электронейтральности примет вид :

$$n + N_a = N_d f^{(0)} = \frac{N_d n_1}{n + n_1}, \quad (2.2.34)$$

или:

$$n^2 + n(N_a + n_1) - n_1(N_d - N_a) = 0,$$

откуда

$$n = \frac{N_a + n_1}{2} \left( \sqrt{1 + \frac{4n_1(N_d - N_a)}{(N_a + n_1)^2}} - 1 \right). \quad (2.2.35)$$

При  $N_a \rightarrow 0$  это выражение переходит в (2.2.29), а при  $N_a > N_d$  вообще не действительно,  $n$  становится отрицательным.

В пределе низких температур, когда  $n_1 \rightarrow 0$ ,  $n_1 \ll N_a$ ,

$$n_{T\downarrow} = \frac{N_d - N_a}{N_a} \frac{g_0}{g_1} N_c \exp\left(-\frac{J_d}{kT}\right), \quad (2.2.36)$$

В этом случае концентрация тоже экспоненциально зависит от температуры, но теперь наклон прямой в зависимости  $\ln\left(nT^{-3/4}\right)$  от  $1/kT$  равен  $-J$ , а не  $-J/2$ , как для некомпенсированного случая, см. (2.2.32).

При высоких температурах, когда  $n_1 \gg N_a$ , получим  $n = N_d - N_a$ , все доноры ионизованы.

---

<sup>1</sup> Эти процессы бывают возможны в узкозонных полупроводниках, при довольно высоких температурах.

Положение уровня Ферми можно определить из (2.2.35), подставив в него  $n$  из (2.2.13) и прологарифмировав. Получим:

$$F = E_c + kT \ln \left\{ \frac{N_a + n_1}{2N_c} \left( \sqrt{1 + \frac{4n_1(N_d - N_a)}{(N_a + n_1)^2}} - 1 \right) \right\}. \quad (2.2.37)$$

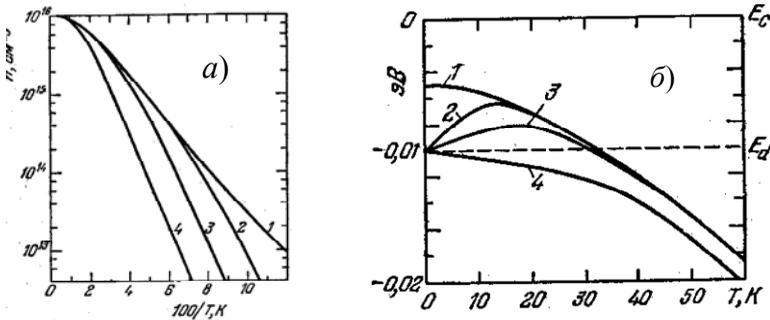


Рис. 2.2.4. Зависимость от обратной температуры а) – концентрации электронов, б) – положения уровня Ферми в частично компенсированном полупроводнике.

Ge, легированный донорами V группы ( $E_c - E_d \approx 0,01$  эВ) и частично компенсированный акцепторами III группы при разных степенях компенсации. Для всех кривых  $N_d - N_a = 10^{16} \text{ см}^{-3}$ .  $N_a =$  : 1 – 0, 2 –  $10^{14} \text{ см}^{-3}$ , 3 –  $10^{15} \text{ см}^{-3}$ , 4 –  $10^{16} \text{ см}^{-3}$ . В расчете принято  $m^*/m_0 = 0,25$ ,  $g_0/g_1 = 1/2$ .

При столь высоких температурах, что  $n_1 \gg N_a$ ,  $N_d - N_a$  получим:

$$F_{T\uparrow} = E_c + kT \ln \frac{N_d - N_a}{N_c}, \quad (2.2.38)$$

т.е. практически то же, то (2.2.31), но концентрация доноров соответственно уменьшена. С ростом температуры уровень Ферми

удаляется от с-зоны. А при низких температурах, когда  $n_1 \ll N_a$ ,  $N_d - N_a$ ,

$$F_{T\downarrow} = E_d + kT \ln \left( \frac{g_0}{g_1} \frac{N_d - N_a}{N_a} \right). \quad (2.2.39)$$

От (2.2.33) это отличается в первую очередь, тем, что при  $T \rightarrow 0$   $F \rightarrow E_d$ , а не к середине промежутка между донорами и зоной. Кроме того, с повышением температуры уровень Ферми движется в направлении, определяемом отношением  $\frac{N_d - N_a}{N_a}$ .

Если  $g_1 N_a < g_0 (N_d - N_a)$ , то вверх, к зоне. Если  $g_1 N_a > g_0 (N_d - N_a)$ , то – вниз.

На рис. 2.2.4 представлены расчетные температурные зависимости концентрации носителей и положения уровня Ферми в Ge при разных степенях компенсации.

#### 2.2.2.4. Полупроводник с многозарядными примесными центрами

Случаи, когда примесный центр может образовать в запрещенной зоне несколько состояний, различающихся зарядами, не редки. Классический пример – золото в Ge, приведен на рис. 2.1.10.

В таком случае количество реальных уровней не определено однозначно. Если центр нейтрален, он может захватить только один носитель, и с ним будет связана некоторая, вполне определенная энергия. Это – ситуация *a* на рис. 2.2.5. по мере продвижения уровня Ферми вверх (для многозарядных акцепторных примесей) заполняются сначала состояния с зарядом (-1), появляется возможность поместить на него еще один электрон, возникает второй уровень (ситуация *b*). И так далее.

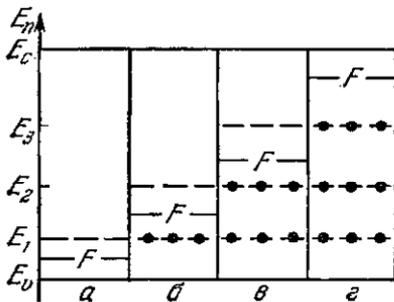


Рис. 2.2.5. Многозарядные центры.

Рассмотрим процесс перезарядки центра как процесс обмена электроном с некой средой, исходно имевшей  $N_0$  электронов и энергию  $E_0$ . Пусть центр захватил  $j$  частиц и находится в  $m$ -ом возбужденном состоянии.  $E_m^{(j)}$  - энергия этого состояния. Тогда при образовании этого состоя-

ния в среде осталось  $(N_0 - j)$  электронов, она имеет энергию  $(E_0 - E_m^{(j)})$ , а энтропия ее изменилась на  $\Delta S_m^{(j)}$ . Следовательно, вероятность этого процесса:

$$f_m^{(j)} = A \exp \frac{\Delta S_m^{(j)}}{k}. \quad (2.2.40)$$

Нужно определить  $\Delta S_m^{(j)}$ . Это можно сделать, рассмотрев свободную энергию Гиббса  $\mathcal{F} = E - ST$ . В нашем процессе она изменилась на:

$$\Delta \mathcal{F}_m^{(j)} = -E_m^{(j)} - T \Delta S_m^{(j)} \quad (2.2.41)$$

Будем считать, что среда достаточно велика для того, чтобы энергия, связанная с обменом одним электроном не зависела от  $j$ . Обозначим ее  $F$ :

$$\left. \frac{d\mathcal{F}}{dj} \right|_{V,T} = F. \quad (2.2.42)$$

Тогда  $T \Delta S_m^{(j)} = jF - E_m^{(j)} \quad (2.2.43)$

$$\text{и} \quad f_m^{(j)} = A \exp \frac{jF - E_m^{(j)}}{kT}. \quad (2.2.44)$$

Величину  $A$  можно найти из очевидного условия нормировки:

$$\sum_{j=0}^M \sum_m f_m^{(j)} = 1, \quad (2.2.45)$$

где  $M$  – максимальный заряд центра.

$$\frac{1}{A} = \sum_{j=0}^M \sum_m \exp \frac{jF - E_m^{(j)}}{kT} \quad (2.2.46)$$

Обозначим  $E_m^{(j)} = E^{(j)} + \varepsilon_m^{(j)}$  и  $f^{(j)} = \sum_m f_m^{(j)}$ , где  $\varepsilon_m^{(j)}$  – энергия  $m$ -го возбуждения центра, содержащего  $j$  электронов,  $f^{(j)}$  – вероятность образования  $j$ -зарядового состояния. И статвес фиксированных состояний  $\beta_m^{(j)}$ . Тогда эффективный статвес  $j$ -зарядового состояния:

$$g_j = \sum_m \beta_m^{(j)} \exp \left( -\frac{\varepsilon_m^{(j)}}{kT} \right). \quad (2.2.47)$$

Окончательно получаем вероятность образования  $j$ -зарядового состояния:

$$f^{(j)} = \frac{g_j \exp \frac{jF - E^{(j)}}{kT}}{\sum_{j=0}^M g_j \exp \frac{jF - E^{(j)}}{kT}} \quad (2.2.48)$$

### 2.2.3. Статистика электронов в квантовых ямах

Число электронов в подзоне можно определить, проинтегрировав по энергии произведение плотности состояний на функцию Ферми.

Плотность состояний в плоской подзоне постоянна и равна:

$$D(E) = \frac{m}{\pi\hbar^2},$$

так что поверхностная концентрация

$$n_s = \nu \frac{m}{\pi\hbar^2} \int_{E_1}^{\infty} \frac{dE}{\exp\left[\frac{E - F}{kT}\right] + 1} = N_c \ln \left\{ 1 + \exp\left[\frac{F - E_1}{kT}\right] \right\}, \quad (2.2.49)$$

где  $F$  – уровень Ферми,  $\nu$  – фактор долинного вырождения, количество долин. По аналогии с трехмерным кристаллом здесь введена эффективная величина

$$N_c = \nu \frac{mkT}{\pi\hbar^2} - \quad (2.2.50)$$

– двумерная эффективная плотность состояний. Она слабее зависит от температуры, чем в трехмерном случае, см. (2.2.11).

Выражение (2.2.49) справедливо для одной подзоны. Если имеется несколько подзон, их концентрации нужно суммировать.

В тех случаях, когда  $F - E_1 > 0$ , т.е. система вырождена, получаем, при  $T \rightarrow 0$ ,

$$n_s = \nu \frac{m}{\pi\hbar^2} (F - E_1) \quad (2.2.51)$$

и, поскольку на уровне Ферми имеются электроны, то их состояния можно характеризовать и волновым вектором Ферми  $k_F$ . Для чисто параболического, изотропного закона дисперсии получим:

$$k_F = \sqrt{\frac{2\pi n_s}{v}}. \quad (2.2.52)$$

В двумерной системе аналогом поверхности Ферми является кривая, или *контур Ферми*. В изотропном случае это – окружность в  $k$ -пространстве. Выражения (2.2.50) ÷ (2.2.52) сохраняются и для эллиптического контура Ферми, и в многодолинном полупроводнике, но в качестве  $m$  следует использовать (как и в трехмерном кристалле) *эффективную массу плотности состояний*. Для квантовых ям она равна:

$$m_d = \sqrt{m_x m_y}. \quad (2.2.53)$$

В отличие от (2.2.21), количество долин здесь в явном виде введено в (2.2.49) ÷ (2.2.52).

## 2.3. КИНЕТИЧЕСКИЕ ПРОЦЕССЫ В ПОЛУПРОВОДНИКАХ

### 2.3.1. Электропроводность

Мы рассматриваем вещества, в которых состояния валентных электронов могут рассматриваться как состояния свободных электронов в периодическом потенциальном поле и могут быть описаны Блоховской функцией:

$$\psi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r}) \exp(i\mathbf{k}\mathbf{r}). \quad (2.3.1)$$

Каждое состояние характеризуется квазиимпульсом  $\hbar\mathbf{k}$  и скоростью

$$\mathbf{v} = \nabla_{\mathbf{k}} \omega(\mathbf{k}) = \nabla_{\mathbf{p}} E(\mathbf{p}) \quad (= \mathbf{p}/m^*), \quad (2.3.2)$$

причем направления  $\mathbf{p}$  и  $\nabla_{\mathbf{p}} E(\mathbf{p})$  совсем не обязательно совпадают, т.е. направление квазиимпульса и скорости тоже могут быть различны. И обратная эффективная масса  $1/m^*$  – тензор (см. разд. 1.6.1.2 в [1]), так что последнее равенство в (2.3.2) применимо только в изотропных кристаллах.

В невозмущенном кристалле распределение носителей можно считать максвелловским с некоторыми средними величинами энергии, скорости, квазиволнового вектора. Но поскольку зона Бриллюэна центрально симметрична, средняя скорость всего ансамбля равна нулю. Электрический ток возникнет только в случае, если изменится именно эта средняя скорость всего ансамбля, скорость его дрейфа  $v_{dr}$ . При этом полностью справедливо рассмотрение, проведенное в разд. 1.6.4.2 для вырожденного ансамбля. Проводимость, т.е. коэффициент пропорциональности между полем  $\mathcal{E}$  и плотностью тока  $\mathbf{j}$ , зависит от концентрации носителей, их заряда, эффективной массы и времени свободного пробега  $\tau_p$ . С другой стороны, эта плотность тока равна произве-

дению плотности заряда (для электронов это  $-en$ ) на дрейфовую скорость его движения:

$$\mathbf{j} = \sigma \mathbf{E} = \frac{e^2 n \tau_p}{m^*} \mathbf{E} = en |v_{dr}|. \quad (2.3.3)$$

Здесь не обязательно учитывать знак заряда носителя. Если изменить знак, изменится и направление скорости, так что направление тока останется неизменным, по направлению поля.

Дрейфовая скорость пропорциональна напряженности поля. Коэффициент пропорциональности

$$\mu = \frac{|v_{dr}|}{E} = \frac{en \tau_p}{m^*} \quad (2.3.4)$$

называется *подвижностью носителей*.

Подвижность и проводимость  $\sigma$  также в общем случае – тензоры, так что в общем случае компоненты  $\mathbf{j}$  и  $\mathbf{E}$  связаны соотношениями:

$$\mathbf{j} = \sigma \mathbf{E} \Rightarrow \begin{cases} j_x = \sigma_{xx} E_x + \sigma_{xy} E_y + \sigma_{xz} E_z \\ j_y = \sigma_{yx} E_x + \sigma_{yy} E_y + \sigma_{yz} E_z \\ j_z = \sigma_{zx} E_x + \sigma_{zy} E_y + \sigma_{zz} E_z \end{cases} \quad (2.3.5)$$

или, сокращенно,

$$j_\alpha = \sigma_{\alpha\beta} E_\beta, \quad \alpha, \beta = x, y, z, \quad (2.3.6)$$

где, как обычно, подразумевается суммирование по повторяющимся индексам.

Наличие рассеяния приводит к тому, что в стационарных состояниях в выражении взаимосвязи внешней силы (силы электрического поля) и приращения импульса (среднего по ансамблю)

нужно учитывать и эффект торможения (трения) со стороны решетки. Т.е. в стационарных состояниях:

$$\mathbf{F} = \frac{d\mathbf{p}}{dt} + \mathbf{F}_{fr} = \frac{d\mathbf{p}}{dt} + \frac{\mathbf{p}}{\tau_p}. \quad (2.3.7)$$

Зависимость импульса от внешней силы в изотропном кристалле:

$$\frac{d\mathbf{p}}{dt} = \mathbf{F} - \frac{\mathbf{p}}{\tau_p} \quad \frac{d\mathbf{p}}{dt} = \mathbf{F} - \frac{\mathbf{p}}{\tau_p} \quad (2.3.8)$$

Время  $\tau_p$  определяет вероятность электрону иметь время свободного пролета в интервале от  $t$  до  $t + dt$ :

$$f(t)dt = \frac{dt}{\tau_p} \exp\left(-\frac{t}{\tau_p}\right). \quad (2.3.9)$$

### 2.3.1.1. Слабая локализация

Эффект слабой локализации – одно из проявлений волновых свойств электрона. Он возникает на фоне большого количества упругих рассеивателей, делающих его движение диффузионным. Электрон, находящийся в момент  $t = 0$  в начале координат  $\mathbf{r} = 0$ , при диффузионном движении со скоростью  $v$  и средней длиной свободного пробега  $l = v\tau$ , через время  $t \gg \tau$  окажется в точке  $\mathbf{r}$  с вероятностью

$$p(\mathbf{r}, t) = (4\pi Dt)^{-d/2} \cdot e^{-r^2/4Dt}, \quad (2.3.10)$$

где  $d$  – размерность пространства, в котором происходит движение,  $r^2 = \sum_1^d x_i^2$ ,  $D = lv/3$  – коэффициент диффузии.

Со временем ширина распределения постепенно увеличивается,

$$\Delta r \approx \sqrt{Dt} \approx l\sqrt{t/\tau} \approx l\sqrt{N}, \quad (2.3.11)$$

$N$  - число шагов в диффузионном процессе, т.е. число упругих столкновений за время  $t$ .

Это – диффузия классической частицы. Если учесть волновые свойства электрона, то точка  $\mathbf{r} = 0$  окажется выделенной. Из-за интерференции в ней сильно изменится функция  $p(\mathbf{r}, t)$ .

Разобьем все возможные траектории, возвращающие электрон в точку  $\mathbf{r} = 0$  в момент  $t$  на пары с одинаковым набором рассеивателей, но с противоположными направлениями движения. Для классической частицы вероятность  $p(0, t)$  есть сумма вероятностей прихода в точку  $\mathbf{r} = 0$  по разным траекториям. Для нее складываются квадраты модулей соответствующих волновых функций. Для квантовой частицы, сохранившей к моменту  $t \ll \tau_\phi$  память об исходной фазе, складываются сами волновые функции, а вероятность  $p(0, t)$  есть квадрат модуля результирующей волновой функции:

Без интерференции 
$$|\psi_1|^2 + |\psi_2|^2 = 2|\psi|^2,$$

С интерференцией 
$$|\psi_1 + \psi_2|^2 = |\psi_1|^2 + |\psi_2|^2 + 2|\psi_1\psi_2| = 4|\psi|^2.$$

Вероятность найти электрон в начале координат  $p(0, t)$  должна из-за интерференции удвоиться (рис. 2.3.1), причем ширина пика определяется соотношением неопределенности  $\delta r \approx \lambda \approx 1/k_F$  ( $\lambda$  – де-бройлевская длина волны,  $k_F$  – волновой вектор Ферми). Разбиение на интерферирующие пары возможно только для замкнутых траекторий, т.е. тех, что заканчиваются в точке  $\mathbf{r} = 0$ , в которой электрон находился в начальный момент времени.

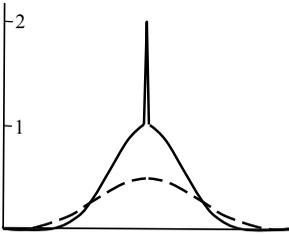


Рис. 2.3.1 *Пространственное распределение электронов с учетом интерференции, через  $t \ll \tau_\phi$  (сплошная линия) и через  $t \gg \tau_\phi$*

Увеличение вероятности для электрона оказаться в точке  $\mathbf{r} = 0$  (т.е. по существу остаться там) и называется слабой локализацией. Оно приводит к поправкам в проводимости  $\delta\sigma$ .

Относительная площадь под дополнительным пиком функции  $p(r, t)$  зависит от  $t$ . Поэтому для оценки относительной величины поправки к проводимости  $\delta\sigma/\sigma$  нужно проинтегрировать по времени.

В трехмерном случае,  $d = 3$ , объем пространства, в котором может находиться электрон в момент времени  $t$ , порядка  $(Dt)^{3/2}$ . Объем, из которого он за время  $dt$  может попасть в начало координат, порядка  $\lambda^2 v dt$ . Отношение этих объемов и определяет относительное количество электронов, побывавших в начале координат за время  $dt$ . Нижний предел интегрирования –  $\tau$ , минимальное время, через которое электрон может вернуться в начало координат, время упругого рассеяния. Верхний предел – максимальное время, через которое электрон, вернувшись, сможет участвовать в интерференции. Это время потери фазы  $\tau_\phi$ . В результате для  $d = 3$  получаем:

$$\begin{aligned} \frac{\delta\sigma}{\sigma} \Big|_3 &\approx - \int_{\tau}^{\tau_\phi} \frac{v\lambda^2 dt}{(Dt)^{3/2}} \approx - \frac{2v\lambda^2}{D^{3/2}} \left( \frac{1}{\tau^{1/2}} - \frac{1}{\tau_\phi^{1/2}} \right) = \\ &= -2 \left( k_F^2 l \right)^{-1} \left( \frac{1}{l} - \frac{1}{L_\phi} \right). \end{aligned} \quad (2.3.12)$$

Здесь величина  $L_\varphi \approx \sqrt{D\tau_\varphi} \approx l\sqrt{\tau_\varphi/\tau} \gg l$  называется диффузионной длиной.

Член с  $L_\varphi$  в выражении (2.3.12) меньше члена с  $l$ , но существенно в нем то, что именно он вносит температурную зависимость в квантовую поправку для проводимости (2.3.12). От температуры зависит  $\tau_\varphi$ .

От размерности  $d$  зависит нормировочный множитель в  $p(r, t)$  в выражении (2.3.10) и, соответственно, знаменатель в подинтегральном выражении в (2.3.12). Функции  $\delta\sigma/\sigma$  для разных  $d$  оказываются существенно различны. Характерным размером, сравнением с которым определяется размерность конкретного образца, является диффузионная длина  $L_\varphi$ .

Пленка толщиной  $b$  и проволока диаметром  $b$  при условии  $b \ll L_\varphi$  – объекты пониженной размерности. Вместо выражения (2.3.12) для них получим:

$$\left. \frac{\delta\sigma}{\sigma} \right|_2 \approx - \int_\tau^{\tau_\varphi} \frac{v\lambda^2 dt}{(Dt)b} \approx - \frac{v\lambda^2}{Db} \ln \frac{\tau_\varphi}{\tau}, \quad (2.3.13)$$

$$\left. \frac{\delta\sigma}{\sigma} \right|_1 \approx - \int_\tau^{\tau_\varphi} \frac{v\lambda^2 dt}{(Dt)^{1/2} b^2} \approx - \frac{2v\lambda^2}{Db^2} (L_\varphi - l). \quad (2.3.14)$$

Условие понижения размерности  $b \ll L_\varphi$  очень мягкое. Оно определяет конфигурацию пространства, в котором осуществляется диффузия. Формально даже очень толстая пленка при достаточно низкой температуре, т.е. больших длинах пробега, должна рассматриваться, как двумерная.

Формулы (2.3.12) – (2.3.14) показывают, что при любой размерности проводимость  $\sigma_i$  падает, начиная с некоторой температуры, с ее понижением (сопротивление растет), причем эффект

интерференции тем сильнее, чем ниже размерность. В трехмерном случае падение  $\sigma_3$  ограничено. При  $T \rightarrow 0$  оно стремится к определенному пределу. При пониженной размерности поправки расходятся, так как и  $\tau_\varphi$ , и  $L_\varphi$  стремятся к бесконечности при  $T \rightarrow 0$ .

Проводимость не может быть отрицательной, так что применимость формул (2.3.12) – (2.3.14) ограничивается требованием относительной малости поправок,  $\delta\sigma \ll \sigma$ .

Этот эффект объясняет логарифмический рост сопротивления многих металлов, в том числе – Cu и Au, при температурах 10–1 К на 3÷15%.

### 2.3.1.2. Отрицательное дифференциальное сопротивление в сильных полях. Диод Ганна

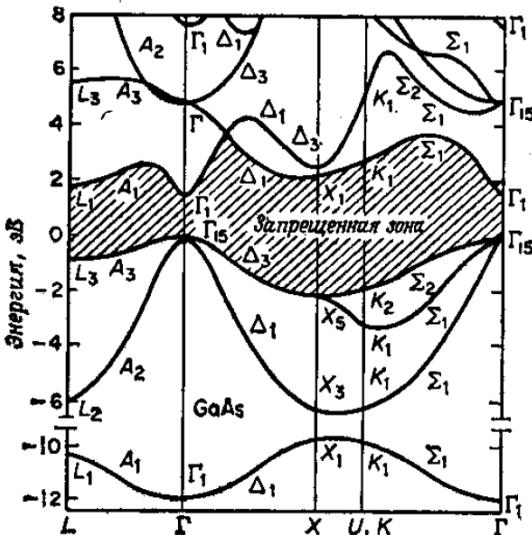


Рис. 2.3.2 Энергетическая диаграмма GaAs [16].

В сильных полях и в однородном образце возможно возникновение области отрицательного дифференциального сопротивления (ОДС). Например, в GaAs, на котором работают диоды Ганна. В этом случае природа ОДС принципиально отлична от диода Есаки (см. разд. 2.4.3.4). Она за-

ключается в особенностях свойств горячих носителей, обусловленных особенностями структуры разрешенных зон.

На рис. 2.3.2 представлена энергетическая диаграмма GaAs.

Экстремум валентной зоны расположен в центре зоны Бриллюэна, в точке  $\Gamma$ . Там же находится и минимум  $c$ -зоны. Но  $c$ -зона имеет еще два локальных экстремума, долины в точках  $L$  и  $X$ .

Энергии долин при 300 К:

$$E_{\Gamma} = 1,42 \text{ эВ}, E_L = 1,71 \text{ эВ}, E_X = 1,90 \text{ эВ}.$$

В  $\Gamma$ -долине изоэнергетическая поверхность сферична, эффективная масса электронов мала,  $m_{\Gamma} = 0.063m_0$  и подвижность велика. В  $L$ - и  $X$ -долинах изоэнергетические поверхности эллипсоидальны. Эффективные массы:  $m_l = 1.9m_0$ ,  $m_t = 0.075m_0$  – в  $L$ -долине и  $m_l = 1.9m_0$ ,  $m_t = 0.19m_0$  – в  $X$ -долине.

При 300 К  $L$ - и  $X$ -долины свободны. Для заселения  $L$ -долины нужна энергия возбуждения  $\delta E = 0,29$  эВ, более чем в 10 раз превышающая  $kT$ . Но в сильных полях за время свободного пробега электрон может набрать энергию, достаточную для междолинного переброса при рассеянии на  $LO$ -фононах. Этот процесс конкурирует с процессами внутридолинной релаксации и обратным переносом. Но здесь существенно то, что, во-первых, плотность состояний в  $L$ -долине больше, чем в  $\Gamma$ -долине, а во-вторых, имеются 8 идентичных  $L$ -долин, так что эффективная масса плотности состояний  $m_L = 0.85m_0$ .

В результате, по мере увеличения поля в  $n$ -GaAs начинается отток электронов из  $\Gamma$ - в  $L$ -долины, но там их подвижность существенно ниже. В результате средняя дрейфовая скорость падает, см. рис. 2.3.3

Предположим теперь, что к образцу приложено поле, близкое к критическому,  $E_c$ . Из-за флуктуаций электронной плотности в какой-либо части образца или на загибе зон в прикатодной области оно может стать несколько больше  $E_c$ . Тогда скорость

дрейфа в этой области упадет. Электроны справа будут "убегать" от нее, слева – накапливаться. При этом напряженность поля станет еще увеличиваться и образуется самоподдерживающаяся область пониженной электронной плотности. Ее называют **доменом**. Домен в образце может быть только один, т.к. на нем падает значительная доля приложенной разности потенциалов и в других областях поле меньше критического.

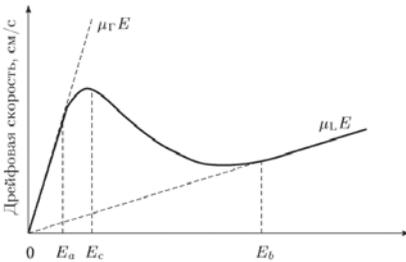


Рис. 2.3.3 Зависимость дрейфовой скорости электронов в GaAs от напряженности поля. [13]

Поскольку домен дрейфует во внешнем поле, то он исчезнет, дойдя до анода. При этом на аноде возникнут флуктуации протекающего тока, его последовательно будут достигать области пониженной и повышенной концентрации электронов.

По исчезновении домена тут же возникнет новый, и т.д. Во внешней цепи возникнут высокочастотные осцилляции, частота которых определяется дрейфовой скоростью электронов и толщиной (длиной) образца. При  $v_{dr} \approx 10^7$  см/с, в образце толщиной  $10^{-3}$  см частота должна быть порядка 10 ГГц.

Этот эффект был обнаружен в 1963г. Ганном и теперь промышленностью выпускаются диоды Ганна, генераторы сверхвысоких частот.

В GaAs *n*-типа, в поле  $E_c \approx 3$  кВ/см, скорость электронов составляет  $v \approx 10^7$  см/с, размер домена обычно составляет несколько микрон, поле в нём 40-200 кВ/см. При  $l = 1$  мм ÷ 5 мкм частота колебаний тока  $f=0,1-20$  ГГц. Рабочие напряжения – порядка 10 В, при этом амплитуды осцилляций – порядка 0,2 В.

## 2.3.2. Носители тока в магнитном поле

### 2.3.2.1. Циклотронный (диамагнитный) резонанс

При наличии магнитного поля полная сила, действующая на электрон, равна:

$$\mathbf{F} = -e(\mathbf{E} + [\mathbf{v}\mathcal{B}]). \quad (2.3.15)$$

Пусть  $\mathcal{E} = 0$ . Выберем систему координат так, чтобы  $\mathcal{B}_x = \mathcal{B}_y = 0$ ,  $\mathcal{B}_z \neq 0$  и устремим  $\tau_p \rightarrow \infty$ .

Уравнения движения электронов в таком случае запишутся в виде:

$$m_c \frac{\partial}{\partial t} \delta v_x = -e\mathcal{B} \cdot \delta v_y, \quad m_c \frac{\partial}{\partial t} \delta v_y = -e\mathcal{B} \cdot \delta v_x. \quad (2.3.16)$$

(здесь  $m_c$  – циклотронная эффективная масса). Их решения:

$$\delta v_x = v_0 \cos \omega_c t, \quad \delta v_y = v_0 \sin \omega_c t, \quad (2.3.17)$$

это – движение по окружности с частотой:

$$\omega_c \equiv \frac{e}{m_c} \mathcal{B}. \quad (2.3.18)$$

Она называется **циклотронной частотой**. Для свободного электрона она равна

$$f_c (\text{МГц}) = \frac{\omega_c}{2\pi} = 2,8 \cdot 10^{-4} B \quad (\text{тесла}).$$

Циклотронная частота не зависит от энергии, если эффективная масса постоянна. От энергии зависит только радиус его орбиты.

С такой частотой будет вращаться электрон по замкнутой орбите в том случае, если он не испытывает столкновений. Или ис-

пытывает, но  $\omega_c \tau_p \gg 1$ , т.е. за время свободного пробега он совершает несколько оборотов. Если же  $\omega_c \tau_p < 1$ , то периодического циклического движения не возникнет, электрон пройдет по круговой орбите лишь ограниченный угол,  $\alpha < 1$  рад.

*Резонансное поглощение высокочастотной энергии на частоте  $\omega_c$  называют циклотронным или диамагнитным резонансом.*

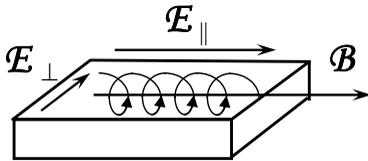


Рис. 2.3.4 *Схема наблюдения циклотронного резонанса.*

В металлах циклотронный резонанс можно наблюдать при низких температурах и в сильных полях, когда выполняется условие  $\omega_c \tau_p > 1$ . Схема эксперимента представлена на рис. 2.3.4. Радиусы циклотронных орбит в поле 1 Тесла – порядка

10 мкм, т.е. существенно больше толщины скинслоя. Поэтому высокочастотное поле  $E_{\perp}$  действует на электрон только на малом отрезке его траектории, близ поверхности. СВЧ энергия будет поглощаться в том случае, если она на каждом обороте электрона будет ускорять его. Для этого частота поля должна быть равна или кратна циклотронной частоте:

$$f = n \frac{\omega_c}{2\pi} = \frac{n}{2\pi} \frac{e}{m_c} B. \tag{2.3.19}$$

При фиксированной частоте  $f$ , максимум поглощения будет наблюдаться в магнитных полях:

$$B = \frac{2\pi f m_c}{en}. \tag{2.3.20}$$

Этот эффект может быть использован для изучения формы поверхности Ферми. Можно показать (см. [4], глава IV, §3 или

[5], глава 12), что циклотронная эффективная масса и, соответственно, частота определяются зависимостью от энергии площади сечения поверхности Ферми плоскостью, перпендикулярной магнитному полю:

$$m_c = \frac{\hbar^2}{2\pi} \frac{\partial S}{\partial E}, \quad \Rightarrow \quad \omega_c = \frac{2\pi eB}{\hbar^2} \frac{\partial E}{\partial S}. \quad (2.3.21)$$

### 2.3.2.2. Экстремальные орбиты

Для идеально сферической поверхности Ферми величина  $\partial S/\partial E$  постоянна для всех сечений сферы, но реально не бывает идеальных сфер. Это, казалось бы, должно "размазать" всю картину, обеспечить поглощение при любых частотах. Но оказывается, что в СВЧ-поглощении в данном случае существенны только так называемые *экстремальные орбиты*, те, у которых обращается в ноль производная от  $\partial S/\partial E$  по проекции  $k_B$ , вдоль магнитного поля. Другие орбиты не дают поглощения, взаимно компенсируя эффект.

Пример экстремальных орбит представлен на рис. 2.3.5 для поверхности Ферми благородных металлов, Cu, Ag, Au. Эти металлы имеют ГЦК-решетку, так что зона Бриллюэна у них имеет форму ячейки Вигнера-Зейтца ОЦК-структуры, см. рисунки 1.2.4 и 1.2.18 в [1]. У нее минимален размер в направлении (111), между противоположными шестиугольными гранями, между точками  $L$ . Из-за расщепления зон на границе зоны Бриллюэна в точке  $L$  энергия оказывается ниже, чем на поверхности сферы Ферми, так что эти

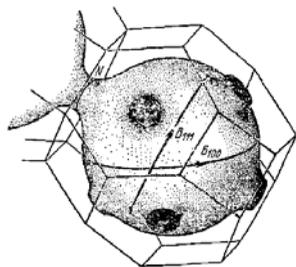


Рис. 2.3.5. Экстремальные орбиты на поверхности Ферми благородных металлов.

состояния оказываются заняты, в периодической зонной схеме отдельные сферы соединены перешейками. На рис. 2.3.5 показаны две экстремальные орбиты для направления магнитного поля (111) и одна для (100).

### 2.3.2.3. Электроны в скрещенных полях

Сила, действующая на заряд во взаимно перпендикулярных электрическом и магнитном полях, равна (2.3.15), так что ее проекции:

$$F_x = eE_x + e \frac{dy}{dt} \mathcal{B}, \quad F_y = eE_y - e \frac{dx}{dt} \mathcal{B}, \quad F_z = eE_z. \quad (2.3.22)$$

Магнитное поле направлено по оси  $z$ .

При свободном движении, между столкновениями, уравнения движения заряда:

$$\frac{d^2x}{dt^2} = \frac{e}{m} E_x + \omega_c \frac{dy}{dt}, \quad \frac{d^2y}{dt^2} = \frac{e}{m} E_y - \omega_c \frac{dx}{dt}, \quad \frac{d^2z}{dt^2} = \frac{e}{m} E_z, \quad (2.3.23)$$

где  $\omega_c \equiv \frac{e}{m} \mathcal{B}$  – циклотронная частота.

Как начальные условия примем при  $t = 0$ :

$$x = y = z = \frac{dx}{dt} = \frac{dy}{dt} = \frac{dz}{dt} = 0. \quad (2.3.24)$$

Решения уравнений (2.3.23): последнего -

$$z = \frac{e}{m} E_z t^2, \quad (2.3.25)$$

первых двух -

$$\begin{aligned} x - a &= -a \cos \omega_c t - b \sin \omega_c t + b \omega_c t, \\ y - b &= -b \cos \omega_c t + a \sin \omega_c t - a \omega_c t, \end{aligned} \quad (2.3.26)$$

$$\text{где} \quad a = \frac{e}{m\omega_c^2} \mathcal{E}_x, \quad b = \frac{e}{m\omega_c^2} \mathcal{E}_y. \quad (2.3.27)$$

В этом можно убедиться непосредственной подстановкой.

Два первых члена в уравнениях (2.3.26) описывают движение по окружности радиуса

$$r = \sqrt{a^2 + b^2} = \frac{e}{m\omega_c^2} \sqrt{\mathcal{E}_x^2 + \mathcal{E}_y^2} \quad (2.3.28)$$

с центром в точке  $(a, b)$ . Последние слагаемые в (2.3.26) описывают дрейф заряда в плоскости  $xу$ , перпендикулярной магнитному полю, со скоростью  $v_t$ . Ее проекции:

$$v_{tx} = b\omega_c = \frac{e}{m\omega_c} \mathcal{E}_y = \frac{\mathcal{E}_y}{\mathcal{B}}, \quad v_{ty} = -a\omega_c = -\frac{e}{m\omega_c} \mathcal{E}_x = -\frac{\mathcal{E}_x}{\mathcal{B}}. \quad (2.3.29)$$

Не трудно убедиться, что эта скорость перпендикулярна электрическому полю.

Итого, получится движение по циклоиде, изображенной на рис. 2.3.4.

Для того, чтобы определить реальные величины средних скоростей электронов при ограниченном времени свободного пробега  $\tau_p$ , нужно определить средние смещения и разделить их на  $\tau_p$ .

Средние смещения определим из (2.3.25), (2.3.26), умножив их на вероятность (2.3.9) испытать столкновение через время  $t$  и проинтегрировав по времени. Здесь появятся интегралы:

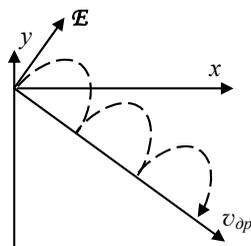


Рис. 2.3.6 Траектория движения заряда во взаимно перпендикулярных электрическом и магнитном полях.

$$\int_0^{\infty} \exp\left(-\frac{t}{\tau_p}\right) \sin \omega_c t \cdot \frac{dt}{\tau_p} = \frac{\omega_c \tau_p}{1 + \omega_c^2 \tau_p^2}$$

и

$$\int_0^{\infty} \exp\left(-\frac{t}{\tau_p}\right) \cos \omega_c t \cdot \frac{dt}{\tau_p} = \frac{1}{1 + \omega_c^2 \tau_p^2}.$$

Используя их, получим из (2.3.25), (2.3.26):

$$\bar{v}_x = \frac{e}{m} \mathcal{E}_x \frac{\tau_p}{1 + \omega_c^2 \tau_p^2} + \omega_c \frac{e}{m} \mathcal{E}_y \frac{\tau_p^2}{1 + \omega_c^2 \tau_p^2}, \quad (2.3.30a)$$

$$\bar{v}_y = \frac{e}{m} \mathcal{E}_y \frac{\tau_p}{1 + \omega_c^2 \tau_p^2} - \omega_c \frac{e}{m} \mathcal{E}_x \frac{\tau_p^2}{1 + \omega_c^2 \tau_p^2}, \quad (2.3.30б)$$

$$\bar{v}_z = \frac{e}{m} \mathcal{E}_z \tau_p. \quad (2.3.30в)$$

Эти выражения справедливы при известном  $\tau_p$ , но время рассеяния зависит от энергии электронов, так что (2.3.30а,б,в) нужно усреднить по всему ансамблю. Для этого можно решить механическую задачу о вероятности соударения, но сейчас, для описания проводимости, нам достаточно ввести эти усредненные величины. Обозначим их  $\zeta_1$ ,  $\zeta_2$ :

$$\zeta_1 \equiv \left\langle \frac{\tau_p}{1 + \omega_c^2 \tau_p^2} \right\rangle, \quad \zeta_2 \equiv \left\langle \frac{\tau_p^2}{1 + \omega_c^2 \tau_p^2} \right\rangle. \quad (2.3.31)$$

При этих обозначениях для компонент плотности тока получим:

$$j_x = en \langle v_x \rangle = \frac{e^2 n}{m} (\zeta_1 \mathcal{E}_x + \omega_c \zeta_2 \mathcal{E}_y) \quad (2.3.32a)$$

$$j_y = en \langle v_y \rangle = \frac{e^2 n}{m} (-\omega_c \zeta_2 \mathcal{E}_x + \zeta_1 \mathcal{E}_y) \quad (2.3.32б)$$

$$j_z = en \langle v_z \rangle = \frac{e^2 n}{m} \langle \tau_p \rangle E_z. \quad (2.3.32\text{в})$$

Следовательно, тензор электропроводности (см. (2.3.5)) в этом случае имеет вид:

$$\begin{aligned} \sigma_{xx} &= \frac{e^2 n}{m} \zeta_1 & \sigma_{xy} &= \omega_c \frac{e^2 n}{m} \zeta_2 & \sigma_{xz} &= 0 \\ \sigma_{yx} &= -\omega_c \frac{e^2 n}{m} \zeta_2 & \sigma_{yy} &= \frac{e^2 n}{m} \zeta_1 & \sigma_{yz} &= 0 \\ \sigma_{zx} &= 0 & \sigma_{zy} &= 0 & \sigma_{zz} &= \frac{e^2 n}{m} \langle \tau_p \rangle \end{aligned} \quad (2.3.33)$$

В отсутствие магнитного поля  $\omega_c = 0$  и  $\zeta_1 \equiv \langle \tau_p \rangle$ , так что все недиагональные члены равны нулю, а диагональные равны между собой. Получаем известную статическую электропроводность:

$$\sigma = \frac{e^2 n}{m} \langle \tau_p \rangle \quad (2.3.34)$$

и подвижность

$$\mu = \frac{e}{m} \langle \tau_p \rangle. \quad (2.3.35)$$

#### 2.3.2.4. Эффект Холла

На рис. 2.3.7 показана схема эксперимента по наблюдению эффекта Холла в веществе с электронной проводимостью. К прямоугольному образцу прикладываются взаимно перпендикулярные электрическое и магнитное поля. На рисунке магнитное поле направлено от нас.

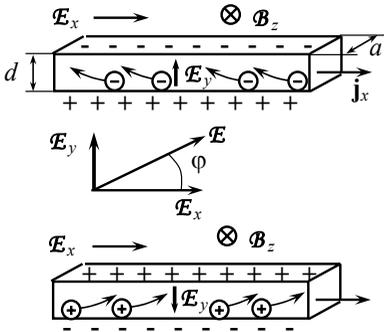


Рис. 2.3.7. Эффект Холла – возникновение поперечной эдс в скрещенных полях. Знак эдс определяется типом носителей.

Направление поля и знак угла зависят от знака заряда носителей.

На рисунке изображен характер движения носителей обоих знаков. Угол Холла показан только для движения электронов.

Опыт показывает, что поле  $E_y$  пропорционально плотности тока  $j_x$  и магнитному полю  $B$ :

$$E_y = \frac{U_y}{d} = R B j_x = R B \frac{i}{ad}, \quad (2.3.36)$$

где  $R$  – **постоянная Холла**. В слабых полях, удовлетворяющих условию

$$\omega_c \tau_p \ll 1, \quad (2.3.37)$$

она не зависит от магнитного поля и характеризует исследуемое вещество.

На электрон действует сила Лоренца, так что по мере набора скорости в электрическом поле он все эффективнее смещается вверх (на рисунке). Это приводит к возникновению электрического поля  $E_y$ , перпендикулярного двум приложенным полям, и разности потенциалов на боковых сторонах образца  $U_y = d \cdot E_y$ . Электрическое поле в образце поворачивается на угол  $\phi$ ,

При  $j_y = 0$  (цепь разомкнута) из (2.3.32б) и (2.3.33) для угла Холла получим:

$$tg\varphi = \frac{E_y}{E_x} = -\frac{\sigma_{yx}}{\sigma_{yy}} = \frac{\sigma_{xy}}{\sigma_{xx}} = \omega_c \frac{\zeta_2}{\zeta_1} = \frac{e}{m} \frac{\zeta_2}{\zeta_1} \mathcal{B} \equiv \mu_H \mathcal{B}, \quad (2.3.38)$$

где введена новая величина, **подвижность Холла**  $\mu_H$ . Эта величина действительно имеет размерность подвижности, но она не равна подвижности, определенной в (2.3.4), как величина, пропорциональная среднему  $\tau_p$ . Холловская подвижность, см. (2.3.31), иначе связана с временем рассеяния импульса:

$$\mu_H = \frac{e}{m} \frac{\zeta_2}{\zeta_1} = \frac{e}{m} \frac{\langle \tau_p^2 \rangle}{\langle \tau_p \rangle}. \quad (2.3.39)$$

Используя (2.3.38) можно найти выражение для постоянной Холла, введенной в (2.3.36). Выразим  $E_x$  через  $E_y$  и компоненты проводимости. Исключив в (2.3.32а)  $E_x$ , получим:

$$E_y = \frac{\sigma_{xy}}{\sigma_{xx} + \sigma_{xy}^2} j_x, \quad (2.3.40)$$

следовательно:

$$R = \frac{1}{\mathcal{B}} \frac{\sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2}. \quad (2.3.41)$$

В слабых полях, удовлетворяющих условию (2.3.37), можно считать, что

$$\sigma_{xy} \ll \sigma_{xx} \approx \sigma = e n \mu.$$

Тогда (2.3.41) можно привести к виду:

$$R \approx \frac{\sigma_{xy}}{\mathcal{B}\sigma^2} = \frac{tg\varphi}{\mathcal{B}\sigma} = \frac{\mu_H}{\mu} \frac{1}{en} = \gamma \frac{1}{en}. \quad (2.3.42)$$

Здесь  $\gamma$  - "холловский фактор", отношение холловской и обычной подвижности. Если он известен, постоянная Холла позволяет определить концентрацию носителей.

Холловский фактор можно определить, зная удельную проводимость  $\sigma$  и подвижность  $\mu$ , поскольку:

$$R\sigma = \mu_H. \quad (2.3.43)$$

Знание  $\gamma = \frac{\langle \tau_p^2 \rangle}{\langle \tau_p \rangle^2}$  позволяет определить характер рассеяния носителей, поскольку механизм рассеяния определяет и зависимость  $\tau_p(E)$ , т.е. различие величин  $\langle \tau_p^2 \rangle$  и  $\langle \tau_p \rangle^2$ .

### 2.3.2.5. Магнетосопротивление

Сопротивление проводников в поперечном магнитном поле возрастает.

Примем, как и прежде, что  $\mathcal{E}_z, j_y = 0$  и магнитное поле направлено по оси  $z$ . Тогда из соотношений (2.3.32а,б) получим:

$$j_x = \frac{\sigma_{xx}^2 + \sigma_{xy}^2}{\sigma_{xx}} \mathcal{E}_x. \quad (2.3.44)$$

Следовательно, проводимость в поперечном магнитном поле равна:

$$\sigma_{\perp}(\mathcal{B}) = \frac{\sigma_{xx}^2 + \sigma_{xy}^2}{\sigma_{xx}}. \quad (2.3.45)$$

Но соотношения (2.3.33), (2.3.42) получены и верны для замкнутых орбит, электронных и дырочных.

### 2.3.2.6. Эффект де Гааза-ван Альфена

В 1930 г. де Гааз и ван Альфен наблюдали [6] осцилляции намагниченности  $M$  (отношения  $M/B$ ) висмута при изменении магнитного поля. Эффект проявлялся только при низкой температуре, в сильных полях. В том же 1930-м г. Ландау дал объяснение этого эффекта, имеющего чисто квантовую природу. Существенно позднее, в 1952 г., Онзагер показал, что эти осцилляции дают возможность определить параметры (площадь) экстремальных орбит (см. разд. 2.3.2.2). Окончательно теория этого эффекта разработана в 1953-1955 гг. [7].

Для простоты *рассмотрим двумерную систему* в магнитном поле, перпендикулярном ее плоскости. И пусть напряженность поля и степень совершенства структуры (определяющая вероятность рассеяния электронов и время  $\tau_p$ ) будут достаточно высоки для выполнения условия  $\omega_c \tau_p \gg 1$ .

В этом случае в отсутствие электрического поля электроны вращаются с частотой  $\omega_c$  вокруг неподвижных точек. При наличии электрического поля  $\mathcal{E}_\perp$ , перпендикулярного  $\mathcal{B}$ , они так же вращаются вокруг точек, дрейфующих перпендикулярно и  $\mathcal{B}$ , и  $\mathcal{E}_\perp$  (разд. 2.3.2.3). Такое круговое движение можно рассматривать как сумму двух линейных осцилляторов с частотой  $\omega_c$ .

Решения уравнения Шредингера для линейного осциллятора известны (см., например, [8], §25):

$$E_n = \hbar\omega_c \left( n + \frac{1}{2} \right). \quad (2.3.46)$$

Получается, что в сильных магнитных полях принципиально изменяются и характер движения, и законы квантования элект-

тронных состояний. Поступательное движение не возможно и  $k$  больше не является определяющим квантовым числом. Электроны движутся по замкнутым орбитам, и квантуется их энергия. Причем шаг квантования,  $\hbar\omega_c$ , существенно превышает энергетические интервалы между блоховскими состояниями в том же макрокристалле вне поля.

На каждом энергетическом уровне (они называются уровнями Ландау) может находиться некоторое количество электронов, которое можно оценить, зная распределение плотности состояний двумерного газа и шаг квантования<sup>1</sup>. Количество состояний на одном уровне, или фактор вырождения  $\xi$ , – число состояний, приходящихся на энергетический интервал  $\hbar\omega_c$  в образце единичной площади, одинаковое для всех орбит. Без учета спинового вырождения (оно снимается в магнитном поле) получим:

$$\xi = \frac{m}{2\pi\hbar^2} \frac{\hbar e}{m} \mathcal{B} = \frac{e}{2\pi\hbar} \mathcal{B} = \frac{1}{l_B^2}. \quad (2.3.47)$$

Здесь  $l_B = \sqrt{\frac{2\pi\hbar}{e\mathcal{B}}}$  – *магнитная длина Ландау*. Можно показать, что это – радиус классической орбиты, соответствующей первому уровню Ландау. Величина  $\frac{2\pi\hbar}{e}$  – *квант потока*, минимальная величина магнитного потока через циклотронную орбиту.

Принципиальным моментом здесь является степень заполненности уровней Ландау. При нулевой температуре будут заняты все нижние состояния, и если число электронов на единицу

---

<sup>1</sup> Это интуитивный подход к вычислению фактора вырождения, но точный расчет дает тот же результат, см. [5], глава 14.

площади образца  $n$  не кратно  $\xi$ , то на верхнем останется  $\text{frac}\left(\frac{n}{\xi}\right)$  электронов.

Но  $\xi$  зависит от  $\mathcal{B}$  и, следовательно, изменяя  $\mathcal{B}$ , мы будем периодически переводить систему в состояния с любой степенью заполненности верхнего уровня, уровня Ферми. Уровень Ферми будет целиком заполнен, если выполняется условие  $n = \xi\lambda$ , где  $\lambda$  – целое. Это состояние будет воспроизводиться с периодом:

$$\Delta\left(\frac{1}{\mathcal{B}}\right) = \frac{\pi\hbar}{ne}. \quad (2.3.48)$$

Де Гааз и ван Альфен наблюдали изменения намагниченности образца висмута. Зависимость магнитной проницаемости от степени заполнения верхнего энергетического состояния можно пояснить, без математических вычислений, следующими соображениями.

Свободный электронный газ обладает ларморовским диамагнетизмом. В магнитном поле электроны прецессируют, частично компенсируя внешнее поле.

Парамагнетизм свободного электронного газа поясняет рис. 2.3.8. В магнитном поле между идентичными состояниями, различающимися только спином, возникает разность энергий

$\Delta E = g\mu_B\mathcal{B} = \frac{g}{2}\hbar\omega_c$ , в результате чего количество электронов со спином по полю увеличивается, против – уменьшается.

При сложении этих двух эффектов парамагнетизм доминирует. Но если верхние

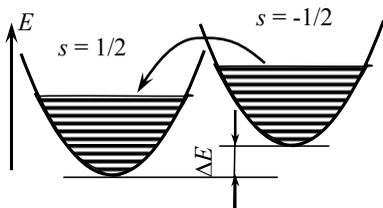


Рис. 2.3.8 К объяснению парамагнетизма свободного электронного газа.

по энергии состояния полностью заняты, то подобная "перекачка", изменение спина части ансамбля, не возможна. Парамагнетизм исключается, образец становится диамагнитным.

*В трехмерной системе* к приведенным соображениям добавляется третья степень свободы,  $k_z$  и связанная с ним энергия. В принципе, описанный эффект должен смазаться за счет вовлечения этой третьей координаты в распределение электронов по состояниям. Энергия, связанная с движением по магнитному полю, не изменяется. Правда, если электроны находятся в состояниях, в которых изменение энергии не связано с  $z$ -координатой, то все приведенные выше соображения остаются в силе, как для 2D-систем. Онзагер показал, что осцилляции должны наблюдаться на орбитах, для которых  $\frac{\partial E}{\partial k_z} = 0$ , т.е. на экстремальных орбитах.

При этом период повторения определяется площадью экстремальных орбит  $S_e$ :

$$\Delta\left(\frac{1}{\mathcal{B}}\right) = \frac{e}{\hbar} \frac{1}{S_e}. \quad (2.3.49)$$

Как правило, при фиксированной ориентации  $\mathcal{B}$  наблюдаются несколько экстремальных орбит. Поэтому в наблюдаемом спектре будут присутствовать несколько частот.

Эффект изменения структуры поверхности Ферми будет проявляться во всех свойствах, для которых существенно перераспределение электронов по спиновым состояниям и просто плотность состояний на уровне Ферми. Так, в [5], глава 14 приведены спектры зависимости от  $\mathcal{B}$  различных свойств, в том числе и термоэдс, и эффекта Пельтье, которые мы будем рассматривать.

### 2.3.2.7. Двумерные структуры. Квантовый эффект Холла

Квантовый эффект Холла (КЭХ, см. [9]; [10], гл.5; [13], §9.6) был обнаружен в 1980 г. и уже в 1985 г. Фон Клитцинг получил

за него Нобелевскую премию. Столь быстрое признание работы объясняется тем, что это уникальное явление – чисто квантовый эффект, в котором проявляются тонкие детали свойств электронных состояний. Но, вероятно, главным все-таки явилось то, что он дал прекрасную возможность точного измерения величины одной из фундаментальных постоянных,  $\frac{e^2}{\hbar}$ , а с ней – и постоянной тонкой структуры. В системе СГСЭ постоянная тонкой структуры равна:

$$\alpha = \frac{e^2}{\hbar c} \approx 137,04. \quad (2.3.50)$$

Сейчас она измерена с точностью до  $10^{-7}$ .

Первая работа [11], в которой был описан КЭХ, и называлась "Новый метод точного определения постоянной тонкой структуры...".

Суть эффекта заключается в том, что в двумерной системе, в которой  $\mathcal{B}$  направлено поперек слоя, энергии и плотность состояний подвижных носителей квантованы и, за отсутствием 3-й координаты, не уширены в подзоны. Эти кванты известны из теории, развитой в связи с обнаружением эффекта де Гааза – ван Альфена, см. выражения для циклотронной частоты, (2.3.18), энергии квантованных состояний, (2.3.46), плотности состояний, (2.3.47).

Энергетические интервалы между квантовыми состояниями,  $\hbar\omega_c = \frac{\hbar e}{m_c} \mathcal{B}$ , оказываются достаточно велики для того, чтобы при низких температурах<sup>1</sup> было возможно создать ситуацию, когда

---

<sup>1</sup> Для наблюдения КЭХ желательны температуры ниже 1 К. При более высоких становится существенным рассеяние на фононах.

несколько квантовых состояний полностью заняты, остальные свободны<sup>1</sup>. В этих случаях точно известна концентрация носителей,

$$n = l\xi = l \frac{e}{2\pi\hbar} \mathcal{B} = l \frac{e}{h} \mathcal{B}, \quad (2.3.51)$$

где  $l = 1, 2, 3, \dots$  – целое число, равное количеству заполненных уровней.

Следовательно, точно определена и величина постоянной Холла (2.3.42):

$$R = \frac{1}{en} = \frac{1}{l} \frac{h}{e^2} \frac{1}{\mathcal{B}}, \quad (2.3.52)$$

и сопротивление Холла (см. (2.3.39)):

$$\rho_{xy} = \frac{\mathcal{E}_y}{j_x} = R\mathcal{B} = \frac{1}{l} \frac{h}{e^2}, \quad (2.3.53)$$

а проводимость, величина, обратная  $\rho_{xy}$ , равна<sup>2</sup>  $l \frac{e^2}{h}$ .

На рис. 2.3.9 показан чертеж образца и схема измерений для наблюдения квантового эффекта Холла. Через образец обычно пропускается постоянный стабилизированный ток  $J_{12}$ . Электроды 3, 4, 5 и 6 – измерительные.

---

<sup>1</sup> Это так называемый "целочисленный" или "классический КЭХ". Здесь количество занятых состояний  $\nu \geq 1$ . При  $\nu < 1$ , говорят о "дробном КЭХ". Его природу и суть проявляемых свойств электронного ансамбля мы рассматривать не будем. С ним можно ознакомиться по специальным изданиям, например, [9].

<sup>2</sup> Это – "целочисленный КЭХ". Эффекты заполнения квантовых состояний при  $l < 1$  – "дробный КЭХ" (см. [10]). Мы его рассматривать не будем.

Выполнить равенство (2.3.51) можно либо в постоянном магнитном поле, изменяя концентрацию носителей, либо, при неизменной концентрации, изменяя магнитное поле. Плотность состояний (2.3.47) на квантовом уровне пропорциональна магнитному полю, так что увеличивая его, можно уменьшать количество уровней, необходимое для размещения имеющихся электронов.

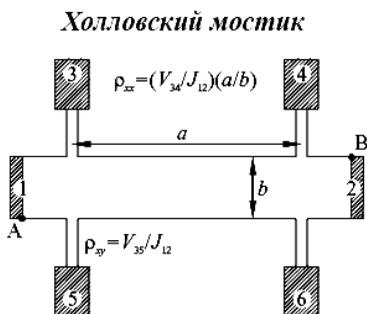


Рис. 2.3.9. Схема измерения квантового эффекта Холла.

Первые опыты по наблюдению КЭХ были проведены в МОП – структуре, в которой потенциалом затвора можно было управлять, изменяя концентрацию электронов в квантовой яме на границе раздела полупроводник/окисел. Энергетическая диаграмма такой системы изображена на рис. 2.3.10а. Концентрация носителей в слое пропорциональна емкости перехода и потенциалу затвора  $V_3$  :

$$n = \frac{C_0 V_3}{e}. \quad (2.3.54)$$

Подробнее характеристики такой системы будут рассмотрены в следующем разделе.

Наблюдения КЭХ при изменении магнитного поля проводятся на слоях с фиксированными концентрациями носителей, созданных на гетеропереходах, рис. 2.3.10б.

Вид ожидаемых зависимостей сопротивления Холла от заполнения состояний поясняет рис. 2.3.11. Принципиальным моментом здесь является рассеяние носителей. На рисунке показаны

положения уровней Ландау и плотности состояний: идеальная и реальная, с учетом рассеяния на нарушениях структуры слоя (беспорядка).

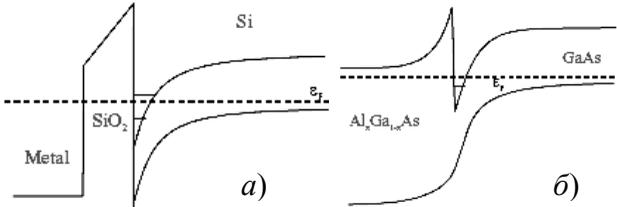


Рис. 2.3.10. Энергетическая диаграмма а) - МОП-системы, к затвору которой приложен потенциал, достаточный для образования вырожденного слоя. б) – гетероперехода p-GaAs/n-Al<sub>x</sub>Ga<sub>1-x</sub>As.

Обозначены также уровни размерного квантования по z-координате.

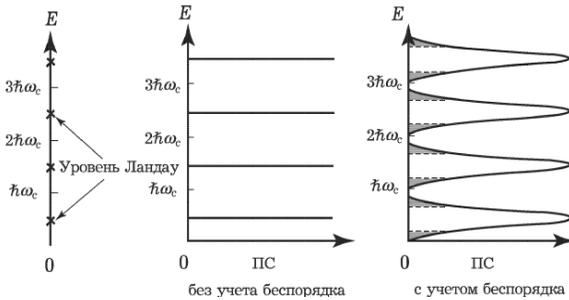


Рис. 2.3.11. Уровни Ландау в квантовой яме. На энергетической оси в единицах  $\hbar\omega_c$  показаны положения уровней Ландау и плотности состояний (ПС) в идеальном и реальном случаях.

Рассеяние приводит к расширению бесконечно узких уровней. Если  $\tau_p$  – время рассеяния импульса, то, согласно соотношению неопределенности, полуширина полосы составит:

$$\Delta E = \frac{\hbar}{\tau_p}. \quad (2.3.55)$$

Существенно, что возникновение состояний на периферии изображенных колоколов мало вероятно, т.е. эти состояния локализованы в пространстве. Эти состояния на рис. 2.3.11 заштрихованы. Резкую границу между обобществленными состояниями и локализованными провести не возможно, но главное – что есть и те, и другие.

На рис. 2.3.11 увеличение концентрации носителей изображится смещением вверх уровня Ферми, разделяющего заполненные и пустые состояния. Если он находится где-то в одной из заштрихованных областей, то изменение его положения влияет только на концентрацию локализованных электронов. Концентрация подвижных остается неизменной. При этом не изменяется и величина холловского сопротивления. В результате на спектрах зависимости  $\rho_{xy}$  от концентрации (от потенциала затвора МОП-структуры) или от напряженности магнитного поля возникают плато, на которых выполняется соотношение (2.3.53).

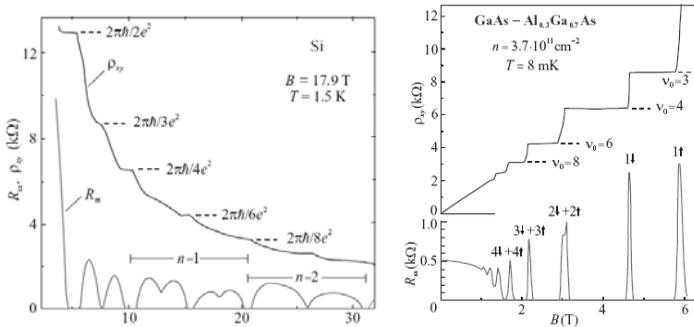


Рис. 2.3.12. Вид спектров КЭХ в двух режимах – в зависимости от концентрации носителей в постоянном поле и при определенной концентрации от напряженности поля.

Спектры, полученные авторами КЭХ для МОП-структуры в Si [11] и на гетеропереходе GaAs/Al<sub>0,3</sub>Ga<sub>0,7</sub>As [12], представлены на рис. 2.3.12.

Подробнее с рассмотрением основ теории целочисленного КЭХ можно ознакомиться по [13], разд. 9.6.

## 2.4. ЯВЛЕНИЯ НА КОНТАКТАХ

В этом разделе мы рассмотрим общие закономерности, проявляющиеся в невырожденных полупроводниках.

### 2.4.1. Контактная разность потенциалов, загиб зон

В разделе 2.2 мы видели, что легированием в полупроводнике можно существенно изменять положение уровня Ферми в запрещенной зоне. Практически – в пределах всей зоны. Следовательно, так же изменяется и термоэлектронная работа выхода  $\Phi_T$ , см. [1], разд. 1.6.1.3 и рис. 2.1.5. Получается, что от полупроводника  $n$ -типа легче оторвать электрон, чем при проводимости  $p$ -типа.

У металлов работы выхода определены более однозначно, но они существенно различаются для разных металлов. У большинства металлов в середине таблицы Менделеева  $\Phi_T = 4-5$  эВ. У платины  $\Phi_T = 5,3$  эВ, у щелочных металлов они меньше, почти до 2 эВ.

На рис. 2.4.1 изображены два образца, металл и полупроводник, с различными работами выхода. Пока они оба электронейтральны, энергия электронов в вакууме, обозначенная как "0", будет одинакова близ обеих поверхностей. Но энергия связи электронов в полупроводнике меньше, чем в металле. Поэтому, если теперь обеспечить между ними контакт по какой-либо внешней цепи, то электроны полупроводника начнут перетекать в металл. Там их энергия, за счет связи с полным ансамблем частиц, меньше, чем в полупроводнике. Это и есть причина различия работ выхода.

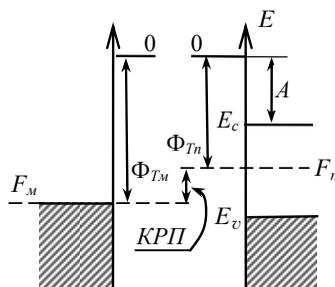


Рис. 2.4.1.

Металл начнет заряжаться отрицательно, полупроводник – положительно. К изображенным на рис. 2.4.1 внутренним энергиям добавятся энергии  $-\epsilon\phi$ . И этот наведенный потенциал  $\phi$  одинаково сместит все энергетические уровни образцов. Ни ширина запрещенной зоны, ни работа выхода, ни величина сродства  $A$ , ни расстояние уровня Ферми от краев зон никак не изменятся.

Процесс перетекания электронов будет продолжаться до тех пор, пока полные энергии электронов (средние) в образцах не выровняются. При этом уровни Ферми у них сравняются, но энергия электронов в вакууме близ их поверхностей изменится. Между поверхностями возникнет разность потенциалов, которая называется **КОНТАКТНОЙ РАЗНОСТЬЮ ПОТЕНЦИАЛОВ** и обозначена на рисунке 2.4.2 как КРП. Получим энергетическую диаграмму рис. 2.4.2.

Существенно, что при такой зарядке можно пренебречь изменением концентраций электронов в образцах. Так, емкость кристалла размером  $1 \text{ см}^3$  – порядка  $10^{-12}$  Ф. Для зарядки на 1 В потребуется  $\approx 10^7$  электронов. А собственных электронов в нем  $\approx 10^{23}$ . Концентрация носителей обычно больше, чем  $10^{15} \text{ см}^{-3}$ .

#### 2.4.1.1. Измерение КРП

Возникшая разность потенциалов измерима и подобные измерения часто используются для быстрого определения работы выхода неизвестного образца, если есть откалиброванный образец, с известной работой выхода.

Система, изображенная на рис. 2.4.2 – плоский конденсатор с некоторой, не нулевой разностью потенциалов  $u_k$ . Следовательно, на его обкладках присутствуют наведенные заряды  $Q = C u_k$ .

При изменении емкости  $C$  будет изменяться и этот заряд, т.е. какое-то количество электричества перетечет по внешней цепи. Емкость можно изменить просто изменив расстояние между по-

верхностями. Если один из электродов сделать вибрирующим, то во внешней цепи возникнет переменный ток на частоте вибрации.

Но во внешнюю цепь можно вставить источник регулируемой эдс. Если подобрать ее равной КРП и соответственно выбрать полярность (на рисунке к полупроводнику нужно приложить "-"), то поле в зазоре будет скомпенсировано, зарядов на конденсаторе не возникнет и изменение емкости не вызовет тока во внешней цепи.

Это – "*метод вибрирующего конденсатора*".

#### 2.4.1.2. *Приповерхностный загиб зон на контакте с металлом*

На рис. 2.4.2 предполагается, что расстояние между металлом и полупроводником достаточно для того, чтобы полями заряженных тел можно было пренебречь. И не учтен наведенный заряд конденсатора. Но, если концентрация электронов близ поверхности уменьшилась, то, значит, увеличилось расстояние от уровня Ферми до дна с-зоны, обозначенное в (2.2.9) как  $\zeta \equiv F - E_c$ .

При плотном контакте поверхностей это изменение будет равно КРП. Локализованный близ поверхности заряд компенсирует поле металла, так что в глубине полупроводника относительное расположение зон и уровня Ферми останется таким же, как на рис. 2.4.2. Но толщина приповерхностной области, в которой будет накоплен этот компенсирующий заряд, будет тем больше, чем меньше концентрация носителей в полупроводнике.

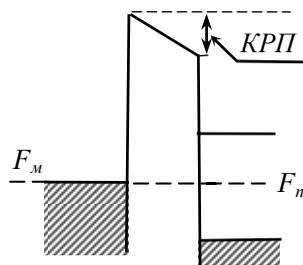


Рис. 2.4.2. *Контактная разность потенциалов равна разности работ выхода контактирующих тел.*

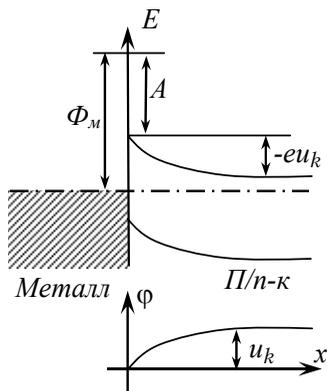


Рис. 2.4.3.

можно пренебречь.

После установления равновесия получим картину, изображенную на рис. 2.4.3. Уровни Ферми одинаковы в обоих телах, в глубине полупроводника энергетическая диаграмма осталась неизменной, но относительно металла полупроводник заряжен до потенциала  $\varphi = u_k = \Phi_{met} - \Phi_{sc}$ . Близ поверхности потенциал зависит от координаты и в результате возникает картина, которую называют **загиб зон**.

В данном случае – **загиб зон вверх** в полупроводнике *n*-типа. Такой поверхностный слой обеднен носителями, поэтому такой контакт и называют **обедненным**.

В общем случае разность работ выхода металла и полупроводника может быть и положительной (как на приве-

Подобная картина должна была бы наблюдаться и в металле, но в металлах высокая концентрация электронов и внешнее поле может быть экранировано на малых расстояниях. Расчетные значения глубин экранирования  $d_{met}$  меньше моноэлемента. Поскольку потоки индукции  $\mathcal{D}$  через поверхности металла и полупроводника одинаковы, то падением потенциала в приповерхностной области металла (порядка  $\Delta u \approx \mathcal{D}d_{met}$ )

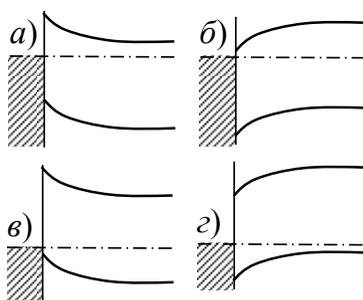


Рис. 2.4.4. Основные типы областей загиба зон.

денных рисунках), и отрицательной. Полупроводник при этом может быть и  $n$ -, и  $p$ -типа. Так что, в принципе, возможны 4 типа поверхностных слоев на контакте с металлом, изображенные на рис. 2.4.4. Здесь  $a)$  и  $b)$  –  $n$ -полупроводник, соответственно **обедненный** и **обогащенный** слои;  $в)$  и  $г)$  –  $p$ -полупроводник, соответственно **обогащенный** и **обедненный** слои. В слоях  $a)$  и  $в)$  – **загиб зон вверх**,  $b)$  и  $г)$  – **загиб зон вниз**.

### 2.4.1.3. Соотношение Эйнштейна

Равновесие систем, изображенных на рис. 2.4.3, 2.4.4, - динамическое. Носители подвижны, так что при наличии поля и градиента концентрации возникнут токи. Соответственно, дрейфовый и диффузионный:

$$j_n = j_{dr} + j_{diff} = en\mu_n E + eD_n \nabla n \quad (2.4.1a)$$

$$j_p = j_{dr} + j_{diff} = ep\mu_p E - eD_p \nabla p. \quad (2.4.1б)$$

Для того, чтобы потоки, вызванные градиентом концентрации можно было рассматривать как диффузионные, пользоваться коэффициентами диффузии  $D_n$  и  $D_p$  и описывать токи соотношениями (2.4.1), должно выполняться условие:

$$|\nabla n|l_n \ll n, \quad |\nabla p|l_p \ll p \quad (2.4.2)$$

где  $l_n, l_p$  – длины свободного пробега электрона и дырки.

Нужно оценить, при каких напряженностях поля близ поверхности эти условия выполняются.

Обозначим  $n_0$  и  $E_{c0}$  – собственную концентрацию электронов и энергию дна  $s$ -зоны вдали от контакта. Потенциал объема примем за ноль, потенциал произвольной точки в области контакта -  $\phi$ .

Тогда в невырожденном полупроводнике концентрация электронов:

$$\begin{aligned} n(x) &= N_c \exp \frac{F - E_c(x)}{kT} = \\ &= N_c \exp \frac{F - [E_{c0} - e\varphi(x)]}{kT} = n_0 \exp \frac{e\varphi(x)}{kT}, \end{aligned} \quad (2.4.3)$$

градиент концентрации:

$$\nabla n = \frac{e}{kT} \nabla \varphi \cdot n_0 \exp \frac{e\varphi}{kT} = -\frac{e}{kT} n \mathcal{E}, \quad (2.4.4)$$

и окончательно условие (2.4.2) примет вид:

$$\frac{\nabla n}{n} = -\frac{e}{kT} \mathcal{E} \Rightarrow \left| \frac{e}{kT} l_n \mathcal{E} \right| \ll 1, \quad |e l_n \mathcal{E}| \ll kT \quad (2.4.5)$$

На длине пробега в приповерхностном слое электрон должен набирать энергию, много меньше  $kT$ . Здесь  $\mathcal{E}$  - некоторое среднее поле в области загиба зон.

Будем считать, что это условие выполняется.

Если полупроводник далек от вырождения и соотношения (2.4.3), (2.4.4) выполняются, то из (2.4.1), (2.4.3), (2.4.4) и из факта равновесия на контакте ( $j = 0$ ) получаем, что подвижность и коэффициент диффузии не независимы, а связаны соотношением:

$$\frac{\mu_n}{D_n} = \frac{e}{kT}. \quad (2.4.6a)$$

Аналогично для дырок мы получили бы

$$\frac{\mu_p}{D_p} = \frac{e}{kT}. \quad (2.4.6б)$$

Эти соотношения называются *соотношения Эйнштейна*.

Можно получить более общую форму записи этого соотношения, применимую и в случае, если полупроводник близок к вырождению.

Для химического потенциала (2.2.9) справедливо:

$$\zeta = F - (E_c - e\varphi) = \zeta_0 + e\varphi, \quad \zeta^* = \zeta_0^* + \frac{e\varphi}{kT}. \quad (2.4.7)$$

Тогда градиент концентрации можно выразить в виде:

$$\nabla n = \frac{dn}{d\zeta^*} \nabla \zeta^* = \frac{dn}{d\zeta^*} \frac{e}{kT} \nabla \varphi = - \frac{dn}{d\zeta^*} \frac{e}{kT} \mathcal{E}. \quad (2.4.8)$$

И из (2.4.1) получим:

$$j_n = en\mu_n \mathcal{E} + eD_n \nabla n = en\mu_n \mathcal{E} - eD_n \frac{dn}{d\zeta^*} \frac{e}{kT} \mathcal{E} = 0, \quad (2.4.8a)$$

$$\frac{\mu_n}{D_n} = \frac{e}{kT} \frac{dn}{nd\zeta^*} = \frac{e}{kT} \frac{d(\ln n)}{d\zeta^*}. \quad (2.4.9)$$

Для дырок мы получили бы:

$$\frac{\mu_p}{D_p} = \frac{e}{kT} \frac{dp}{pd\eta^*} = \frac{e}{kT} \frac{d(\ln p)}{d\eta^*}. \quad (2.4.10)$$

#### 2.4.1.4. Условия равновесия контактирующих тел

Поскольку уровень (энергия) Ферми – средняя энергия электронов в среде, естественно ожидать, что наличие или отсутствие потоков электронов (при наличии свободных носителей) определяется именно градиентом уровня Ферми.

Это действительно так, что легко показать, подставив (2.4.8) и (2.4.8a) в (2.4.1a). Получим:

$$j_n = n\mu_n \left( -e\nabla\varphi + \frac{e}{kT} \frac{D_n}{\mu_n} \frac{d(\ln n)}{d\zeta^*} \nabla\zeta \right). \quad (2.4.11)$$

Но 
$$\frac{e}{kT} \frac{D_n}{\mu_n} \frac{d(\ln n)}{d\zeta^*} = 1,$$

и отсюда:

$$j_n = n\mu_n \nabla(\zeta - e\varphi) = n\mu_n \nabla F. \quad (2.4.12a)$$

Для потока дырок мы так же получили бы:

$$j_p = p\mu_p \nabla F \quad (2.4.12б)$$

Этот результат очень существенен для дальнейшего: **поток носителей тока определяется, кроме их концентрации и подвижности, градиентом уровня Ферми.**

### 2.4.2. Слой объемного заряда

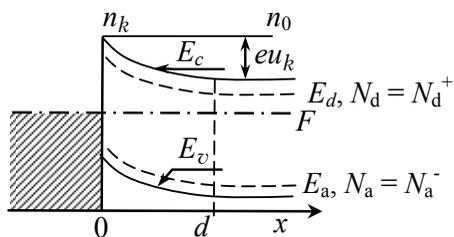


Рис. 2.4.5. Загиб зон на контакте с металлом. Обедненный слой в *n*-полупроводнике. Здесь *d* – толщина слоя, в котором существен объемный заряд.

Теперь нужно найти зависимость  $\varphi(x)$  и  $n(x)$  близ поверхности, на контакте с металлом.

Будем считать, что все имеющиеся примеси достаточно мелки и полностью ионизованы во всей области объемного заряда. Следовательно, объемный заряд в приповерхностном слое обеспечен только

изменением концентрации носителей. Для определенности примем, что наш образец имеет электронную проводимость. Концентрация электронов в объеме, вдали от контакта, равна  $n_0$ , а концентрация дырок пренебрежимо мала. Тогда объемный заряд будет равен:

$$\rho = e(N_d^+ - N_a^- - n). \quad (2.4.13)$$

или, поскольку мы считаем, что все примеси остаются ионизованными:

$$\rho(\infty) = 0, \quad \rho(x) = e[n_0 - n(x)]. \quad (2.4.14)$$

Концентрация электронов на поверхности в отсутствие тока, при  $j = 0$  равна (в случае, далеком от вырождения):

$$n(0) = n_k = n_0 \exp\left(-\frac{eu_k}{kT}\right). \quad (2.4.15)$$

При прохождении тока уже нельзя считать, что концентрация на поверхности,  $n(0)$ , остается прежней. Но можно оценить, в каких случаях с достаточной точностью можно считать это отклонение пренебрежимо малым.

Отсутствие тока,  $j = 0$ , не означает, что контактирующие тела не обмениваются электронами. На границу раздела из полупроводника в металл и обратно текут одинаковые токи:

$$|j_1| = |j_2| = \frac{1}{4} en_k v_T. \quad (2.4.16)$$

Наблюдаемый ток – результат нарушения этого баланса,

$$|j| = |j_1 - j_2| = \frac{ev_T}{4} [n_k - n(0)]. \quad (2.4.17)$$

Если эта разность мала по сравнению с каждым из токов, то, очевидно, ею действительно можно пренебречь. Оценка для концентрации носителей  $10^{13} \text{ см}^{-3}$  дает для  $j_1, j_2$  величину порядка  $10 \text{ А}\cdot\text{см}^{-2}$ . А концентрация  $10^{13} \text{ см}^{-3}$  соответствует расстоянию между уровнем Ферми и ближайшей зоной более  $(12 \div 13)kT$ . При комнатной температуре это  $0,3 \div 0,325 \text{ эВ}$ . Так что при плотностях тока порядка  $0,1 \text{ А}\cdot\text{см}^{-2}$  наше предположение о том, что  $n(0) = n_k$  вполне оправдано.

Тогда граничные условия для нашей задачи, если к переходу приложено внешнее напряжение  $u$ , можно записать в виде:

$$\begin{aligned} \varphi(0) &= 0, & \varphi(\infty) &= u_k + u, \\ n(0) &= n_k, & n(\infty) &= n_0. \end{aligned} \quad (2.4.18)$$

Для решения задачи запишем уравнение для плотности тока:

$$j_n = en\mu_n E + \mu_n kT \frac{\partial n}{\partial x}, \quad (2.4.19)$$

уравнение непрерывности:

$$\frac{\partial \rho}{\partial t} = - \frac{\partial j}{\partial x} \quad (2.4.20)$$

и уравнение Пуассона:

$$\frac{\partial E}{\partial x} = \frac{4\pi\rho}{\varepsilon}, \quad \frac{\partial^2 \varphi}{\partial x^2} = - \frac{\partial E}{\partial x} = \frac{4\pi e}{\varepsilon} [n - n_0]. \quad (2.4.21)$$

Для концентрации в произвольной точке, очевидно, справедливо:

$$n(x) = n_k \exp\left(\frac{e\varphi(x)}{kT}\right) = n_0 \exp\left(\frac{e[\varphi(x) - u_k]}{kT}\right). \quad (2.4.22)$$

Используя это соотношение и (2.4.15), перепишем (2.4.21) в виде:

$$\begin{aligned} \frac{\partial^2 \varphi}{\partial x^2} &= \frac{4\pi e}{\varepsilon} \cdot n_k \cdot \left[ \exp \frac{e\varphi}{kT} - \exp \frac{eu_k}{kT} \right] = \\ &= \frac{4\pi e}{\varepsilon} \cdot n_0 \cdot \left[ \exp \frac{e(\varphi - u_k)}{kT} - 1 \right]. \end{aligned} \quad (2.4.23)$$

### 2.4.2.1. Дебаевская длина экранирования

При малом загибе зон, если  $\left| \frac{eu_k}{kT} \right| \ll 1$  и, следовательно,  $\left| \frac{e\varphi}{kT} \right| \ll 1$ , экспоненты в (2.4.23) можно разложить в ряд и ограничиться первыми членами:

$$\begin{aligned} \exp \frac{e\varphi}{kT} &\approx 1 + \frac{e\varphi}{kT}, & \exp \frac{eu_k}{kT} &\approx 1 + \frac{eu_k}{kT}, \\ \exp \frac{e(\varphi - u_k)}{kT} &\approx 1 + \frac{e(\varphi - u_k)}{kT}. \end{aligned}$$

Тогда из (2.4.23) получим:

$$\frac{\partial^2 \varphi}{\partial x^2} = \frac{4\pi e}{\varepsilon} \cdot n_0 \cdot \left[ \exp \frac{e(\varphi - u_k)}{kT} - 1 \right] \approx \frac{\varphi - u_k}{L_D^2}, \quad (2.4.24)$$

где 
$$L_D^2 \equiv \frac{\varepsilon kT}{4\pi e^2 n_0}. \quad (2.4.25)$$

Решение (2.4.25) имеет вид:

$$\varphi = u_k \left[ 1 - \exp \left( -\frac{x}{L_D} \right) \right]. \quad (2.4.26)$$

Величина  $L_D$  имеет размерность длины и определяет скорость экспоненциального затухания потенциала в пространстве, с удалением от границы раздела. Она называется *дебаевская длина экранирования*.

Длина экранирования растет с уменьшением концентрации носителей. При  $n_0 = 10^{15} \text{ см}^{-3}$  она равна примерно 100 нм.

#### 2.4.2.2. Перезарядка примесей в слое объемного заряда

Согласно (2.4.25),  $L_D \rightarrow \infty$  при  $n_0 \rightarrow 0$ . Но мы не учитывали, что возможно изменение заполненности примесных центров. А это должно сделать экранирование более эффективным.

Предположим, что мы имеем  $n$ -полупроводник, в котором акцепторные примеси находятся вдали от уровня Ферми и незначительное изменение этого расстояния ничего принципиально не изменят. А изменение заряда на донорах нужно учесть в выражении для объемного заряда.

Ранее, (2.2.26), мы получили выражение для степени занятости состояний. Для заряда на донорах справедливо:

$$N_d^+ = N_d \frac{n_1}{n + n_1}, \quad (2.4.27)$$

где 
$$n_1 = \frac{g_0}{g_1} N_c \exp\left(-\frac{J}{kT}\right). \quad (2.4.28)$$

тогда плотность заряда:

$$\rho = e \left( N_d \frac{n_1}{n + n_1} - N_a - n \right), \quad (2.4.29)$$

а вдали от контакта:

$$\rho(\infty) = e \left( N_d \frac{n_1}{n_0 + n_1} - N_a - n_0 \right) = 0. \quad (2.4.30)$$

Считая, как и прежде, что загиб зон мал настолько, что

$$n \approx n_0 \left( 1 + \frac{e(\varphi - u_k)}{kT} \right),$$

получим:

$$\frac{n_1}{n + n_1} \approx \frac{n_1}{n_0 + n_1} \left( 1 + \frac{e(\varphi - u_k)}{kT} \frac{n_1}{n_0 + n_1} \right). \quad (2.4.31)$$

Тогда из (2.4.29), с учетом (2.4.30), плотность заряда:

$$\rho = \frac{e^2 n_0}{kT \psi^2} (\varphi - u_k), \quad (2.4.32)$$

где 
$$\psi^2 = \frac{n_0 + n_1}{N_a + n_1 + 2n_0}. \quad (2.4.33)$$

Теперь уравнение Пуассона (2.4.24) приходит к виду:

$$\frac{\partial^2 \varphi}{\partial x^2} \approx \frac{\varphi - u_k}{L_D^2 \psi^2}, \quad (2.4.34)$$

т.е. длина экранирования изменяется:

$$L_{\varphi} = L_D \psi \quad (2.4.35)$$

И существенную роль здесь играет концентрация акцепторов, которые являются компенсирующей примесью.

Для оценки степени влияния этого эффекта рассмотрим несколько крайних случаев.

Если  $N_a, n_0 \ll n_1$ , т.е. примесей обоих типов мало, то  $\psi = 1$ , они не могут изменить длину экранирования.

Если образец сильно компенсирован, так что  $N_a \gg n_0$ , то уровень Ферми находится близ уровня доноров и  $n_0 \approx n_1$ . При этом из (2.4.34) получаем:

$$\psi^2 = \frac{2n_0}{N_a + 3n_0} \approx \frac{2n_0}{N_a} \ll 1 \quad (2.4.36)$$

и длина экранирования становится равна:

$$L_{\psi}^2 \equiv \frac{\varepsilon kT}{2\pi e^2 N_a}. \quad (2.4.37)$$

В таком случае она полностью определяется перезарядкой примесей.

### 2.4.2.3. Обогащенный слой

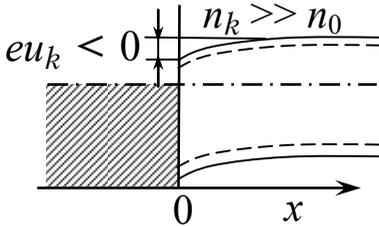


Рис. 2.4.6. Обогащенный слой в  $n$ -полупроводнике. Обозначения те же, что на рис. 2.4.5.

Мы рассмотрим распределение потенциала и концентрации носителей в отсутствие тока. Прохождение тока через такой контакт – см. разд. 2.4.3.1.

Обогащенный слой в  $n$ -полупроводнике изображен на рис. 2.4.6. зоны загнуты вниз, концентрация на поверхности много больше объемной. Поэтому

можно считать, что в приповерхностном слое концентрация электронов велика по сравнению с объемной, т.к.:

$$\exp \frac{e\phi}{kT} \gg \exp \frac{eu_k}{kT}. \quad (2.4.38)$$

Тогда (2.4.23) существенно упрощается:

$$\frac{\partial^2 \varphi}{\partial x^2} = \frac{4\pi e}{\varepsilon} \cdot n_k \cdot \exp \frac{e\varphi}{kT}. \quad (2.4.39)$$

Домножив это равенство на  $\frac{\partial \varphi}{\partial x}$  и проинтегрировав по  $x$ , получим:

$$\left( \frac{\partial \varphi}{\partial x} \right)^2 = \frac{8\pi n_k kT}{\varepsilon} \cdot \exp \frac{e\varphi}{kT} + C. \quad (2.4.40)$$

Поскольку вдали от перехода  $\frac{\partial \varphi}{\partial x} = 0$ , и  $\varphi = u_k$ , постоянная интегрирования равна:

$$C = -\frac{8\pi n_k kT}{\varepsilon} \cdot \exp \frac{eu_k}{kT}, \quad (2.4.41)$$

т.е., в силу нашего допущения (2.4.38), ею можно пренебречь. Остается:

$$\frac{\partial \varphi}{\partial x} = -\sqrt{\frac{8\pi n_k kT}{\varepsilon}} \cdot \exp \frac{e\varphi}{2kT}. \quad (2.4.42)$$

Положительный корень не соответствует начальным условиям, согласно которым потенциал уменьшается с удалением от поверхности. Интегрируя это равенство, получаем:

$$e\varphi = -2kT \ln \left( 1 + \frac{x}{a} \right), \quad (2.4.43)$$

где 
$$a = \sqrt{\frac{\varepsilon kT}{2\pi e^2 n_k}}. \quad (2.4.44)$$

Потенциал зависит от расстояния логарифмически и простирается на глубину примерно на расстояние  $a$ . По форме выражение для  $a$  – то же, что и для дебаевской длины экранирования, но здесь фигурирует  $n_k$ , а не  $n_0$ , так что глубины, на которые про-

стирается приповерхностное поле, в этом случае существенно меньше. Концентрация электронов тоже определяется величиной  $n_k$ , а не  $n_0$ :

$$n = n_k \exp\left(\frac{e\varphi}{kT}\right) = n_k \left(\frac{a}{a+x}\right)^2. \quad (2.4.45)$$

Вдали от контакта, где  $\varphi = u_k$ , получим:

$$n = n_k \exp\left(\frac{eu_k}{kT}\right) = n_0. \quad (2.4.46)$$

Реально на поверхности даже высокоомного полупроводника можно создать хорошую проводимость. Этим пользуются, когда надо подвести к полупроводнику контакт, не вносящий искажений в электрическую цепь. Обогащенный слой не изменит сопротивления последовательной цепи. Поэтому такие контакты называются *антизапорными*.

#### 2.4.2.4. Обедненный (истощенный) слой

Энергетическая диаграмма обедненного слоя в  $n$ -полупроводнике представлена на рис. 2.4.5. В отличие от предыдущего случая такой контакт называют *запорным*.

Предположим, что к контакту металл/полупроводник приложено и внешнее напряжение  $u$ , обедняющее приповерхностный слой. В этом случае  $eu_k > 0$  и

$$\exp\frac{e(\varphi - u_k)}{kT} \ll 1 \quad (2.4.47)$$

При этом уравнение Пуассона (2.4.23) совсем упрощается:

$$\frac{\partial^2 \varphi}{\partial x^2} = -\frac{4\pi e}{\varepsilon} \cdot n_0 = Const \quad (2.4.48)$$

Проинтегрируем это уравнение, и для определения постоянных интегрирования учтем граничные условия:

$$\varphi(0) = 0; \quad \varphi(d) = u + u_k, \quad \left. \frac{d\varphi}{dx} \right|_{x=d} = 0. \quad (2.4.49)$$

Здесь мы считаем, что область объемного заряда равна  $d$  (см. рис. 2.4.5) и потенциал во всей этой области изменяется плавно.

Решение (2.4.48) дает:

$$-\frac{\partial\varphi}{\partial x} = \mathcal{E} = -\frac{4\pi en_0}{\varepsilon}(d-x), \quad (2.4.50)$$

$$\varphi = u + u_k - \frac{2\pi en_0}{\varepsilon}(d-x)^2 \quad (2.4.51)$$

и, поскольку  $\varphi(0) = 0$ , из последнего получаем глубину экранирования:

$$d = \sqrt{\frac{\varepsilon(u_k + u)}{2\pi en_0}}. \quad (2.4.52)$$

В нашей модели, следовательно, поле в области перехода падает линейно, а потенциал – квадратично (при отсчете от края области перехода).

Существенно, что здесь все экранирование обеспечивается локальными заряженными центрами. Величина  $en_0$  в (2.4.52) равна их суммарному заряду в объеме полупроводника.

#### **2.4.2.5. *p-n-переход***

До сих пор мы рассматривали только контакты с металлом, что позволило нам упростить задачу, оставить в рассмотрении лишь один из контактирующих материалов. Теперь можно рассмотреть контакт двух полупроводников с различной проводимостью.

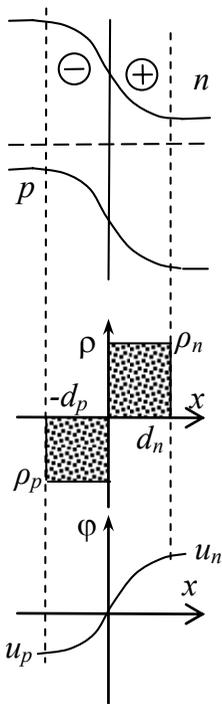


Рис. 2.4.7.

*p-n-переход.*

Предположим, что этот контакт – плотный и контактирующие *n*- и *p*-полупроводники созданы легированием, на одной основе. Например, оба – легированием кремния. В этом случае не только одинаковы ширины запрещенных зон, но и положения краев зон относительно уровня вакуума одинаковы (до контакта). Следовательно, разность энергий в зонах может определяться только наведенным или приложенным извне потенциалом. И, что нам тоже поможет, одинаковы диэлектрические проницаемости.

Здесь так же, как и на контакте с металлом, происходит перенос заряда с одного контактера на другой, причем в переходной области *p-n*-перехода всегда возникает обедненный слой. Его называют **запорным**.

Будем рассматривать резкий переход, т.е. такой, при котором концентрации легирующих примесей скачком изменяются на контакте, а далее не изменяются.

Обозначим толщины запорных слоев как  $d_p$  и  $d_n$ .

Равновесные объемные концентрации носителей:

$$p_0 \approx N_a - N_d, \text{ (при } x < -d_p) \text{ и } n_0 \approx N_d - N_a \text{ (при } x > d_n).$$

Тогда объемный заряд близ контактов равен  $-ep_0$  и  $en_0$  в этих областях, соответственно. Граничные условия в такой системе:

$$\begin{aligned} \varphi(d_n) = u_n, \quad \varphi(-d_p) = u_p, \quad \left. \frac{d\varphi}{dx} \right|_{d_n} = \left. \frac{d\varphi}{dx} \right|_{-d_p} = 0, \\ \varphi_n(0) = \varphi_p(0) = 0 \end{aligned} \quad (2.4.53)$$

Здесь  $u_p$ ,  $u_n$  - падения напряжения соответственно в  $p$ - и  $n$ -областях.

К обеим сторонам  $p$ - $n$ -перехода приложимы результаты предыдущего раздела, поэтому сразу можно написать:

$$\varphi_p = u_p + \frac{2\pi e p_0}{\varepsilon} (x + d_p)^2, \quad (2.4.54)$$

$$\varphi_n = u_n - \frac{2\pi e n_0}{\varepsilon} (x - d_n)^2$$

и 
$$u_n - u_p = u + u_k = \frac{2\pi e}{\varepsilon} \{ p_0 d_p^2 + n_0 d_n^2 \}. \quad (2.4.55)$$

Далее, поскольку электрическая индукция на контакте непрерывна и мы считаем, что  $\varepsilon$  с обеих сторон равны, то:

$$\mathcal{E}_n(0) = \left. \frac{d\varphi}{dx} \right|_{x=0} = \mathcal{E}_p(0) = \left. \frac{d\varphi}{dx} \right|_{x=0},$$

откуда, дифференцируя (2.4.54), получаем:

$$p_0 d_p = n_0 d_n. \quad (2.4.56)$$

Теперь можно из (2.4.55) и (2.4.56) определить полную ширину перехода:

$$d = d_p + d_n = \left( \frac{\varepsilon(u + u_k)}{2\pi e} \frac{p_0 + n_0}{p_0 n_0} \right)^{1/2}. \quad (2.4.57)$$

Как и следовало ожидать, полная ширина перехода при существенном неравенстве равновесных концентраций определяет-

ся той стороной, в которой концентрация меньше. Так, если  $p_0 \gg n_0$ ,

$$d \approx d_n = \left( \frac{\varepsilon(u + u_k)}{2\pi e} \frac{1}{n_0} \right)^{1/2}. \quad (2.4.58)$$

Используя равенство (2.4.56) и полагая в (2.4.54)  $x = 0$ , получим соотношение падений напряжения в обеих областях:

$$\left| \frac{u_p}{u_n} \right| = \frac{d_p}{d_n} = \frac{n_0}{p_0}. \quad (2.4.59)$$

Если переход не резок, разность концентраций примесей изменяется вдоль координаты  $x$  по какому-то плавному закону, то (2.4.57), очевидно, будет не справедливо. Для того, чтобы рассчитать ширину перехода и распределение заряда и потенциала в ней, нужно точно знать этот закон.

#### 2.4.2.6. Емкость $p$ - $n$ -перехода

Из (2.4.57) видно, что ширина перехода зависит от напряжения, приложенного извне. Но с шириной будет изменяться и заряд, накопленный в переходе:

$$|q| = ep_0d_p = en_0d_n = \frac{\varepsilon(u + u_k)}{2\pi} \frac{1}{d} = \left( \frac{\varepsilon e(u + u_k)}{2\pi} \frac{p_0n_0}{p_0 + n_0} \right)^{1/2}. \quad (2.4.60)$$

Электрическая емкость такого перехода – производная от заряда по напряжению:

$$C = \frac{dq}{du} = \left( \frac{\varepsilon e}{8\pi(u + u_k)} \frac{p_0n_0}{p_0 + n_0} \right)^{1/2} = \frac{\varepsilon}{4\pi d}, \quad (2.4.61)$$

как в плоском конденсаторе.

Этот эффект изменения емкости перехода при изменении поданного напряжения используется в электронных приборах – варикапах, на которых, в частности, построены все системы сканирования частот в радио- и телеприемниках.

Выражение (2.4.61) получено и справедливо только для резкого перехода. В плавных переходах зависимость  $C(u)$  будет иной. Ее можно измерить и получить тем самым дополнительную информацию о структуре перехода.

### 2.4.3. Прохождение токов через контакт

#### 2.4.3.1. Антισταпорный слой. Токи, ограниченные пространственным зарядом

Рассмотрим антισταпорный контакт на границе  $n$ -полупроводника с металлом. На рис. 2.4.8 изображен ход энергии дна  $s$ -зоны в таком переходе без внешнего поля и при разностях потенциалов, приложенных в направлении, вытягивающем основные носители из металла в полупроводник.

Самый существенный для нас момент отмечен на кривой, соответствующей потенциалу  $u_1$ . Во всех случаях, вычисляя проходящие токи, мы должны учитывать две их составляющих, дрейфовую и диффузионную. Очевидно, вдали от поверхности электрическое поле вытягивает электроны и дрейфовая составляющая тока вели-

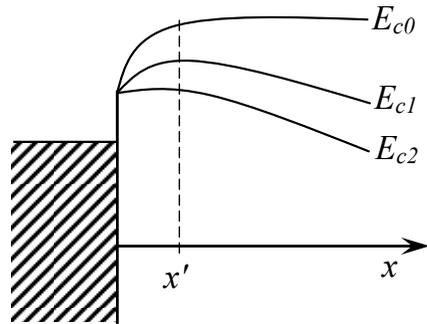


Рис. 2.4.8. Антισταпорный контакт при разных вытягивающих напряжениях  $u_2 > u_1 > u_0 = 0$ .

ка. Но в области объемного заряда поле действует в противоположном направлении. В отсутствии тока через контакт возникающий здесь дрейфовый ток компенсируется противоположно направленным диффузионным, см. разд. 2.4.2.3, стр. 114.

А на некотором расстоянии  $x_1$  от поверхности электрическое поле равно нулю,  $\frac{dE_c}{dx} = 0$ . Следовательно, в этом сечении дрейфовый ток вообще отсутствует, остается только диффузионная составляющая. Получается, что эта плоскость работает как бы виртуальным катодом, резервуаром, поставляющим электроны в полупроводник. Увеличение вытягивающего потенциала приводит к сдвигу плоскости  $x'$  к металлу и одновременно уменьшению высоты барьера, т.е. и повышению концентрации электронов в ней. Это естественно, т.к. увеличение вытягивающего потенциала влечет за собой увеличение градиента энергии Ферми и увеличение плотности тока в объеме полупроводника, а, значит, и во всей рассматриваемой последовательной цепи. И этот возросший ток по-прежнему равен диффузионному току в плоскости  $x_1$ .

При дальнейшем увеличении вытягивания можно придти к такой ситуации, когда барьер  $x'$  вообще исчезнет. На рис. 2.4.8 при потенциале  $u_2$  изображена пороговая точка этих изменений, поле отсутствует на самой поверхности. Затем и на ней будет вытягивающее поле.

Для рассмотрения вольтамперных зависимостей в этой системе предположим, как и ранее, что перезарядку примесей можно не учитывать, и будем отталкиваться от уравнений для плотности тока (2.4.19) и уравнения Пуассона (2.4.21). Исключив из них концентрацию электронов, получим:

$$\frac{\partial^2 \mathcal{E}}{\partial x^2} + \frac{e}{kT} \mathcal{E} \frac{\partial \mathcal{E}}{\partial x} - \frac{4\pi\sigma_0}{\epsilon\mu} \frac{e}{kT} \mathcal{E} + \frac{4\pi j}{\epsilon\mu} \frac{e}{kT} = 0. \quad (2.4.62)$$

Здесь  $\sigma_0 = en_0\mu_n$  – проводимость объема при неизменной концентрации  $n_0$ , а плотность тока  $j$  рассматривается как параметр. Мы рассматриваем стационарные состояния, так что  $j = Const$ .

Это – уравнение второго порядка. Но задача упростится, если первым членом, определяемом диффузией, можно пренебречь. Нас интересует вольтамперная характеристика всей системы металл-переходный слой-полупроводник. У первых двух элементов проводимость существенно выше, чем у объема полупроводника. Основной вклад в сопротивление дадут области вдали от контакта, имеющие концентрацию носителей  $n_0$ , так что возможность пренебрежения диффузией – вполне закономерный вопрос. Он решается соотношением величин первых двух членов. Последние члены не зависят явно от координаты.

Проинтегрировав (2.4.62) один раз, получим:

$$\frac{\partial \mathcal{E}}{\partial x} - \frac{e}{2kT} \mathcal{E}^2 = Const$$

Для оценки величин присутствующих здесь членов можно принять, что  $\frac{\partial \mathcal{E}}{\partial x} \approx \frac{\bar{\mathcal{E}}}{L}$ , и  $\mathcal{E}^2 \approx \bar{\mathcal{E}}^2$ , где  $\bar{\mathcal{E}} = \frac{u}{L}$ . Отсюда получаем, что  $\left| \frac{\partial \mathcal{E}}{\partial x} \right| \ll \left| \frac{e}{kT} \mathcal{E}^2 \right|$ , если  $u \approx \bar{\mathcal{E}}L \gg \frac{kT}{e}$ . Поскольку при комнатной температуре  $kT \approx 0,025$  эВ, это условие не трудно выполнить. И тогда (2.4.62) упрощается:

$$\frac{e}{kT} \mathcal{E} \frac{\partial \mathcal{E}}{\partial x} - \frac{4\pi\sigma_0}{\epsilon\mu} \frac{e}{kT} \mathcal{E} + \frac{4\pi j}{\epsilon\mu} \frac{e}{kT} = 0. \quad (2.4.63)$$

Для дальнейшего упрощения предстоящих вычислений введем безразмерные величины:

$$x_1 = \frac{en_0}{j\tau_M} x, \quad \mathcal{E}_1 = \frac{\sigma_0}{j} \mathcal{E} = \frac{n_0}{n}, \quad \varphi_1 = \frac{en_0\sigma_0}{j^2\tau_M} \varphi, \quad (2.4.64)$$

где 
$$\tau_M = \frac{\varepsilon}{4\pi\sigma_0} \quad (2.4.65)$$

– максвелловское время релаксации, время установления зарядовых равновесий.

В этих обозначениях (2.4.63) существенно упрощается:

$$\mathcal{E}_1 \frac{\partial \mathcal{E}_1}{\partial x} - \mathcal{E}_1 + 1 = 0. \quad (2.4.66)$$

Его решение:

$$\mathcal{E}_1 + \ln[1 - \mathcal{E}_1] = x_1 + C, \quad (2.4.67)$$

На границе, при  $x_1 = 0$ ,  $\mathcal{E}_1(0) = \frac{n_0}{n_k}$ . Отсюда постоянная интегрирования  $C$ :

тегрирования  $C$ :

$$\begin{aligned} C &= \mathcal{E}_1(0) + \ln[1 - \mathcal{E}_1(0)] = \\ &= \mathcal{E}_1(0) + \left[ -\mathcal{E}_1(0) - \frac{1}{2} \mathcal{E}_1^2(0) - \dots \right] \approx -\frac{1}{2} \mathcal{E}_1^2(0) = -\frac{1}{2} \left( \frac{n_0}{n_k} \right)^2. \end{aligned} \quad (2.4.68)$$

Поскольку  $\mathcal{E}_1(0) = \frac{n_0}{n_k} \ll 1$ , здесь вполне можно обойтись

двумя первыми членами разложения логарифма в ряд.

Теперь можно найти и безразмерный потенциал:

$$\varphi_1 = -\int_0^{x_1} \mathcal{E}_1 dx = -\int_{\mathcal{E}_1(0)}^{\mathcal{E}_1} \mathcal{E}_1 \frac{dx}{d\mathcal{E}_1} d\mathcal{E}_1 = \int_{\mathcal{E}_1(0)}^{\mathcal{E}_1} \frac{\mathcal{E}_1^2 d\mathcal{E}_1}{1 - \mathcal{E}_1}. \quad (2.4.69)$$

Полагая  $\mathcal{E}_1(0) \approx 0$ , при интегрировании получим:

$$\varphi_1 = \frac{1}{2} \mathcal{E}_1^2 + \mathcal{E}_1 + \ln(1 - \mathcal{E}_1) = \frac{1}{2} \mathcal{E}_1^2 + x_1 + C. \quad (2.4.70)$$

Здесь в последнем равенстве использовано решение (2.4.67).

Это и есть вольтамперная зависимость, поскольку в присутствующие здесь безразмерные величины плотность тока  $j$  входит как параметр. Уравнение (2.4.70) вместе с определениями (2.4.64) и равенством (2.4.67) описывают, следовательно, зависимость падения напряжения от протекающего тока. Но решение его сложно и в общем случае требует использования численных методов.

Можно, однако, его несколько упростить для особых случаев. Мы рассмотрим три, слабые и сильные токи, причем получаются различные решения для "не очень сильных" и "очень сильных" токов.

### 1. Слабые токи.

Независимо от токов  $\mathcal{E}_1 = \frac{n_0}{n} \leq 1$ ,  $|C| = \frac{1}{2} \left( \frac{n_0}{n_k} \right)^2 \ll 1$ . Если

же токи настолько малы, что  $|x_1| = \left| \frac{en_0}{j\tau_M} x \right| \gg 1$ , то (2.4.67) имеет

решения только при  $x_1 < 0$  и при этом  $\mathcal{E}_1 \approx 1$ . Следовательно, ток  $j < 0$ , т.е. он течет от полупроводника в металл. Это естественно, т.к. мы рассматриваем инжекцию электронов в полупроводник.

Тогда для реальной напряженности поля получаем:

$$\mathcal{E} = \frac{j}{\sigma_0}, - \quad (2.4.71)$$

– закон Ома, причем проводимость определяется равновесной концентрацией носителей в полупроводнике. Контакт никак не влияет на условия протекания тока.

Получили вольтамперную зависимость для образца толщиной  $L$ :

$$u = \int_0^L \mathcal{E} \cdot dx = L \frac{j}{\sigma_0}. \quad (2.4.72)$$

## 2. Сильные токи.

При сильных токах можно считать, что  $\mathcal{E}_1 = \frac{\sigma_0}{j} \mathcal{E} \ll 1$ . Тогда в (2.4.67) можно разложить логарифм в ряд и ограничиться первыми двумя членами. Тогда слева остается всего один член и уравнение перестает быть трансцендентным:

$$\mathcal{E}_1 + \ln[1 - \mathcal{E}_1] = \mathcal{E}_1 + \left[ -\mathcal{E}_1 - \frac{1}{2} \mathcal{E}_1^2 - \dots \right] \approx -\frac{1}{2} \mathcal{E}_1^2 = x_1 + C. \quad (2.4.73)$$

Его решение:

$$\mathcal{E}_1 = \sqrt{2} \sqrt{|x_1| + |C|}. \quad (2.4.74)$$

Здесь возможны два случая.

**Не очень сильные токи**,  $|x_1| \gg |C|$ ,  $\mathcal{E}_1 \approx \sqrt{2|x_1|}$ .

В этом случае:

$$\mathcal{E} = \sqrt{\frac{2j}{\tau_M \cdot \sigma_0 \cdot \mu}} x, \quad (2.4.75)$$

$$u = \int_0^L \mathcal{E} \cdot dx = \sqrt{\frac{8jL^3}{9\tau_M \cdot \sigma_0 \cdot \mu}}, \quad (2.4.76)$$

$$j = \frac{9}{8} \tau_M \cdot \sigma_0 \cdot \mu \frac{u^2}{L^3}. \quad (2.4.77)$$

Ток квадратично зависит от прикладываемого напряжения (закон Мотта)! Но более того, отличие от закона Ома здесь не только в том, что ток определяется квадратом, а не первой степенью среднего поля  $\bar{E} = \frac{u}{L}$  в образце, но также зависит и от толщины образца:

$$j = \frac{9}{8} \tau_M \cdot \sigma_0 \cdot \mu \cdot \bar{E}^2 \frac{1}{L}. \quad (2.4.78)$$

**Очень сильные токи**,  $|x_1| \ll |C|$ ,  $E_1 \approx \sqrt{2|C|}$ .

Тогда:

$$E = \frac{j}{\sigma_0} \frac{n_0}{n_k}, \quad j = e\mu n_k \frac{u}{L} \quad (2.4.79)$$

$$u = j \frac{L}{e\mu} \frac{1}{n_k} \quad (2.4.80)$$

Здесь опять выполняется закон Ома, но концентрация носителей уже не  $n_0$ , а  $n_k$ , т.е. проводимость существенно увеличена!

Итого, мы получили три типа зависимостей:

$$\begin{aligned} 1) \quad u &= j \frac{L}{\sigma_0} = j \frac{L}{e\mu} \frac{1}{n_0}, & j &= u \frac{e\mu n_0}{L} = u \frac{\sigma_0}{L}, \\ 2) \quad u &= \sqrt{j} \sqrt{\frac{8L^3}{9\tau_M \cdot \sigma_0 \cdot \mu}}, & j &= u^2 \frac{9}{8} \frac{\tau_M \cdot \sigma_0 \cdot \mu}{L^3} = u^2 \frac{9}{8} \frac{\mu\varepsilon}{4\pi} \cdot \frac{1}{L^3}, \\ 3) \quad u &= j \frac{L}{e\mu} \frac{1}{n_k}, & j &= u \frac{e\mu n_k}{L}. \end{aligned}$$

На рис. 2.4.9 этим зависимостям соответствуют отрезки  $a$ - $b$ ,  $b$ - $c$ ,  $c$ - $d$ , обозначенные жирными линиями.

Для того, чтобы выяснить причины изменения вольтамперной зависимости, проверим, при каких условиях это происходит.

Приравняем величины токов, которые дает пара зависимостей в точках перехода  $b$  и  $c$ .

Переход к квадратичному закону:

$$u^2 \frac{9}{8} \frac{\mu \varepsilon}{4\pi} \cdot \frac{1}{L^3} = u \frac{e\mu n_0}{L} = u \frac{\sigma_0}{L} \Rightarrow$$

$$\Rightarrow \tau_M = \frac{\varepsilon}{4\pi\sigma_0} = \frac{8}{9} \frac{L^2}{\mu u} \approx \frac{L}{\mu E} = \frac{L}{\bar{v}} = t_{dr}. \quad (2.4.81)$$

Оказывается, переход к квадратичному закону происходит тогда, когда время дрейфа электрона через образец становится равно времени релаксации.

Возврат к закону Ома:

$$u^2 \frac{9}{8} \frac{\tau_M \cdot \sigma_0 \cdot \mu}{L^3} = u \frac{e\mu n_k}{L}.$$

Здесь – то же соотношение, но время дрейфа становится равно времени релаксации при концентрации носителей, увеличенной до  $n_k$ :

$$t_{dr} = \frac{\varepsilon}{4\pi e\mu n_k} = \tau'_M. \quad (2.4.82)$$

Следовательно, на квадратичном участке происходит эффективная инжекция, увеличивающая среднюю концентрацию носителей, поскольку электрон начинает пролетать через образец быстрее, чем может установиться равновесная концентрация. Идет "пропитка" полупроводника избыточными электронами. Но это тем труднее сделать, чем больше толщина образца. Отсюда – третья степень толщины образца в (2.4.77).

В проведенном рассмотрении мы не учитывали возможность перезарядки локальных центров. А этот процесс должен сказаться именно на квадратичном участке. Здесь растет средняя концен-

трация носителей, следовательно, должна изменяться и заполненность примесных центров. Это приводит к расширению омического интервала  $a$ - $b$ . Реальные зависимости содержат участок  $g$  (на рис. 2.4.9). Здесь и происходит перезарядка примесных центров. Но заполнение может стать полным и закончиться при потенциалах, меньших точки  $c$ , а когда оно заканчивается, ситуация возвращается к рассмотренной и ВАХ очень круто достигает прямой  $bc$  (отрезок  $e$ - $f$ ).

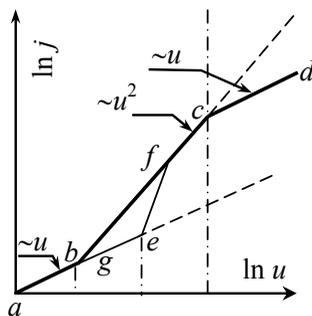


Рис. 2.4.9. Ожидаемая вольтамперная характеристика ТОПЗ.

Подробнее с особенностями ТОПЗ при наличии примесей и с временными характеристиками процесса установления равновесия можно ознакомиться, например, по [14], глава 4.

### 2.4.3.2. Ток через запорный контакт

На рис. 2.4.10 изображен запорный слой при различных смещениях. Во всех случаях можно считать, что падение напряжения, приложенного извне, происходит в области объемного заряда, поскольку там меньше концентрация носителей и проводимость. Поэтому в объеме нет падения напряжения и на рисунке уровень Ферми и энергии краев зон в объеме не зависят от координаты.

В отсутствие смещения токи на поверхности идут в обе стороны, но они равны. При небольшом смещении в прямом направлении ("-" приложен к  $n$ -полупроводнику) высота барьера для потока электронов из металла в полупроводник (ток  $j_2$ ) не изменяется и сам этот ток не изменяется. Обратный же перенос (ток  $j_1$ ) существенно возрастает, поскольку здесь уменьшается высота

барьера относительно уровня Ферми в непосредственной близости от контакта. Здесь следует ожидать экспоненциальной зависимости прямого тока от напряжения.

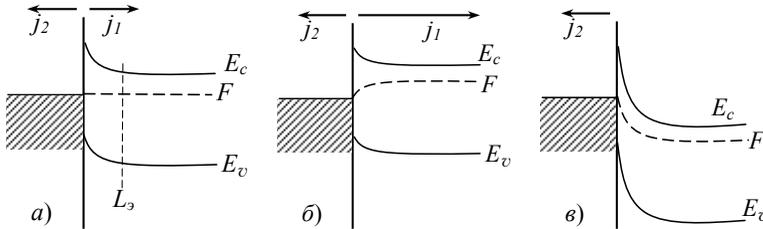


Рис. 2.4.10. Запорный слой в  $n$ -полупроводнике: а) – без смещения; б) – смещенный в прямом направлении; в) – смещенный в обратном (запорном направлении).

При отрицательном смещении ток  $j_2$  опять-таки остается неизменным, но ток  $j_1$  может быть уменьшен до нуля. Правда, при очень больших отрицательных смещениях ширина барьера для перехода электронов из металла может быть уменьшена настолько, что станет возможным туннелирование. Возникнет пробой запирающего промежутка. Но до этого обратный ток в широком диапазоне напряжений будет постоянен, равен  $j_2$ .

Общие соображения относительно особенностей переноса на таком контакте касаются механизма преодоления барьера. Во-первых, здесь можно говорить о токах термоэмиссии, – прохождении электронов над барьером. Во-вторых, возможен и туннельный механизм переноса.

Туннелирование можно считать вероятным в случае, если длина экранирования  $L_s$  не превышает длины волны электрона. Это очень грубая оценка, так как вероятность туннелирования зависит и от высоты барьера, но порядок толщины туннельного барьера она, все таки, позволяет определить.

Длина экранирования – порядка 100 нм при концентрации электронов  $10^{15} \text{ см}^{-3}$ , а длины волн – порядка 10 нм. Поэтому чаще можно пренебрегать туннелированием. Для него нужно существенно уменьшать ширину барьера, т.е. повышать концентрацию носителей. Эффект туннелирования используется, например, в туннельных диодах (см. раздел 2.4.3.4 и [15], глава VIII). Но в них легированием создаются вообще вырожденные полупроводники, с предельно малой глубиной экранирования.

Так что наш случай – прохождение над барьером. Этот процесс тоже может происходить по-разному. Если ширина барьера все-таки не очень велика, или полупроводник достаточно совершенен, или температуры низки, то длина свободного пробега электрона может оказаться больше  $L_3$ . Тогда барьер преодолевается на прямом пролете и в этом случае вполне приложимы представления о термоэмиссии. Напряженность поля в области загиба зон оказывается не существенна. Во всех реальных случаях энергия, приобретаемая от поля на длине пробега мала.

Если же длина свободного пробега мала для этого, то возможна только диффузия. Мы рассмотрим прямой пролет над барьером (диодная теория выпрямления Бете). Рассмотрение диффузионного механизма (диффузионной теории выпрямления) сложнее, но результат остается практически тем же. С ним можно ознакомиться по [15], §VI.12. Мы же только приведем его результаты.

### *Диодная теория.*

Механизм термоэлектронной эмиссии мы рассмотрели в [1], раздел 1.6.1.3, и получили выражение:

$$j = \frac{4\pi m_0 e k^2 T^2}{(2\pi\hbar)^3} \exp\left(\frac{F}{kT}\right) = AT^2 \cdot \exp\left(\frac{F}{kT}\right),$$

где  $A = \frac{4\pi m_0 e k^2}{(2\pi\hbar)^3}$ , а  $F$  – энергия Ферми, отсчитанная от нулевого

уровня вакуума, т.е. высота преодолеваемого барьера. В нашем случае высота барьера равна

$$\Phi = e(u + u_k) + (E_c - F), \quad (2.4.83)$$

так что:

$$j_1 = \frac{4\pi m_0 e k^2}{(2\pi\hbar)^3} T^2 \exp\left(-\frac{e(u + u_k) + (E_c - F)}{kT}\right). \quad (2.4.84)$$

Сюда можно подставить среднюю тепловую скорость электронов  $v_T = \sqrt{\frac{8kT}{\pi m}}$  и концентрацию электронов в объеме, которая (в случае, далеком от вырождения, см. (2.2.10), (2.2.13)), равна:

$$n_0 = 2 \left( \frac{2\pi m_n kT}{(2\pi\hbar)^2} \right)^{3/2} \exp\left(-\frac{E_c - F}{kT}\right). \quad (2.4.85)$$

Получим:

$$j_1 = \frac{1}{4} e n_0 v_T \exp\left(-\frac{e(u + u_k)}{kT}\right). \quad (2.4.86)$$

Это – поток электронов из полупроводника в металл. Обратный равен ему при  $u = 0$ , так что

$$j_2 = \frac{1}{4} e n_0 v_T \exp\left(-\frac{e u_k}{kT}\right). \quad (2.4.87)$$

Разность этих токов – результирующий ток через контакт:

$$j = j_1 - j_2 = \frac{1}{4} en_0 v_T \exp\left(-\frac{eu_k}{kT}\right) \left[ \exp\left(-\frac{eu}{kT}\right) - 1 \right] = \quad (2.4.88)$$

$$= j_s \left[ \exp(-\alpha u) - 1 \right],$$

где 
$$j_s = \frac{1}{4} en_0 v_T \exp\left(-\frac{eu_k}{kT}\right), \quad \alpha = \frac{e}{kT}. \quad (2.4.89)$$

График этой зависимости представлен на рис. 2.4.11. При положительных напряжениях  $u$  (запорное направление) экспонента в квадратных скобках (2.4.88) быстро уменьшается и уже при  $eu = (3..4)kT$  становится пренебрежимо мала. Обратный ток достигает насыщения,  $j = -j_s$ .

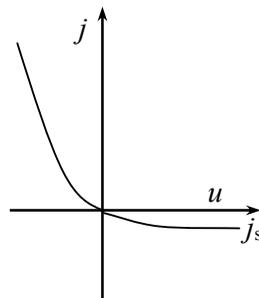


Рис. 2.4.11. Вольт-амперная характеристика запорного контакта.

Напротив, при отрицательных  $u$  уже при  $eu = (3..4)kT$  можно пренебречь единицей в скобках (2.4.88) и зависимость тока от напряжения становится чисто экспоненциальной:

$$j = j_s \exp(-\alpha u). \quad (2.4.90)$$

Эта характеристика запорного контакта широко используется в электронной технике в потенцирующих, а чаще — в логарифмирующих аналоговых устройствах.

**Диффузионная теория** дает практически тот же результат:

$$j = j_s \left[ \exp(-\alpha u) - 1 \right], \text{ но } j_s = en_0 \mu |E(0)| \exp\left(-\frac{eu_k}{kT}\right). \quad (2.4.91)$$

Здесь различия (кроме коэффициента  $1/4$ ) — в том, что при диффузии направленный электронный поток определяется не

средней скоростью хаотического движения свободных электронов, а скоростью их дрейфа, равной произведению подвижности на среднее поле в области объемного заряда.

### 2.4.3.3. Эффект Шоттки

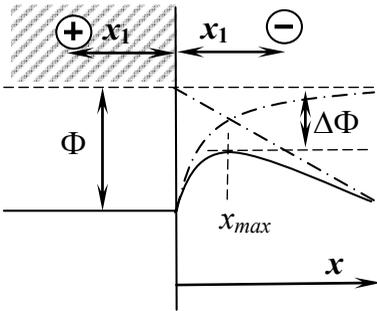


Рис. 2.4.12. Эффект Шоттки.

Реальные ВАХ в за- порной области не очень хорошо описываются полу- ченной зависимостью. Основ- ная причина этого – уменьшение высоты потен- циального барьера в боль- ших вытягивающих полях. На рис. 2.4.10 это не отра- жено, а на рис. 2.4.12 увели- чен масштаб по оси  $x$ .

И в отсутствие вытяги- вающего поля барьер на поверхности – не прямоугольный. Работа выхода электрона – работа по преодолению сил зеркального ото- бражения и зависимость энергии связи электрона от координаты отображена на рисунке штрихпунктирной линией, при наличии поля – сплошной. Она описывается законом:

$$U = -\frac{e^2}{4\epsilon x} - eEx. \quad (2.4.92)$$

Вершина барьера находится в точке, в которой поле притя- жения к поверхности равно вытягивающему полю:

$$U'_x = \frac{e^2}{4\epsilon x^2} - eE = 0 \Rightarrow x_{\max} = \sqrt{\frac{e}{4\epsilon E}}. \quad (2.4.93)$$

Уменьшение высоты барьера – модуль энергии (2.4.92) в точке  $x_{\max}$ . Она равна:

$$\Delta\Phi = \sqrt{\frac{e^3\mathcal{E}}{\varepsilon}}. \quad (2.4.94)$$

Величина  $x_{\max}$  много меньше ширины барьера. Поэтому в (2.4.94) лучше брать не среднее поле в запиорном слое, а  $\mathcal{E}(0)$ , поле на границе раздела. Согласно (2.3.55), оно равно

$$\mathcal{E} = -\sqrt{\frac{8\pi en_0(u_k + u)}{\varepsilon\pi}}. \quad (2.4.95)$$

Диоды типа металл/полупроводник называются диодами Шоттки.

#### 2.4.3.4. Туннелирование, туннельный диод

Лео Есаки<sup>1</sup> создал (1958 г) диод с необычными свойствами. У него наблюдался участок с отрицательным сопротивлением и способностью усиливать сигнал. При изучении внутренней полевой эмиссии в вырожденном германиевом  $p$ - $n$  переходе он обнаружил эту "аномальную" ВАХ, – дифференциальное сопротивление на одном из участков характеристики было отрицательным. Он сумел объяснить этот эффект, который заключается в квантово-механическом туннелировании через изоляционный промежуток.

Суть туннелирования заключается в том, что в области пространства, в которой полная энергия  $E$  меньше потенциальной  $eU$ , волновой вектор

$$k = \sqrt{2m^*(E - eU)}/\hbar = i\sqrt{2m^*(eU - E)}/\hbar \equiv ik_2$$

– мнимая величина, так что волновая функция электрона,

$$\Psi_k(x) = u_k(x) \cdot e^{ikx} = u_k(x) \cdot e^{-k_2x},$$

---

<sup>1</sup> Иногда эту фамилию пишут как Эсаки.

экспоненциально спадает в пространстве. Электрон не может проникнуть "под барьер" на большую глубину и непременно отразится от толстого барьера. Но на какую-то глубину, с не очень большой вероятностью, он все же проникает. В частности, на поверхности, на границе кристалл/вакуум, это приводит к возникновению уровней Тамма (см. разд. 2.1.3.2).

Чем больше высота барьера, тем больше  $k_2$  и тем резче обнуляется волновая функция на границе барьера. Но в реальном случае, если барьер имеет малую толщину, имеется отличная от нуля вероятность прохождения сквозь барьер. В случае, если барьер не очень крут, так что

$$\frac{d}{dx} \left\{ \frac{\hbar}{\sqrt{2m(eU(x) - E)}} \right\} \ll 1, \quad (2.4.96)$$

расчет прозрачности может быть проведен на основе классических представлений и тогда он дает:

$$D = D_0 \exp \left[ \frac{-2}{\hbar} \int_{x_1}^{x_2} dx \sqrt{2m(eU(x) - E)} \right], \quad D_0 \approx 1. \quad (2.4.97)$$

Условие (2.4.96) и назовем условием квазиклассичности барьера.

В работе Есаки было два существенных технических момента. Ему удалось:

- ввести большие концентрации примесей в Si, донорных и акцепторных, так что получились вырожденные полупроводники;
- сделать достаточно резкий переход между вырожденными  $n$ - и  $p$ -областями.

Энергетическая диаграмма получившегося перехода представлена на рис. 2.4.13. Разрыв между  $n$ - и  $p$ -областями, в кото-

ром нет носителей, достаточно тонок для того, чтобы туннелирование было возможно.

В отсутствие внешнего напряжения уровни Ферми выровнены, обмена носителями нет. Если к переходу приложить обратное смещение, "-" к  $p$ -области, то электроны из заполненной части  $v$ -зоны  $p$ -проводника беспрепятственно пройдут на свободные состояния  $s$ -зоны  $n$ -проводника и этот "отрицательный" ток будет только возрастать с напряжением.

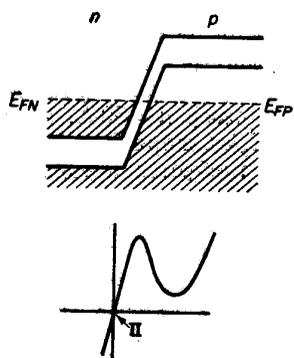


Рис. 2.4.13. Энергетическая диаграмма и ВАХ туннельного диода. [16]

Процессы, происходящие при положительном смещении, иллюстрирует рис. 2.4.14.

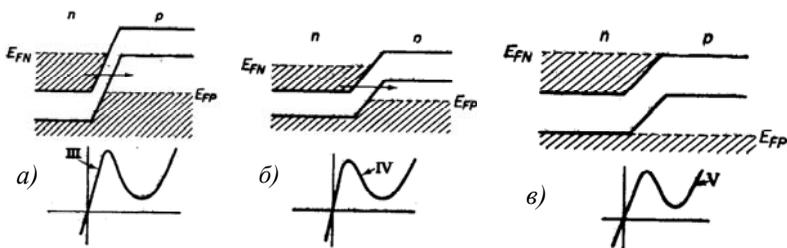


Рис. 2.4.14. ВАХ туннельного диода при положительном смещении. [16]

При уменьшении энергии электронов в  $p$ -области свободная часть ее  $v$ -зоны начинает перекрываться с заполненной частью  $s$ -зоны  $n$ -области. По мере того, как с повышением напряжения это перекрывание увеличивается, растет и ток, область III на ВАХ. При дальнейшем увеличении напряжения это перекрывание уменьшается, так что возникает область с отрицательным диффе-

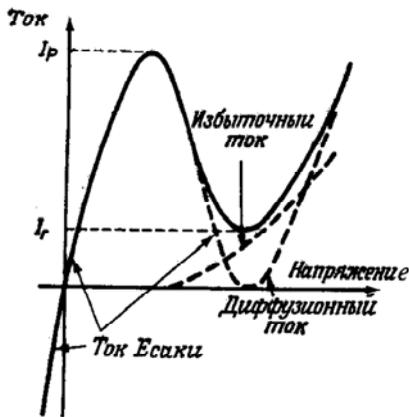


Рис. 2.4.15 Составляющие ВАХ туннельного диода.

ки, связанные с примесями в запирающем слое. Эту составляющую называют **избыточным током**. На рис. 2.4.15 изображена полная ВАХ туннельного диода и отмечены составляющие разной природы.

ренциальным сопротивлением, область IV. Но затем смещение становится уже столь велико, что электроны смогут свободно диффундировать в зону  $p$ -проводника. Тогда ток начинает неограниченно возрастать, область V на рис. 2.4.14.

Реально ток через диод не уменьшается до нуля из-за неизбежных несовершенств структуры. В том числе возникают токи

## 2.5. НЕРАВНОВЕСНЫЕ НОСИТЕЛИ

В разделе 2.2 мы рассматривали статистику заполнения электронных состояний в условиях *термодинамического равновесия*. Вообще любое равновесие в многочастичной квантовой системе – равенство скоростей (темпов) переходов частиц между состояниями  $\Omega_1$  и  $\Omega_2$  в обе стороны. Если мы рассматриваем именно термодинамическое равновесие, то все возбуждения в системе происходят за счет тепловой энергии, в том числе – теплового излучения. При этом возможно рассчитать вероятность (темп) прямого возбуждения электронов из состояния  $\Omega_1$  в состояние  $\Omega_2$ .

Релаксация возбуждения, в принципе, может идти любыми путями, через любые промежуточные состояния. Но в дальнейшем мы увидим, что в термодинамически равновесных системах справедлив так называемый *принцип детального равновесия*, согласно которому темпы переходов  $\Omega_1 \Rightarrow \Omega_2$  и  $\Omega_2 \Rightarrow \Omega_1$  равны. Например, равновесная концентрация носителей, рассмотренная в разд. 2.2, поддерживается непрерывно следующими актами возбуждения и рекомбинации электронно-дырочных пар. При этом можно утверждать, что равны частоты переходов вверх и вниз не только между зонами, но и конкретно между фиксированными группами состояний,  $v_{12} = v_{21}$  на рис. 2.5.1.

Между тем, в хорошем приближении можно считать, что вероятность электрону рекомбинировать не зависит от вероятности возбуждения, от температуры. Возбуждение и рекомбинация – разные процессы. Если температура возрастет, то увеличится темп генерации. Для восстановления равновесия должен увеличиться и темп рекомбинации. А при неизменном ее механизме и неизменной вероятности рекомбинации это возможно только за счет увеличения концентрации электронов.

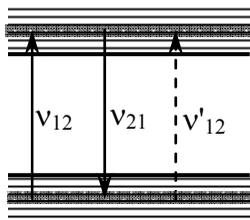


Рис. 2.5.1 *Возбуждение и рекомбинация носителей в полупроводниках.*

Но возбуждение может быть вызвано и внешними воздействиями. Например, инжекцией носителей (см. разд. 2.4.3.1) или облучением,  $v'_{12}$  на рис. 2.5.1. При постоянном внешнем воздействии тоже образуется равновесная система, но в ней не справедлив принцип детального равновесия и для определения эффекта, производимого этим внешним возбуждением, нужно учесть и рассчитать все возможные каналы релаксации.

В любом случае это возбуждение приведет к увеличению концентрации носителей и изменению их распределения по состояниям. После снятия возбуждения они довольно быстро вернуться к исходным, термически равновесным. Эту дополнительную концентрацию, созданную внешним возбуждением, мы и будем считать *неравновесной*. Наша задача – рассмотреть процессы генерации, релаксации и рекомбинации носителей при внешнем возбуждении.

### 2.5.1. *Время жизни носителей заряда*

Введем обозначения:

$\delta n, \delta p$  - избыточные (неравновесные) концентрации носителей;

$(\delta n)_s, (\delta p)_s$  - стационарные концентрации неравновесных носителей, установившиеся после выдержки полупроводника в неизменных условиях;

$G_n, G_p$  - полные темпы генерации свободных электронов и дырок;

$r_n, r_p$  - полные темпы их рекомбинации, исчезновения в результате захвата либо носителями противоположного знака, либо локальными центрами;

$g_{nT}, g_{pT}$  - темпы термической генерации свободных электронов и дырок;

$R_n, R_p$  - темпы рекомбинации неравновесных носителей;

$\tau_n, \tau_p$  - времена жизни неравновесных носителей.

Скорость изменения концентраций неравновесных носителей:

$$\frac{d\delta n}{dt} = g_n - R_n, \quad \frac{d\delta p}{dt} = g_p - R_p. \quad (2.5.1)$$

Поскольку в термодинамическом равновесии темпы генерации и рекомбинации равны, то

$$R_n = r_n - g_{nT}, \quad R_p = r_p - g_{pT}. \quad (2.5.2)$$

Очевидно,

$$R_n = \frac{\delta n}{\tau_n} = \frac{n - n_0}{\tau_n}, \quad R_p = \frac{\delta p}{\tau_p} = \frac{p - p_0}{\tau_p}. \quad (2.5.3)$$

Следовательно,  $\frac{1}{\tau_n}, \frac{1}{\tau_p}$  - вероятности исчезновения из зоны одного носителя за единицу времени. Если эти величины не сильно зависят от концентраций или  $\delta n, \delta p$  в (2.5.3) не велики, то термические темпы генерации:

$$g_{nT} = \frac{n_0}{\tau_n}, \quad g_{pT} = \frac{p_0}{\tau_p} \quad (2.5.4)$$

и тогда (2.5.1) примет вид:

$$\frac{d\delta n}{dt} = g_n - \frac{\delta n}{\tau_n}, \quad \frac{d\delta p}{dt} = g_p - \frac{\delta p}{\tau_p}. \quad (2.5.5)$$

Стационарные концентрации неравновесных носителей заряда, устанавливающиеся после длительного воздействия внешней генерации, получим отсюда при условии:

$$\frac{d\delta n}{dt} = 0, \quad \frac{d\delta p}{dt} = 0$$

Они будут равны:

$$(\delta n)_s = g_n \tau_n, \quad (\delta p)_s = g_p \tau_p. \quad (2.5.6)$$

В общем случае  $\tau_n, \tau_p$  зависят не только от характеристик полупроводника и неравновесных концентраций  $\delta n, \delta p$ , но и от условий эксперимента. Если все же считать  $\tau_n, \tau_p$  заданными величинами, определяющими кинетику электронных процессов, не зависящими от  $\delta n, \delta p$ , то интегрирование уравнений (2.5.5) дает простые экспоненциальные зависимости:

$$\delta n = g_n \tau_n - C \exp(-t/\tau_n), \quad \delta p = g_p \tau_p - C \exp(-t/\tau_p).$$

Постоянные интегрирования  $C$  могут быть определены из начальных условий. При  $t = 0$   $\delta n, \delta p = 0$ , так что

$$C = g_n \tau_n = (\delta n)_s, \quad C = g_p \tau_p = (\delta p)_s,$$

и окончательно: при включении генерации:

$$\delta n = g_n \tau_n [1 - \exp(-t/\tau_n)], \quad \delta p = g_p \tau_p [1 - \exp(-t/\tau_p)]; \quad (2.5.7a)$$

При выключении, если неравновесные концентрации к моменту выключения составляли  $\delta n_1, \delta p_1$ :

$$\delta n = \delta n_1 \exp(-t/\tau_n), \quad \delta p = \delta p_1 \exp(-t/\tau_p). \quad (2.5.7б)$$

### 2.5.2. Уравнение непрерывности

Если в полупроводнике имеются электрические токи, то их тоже нужно учитывать как один из каналов изменения концентрации носителей. Вклад токов можно найти, рассматривая элементарный объем, например, в виде прямоугольного параллелепипеда с ребрами, параллельными прямоугольным осям координат, и вычисляя втекающие и вытекающие из него потоки частиц. Это дает

$$\begin{aligned} \left(\frac{\partial p}{\partial t}\right)_{dr} &= -\frac{1}{e} \left( \frac{\partial j_{px}}{\partial x} + \frac{\partial j_{py}}{\partial y} + \frac{\partial j_{pz}}{\partial z} \right) = -\frac{1}{e} \operatorname{div} j_p, \\ \left(\frac{\partial n}{\partial t}\right)_{dr} &= \frac{1}{e} \left( \frac{\partial j_{nx}}{\partial x} + \frac{\partial j_{ny}}{\partial y} + \frac{\partial j_{nz}}{\partial z} \right) = \frac{1}{e} \operatorname{div} j_n. \end{aligned} \quad (2.5.8)$$

Здесь

$$\begin{aligned} j_p &= \sigma_p \mathbf{E} - e D_p \nabla p, & \sigma_p &= e p \mu_p, \\ j_n &= \sigma_n \mathbf{E} + e D_n \nabla n, & \sigma_n &= e n \mu_n. \end{aligned} \quad (2.5.9)$$

Итого:

$$\frac{\partial p}{\partial t} = g_p - \frac{1}{e} \operatorname{div} j_p - \frac{\delta p}{\tau_p}, \quad \frac{\partial n}{\partial t} = g_n + \frac{1}{e} \operatorname{div} j_n - \frac{\delta n}{\tau_n}. \quad (2.5.10)$$

При изменении концентраций носителей заполнение локальных центров тоже может измениться, так что полный объемный заряд:

$$\rho = e(p + p_t - n - n_t).$$

Уравнение Пуассона в этом случае примет вид:

$$\operatorname{div} \mathbf{E} = \frac{4\pi}{\varepsilon} \rho = \frac{4\pi}{\varepsilon} e(p + p_t - n - n_t). \quad (2.5.11)$$

Отсюда, при известных граничных условиях, можно определить напряженность поля  $\mathcal{E}$ .

Для расчета концентраций уравнений (2.5.9) - (2.5.11) достаточно, но в ряде случаев можно обойтись и без уравнения Пуассона. Это возможно, если мы можем считать, что  $\rho = 0$ . Но мы рассматриваем и генерацию, и перемещение носителей в пространстве. Очевидно, что для такого допущения мы должны принять, что все интересующие нас процессы достаточно медленны по сравнению с максвелловским временем релаксации. Оно должно быть существенно меньше и времени жизни генерированного носителя, и времени его пролета через образец,

$$\tau_M \ll \tau_n, \tau_p, \quad \tau_M \ll t_{dr}. \quad (2.5.12)$$

Но в этом случае упрощаются и уравнения (2.5.10). Поскольку темп генерации носителей равен темпу увеличения их концентрации минус темп их захвата, а захват возможен либо на рекомбинацию, либо на перезарядку локальных центров, то очевидно равенство:

$$\left( g_p - \frac{\delta p}{\tau_p} \right) - \left( g_n - \frac{\delta n}{\tau_n} \right) = \frac{\partial}{\partial t} (n_t - p_t).$$

Тогда, вычитая друг из друга уравнения (2.5.10) и учитывая это равенство, получаем:

$$\frac{\partial \rho}{\partial t} = -\frac{1}{e} \operatorname{div} j, \quad (2.5.13)$$

т.е. уравнение непрерывности в его классическом виде. И при  $\rho = 0$ ,

$$\operatorname{div} j = 0. \quad (2.5.14)$$

С другой стороны, если выразить  $\rho$  из уравнения Пуассона и подставить его в (2.5.13), получим

$$\operatorname{div}\left(j + \frac{\varepsilon}{4\pi} \frac{\partial \mathcal{E}}{\partial t}\right) = 0. \quad (2.5.15)$$

где  $\frac{\varepsilon}{4\pi} \frac{\partial \mathcal{E}}{\partial t}$  - ток смещения. Равенства (2.5.14) и (2.5.15) могут выполняться одновременно только в случае, если токи смещения пренебрежимо малы.

Интересно отметить и такой факт, что если концентрация локальных центров мала по сравнению с  $\delta n$ ,  $\delta p$ , то их перезарядка не существенна и

$$\rho = \delta p - \delta n = 0, \quad \delta p = \delta n. \quad (2.5.16)$$

Если к тому же генерация происходит за счет перебросов зона-зона, т.е.  $g_n = g_p$ , то равны и времена жизни электронов и дырок:

$$g_n = \frac{d\delta n}{dt} + \frac{\delta n}{\tau_n} = \frac{d\delta p}{dt} + \frac{\delta p}{\tau_p} \Rightarrow \tau_n = \tau_p = \tau. \quad (2.5.17)$$

### 2.5.3. Фотопроводимость

Фотопроводимость – увеличение проводимости полупроводника при облучении. Она возникает как следствие увеличения концентрации носителей при оптической генерации.

На рис. 2.5.2 изображены возможные переходы в полупроводнике с одним типом локальных центров. В первую очередь это – межзонная генерация  $g$ . Если энергия падающих фотонов достаточна и именно она ответственна за увеличение концентрации, то такая фотопроводимость и такой фотопроводник называются *собственными*. Возможна и прямая рекомбинация,  $R$ .

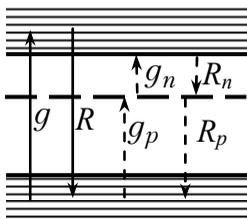


Рис. 2.5.2. Схема переходов при фотопроводимости в полупроводнике с одним типом локальных центров.

Правда, здесь возникают сложности с сохранением энергии. Сталкиваются две свободные частицы, электрон и дырка. В результате обе они исчезают. Можно представить себе, что они имели одинаковые по величине разнонаправленные импульсы, так что закон сохранения импульса не будет нарушен. Но энергию надо куда-то выделить. В принципе, возможно только два канала – генерация облака фононов или фотона, оптическое излучение. Оба они мало вероятны. Генерация фононов – потому, что в одном акте должно быть

генерировано множество частиц. Излучение, как мы увидим позже, сильно зависит от энергии, от ширины зоны, и в широкозонных полупроводниках практически исключено.

Как правило, рекомбинация происходит через локальные центры, в переходах, обозначенных  $R_n$ ,  $R_p$ . Она происходит в две ступени – захват центром электрона и захват дырки.

В принципе, генерация возможна и при оптических переходах, обозначенных  $g_n$ ,  $g_p$ . Но только в том случае, если концентрация центров достаточно велика. Обычно примесное поглощение на несколько порядков слабее собственного.

Темп генерации при собственном поглощении пропорционален плотности поглощенной энергии. Ослабление потока фотонов в пространстве дает закон Бугера-Ламберта:

$$I(x) = I_0 \exp[-\gamma(\omega)x] \quad (2.5.18)$$

Здесь  $\gamma(\omega)$  - коэффициент поглощения.

Плотность поглощаемой энергии:

$$\frac{dI}{dx} = -\gamma(\omega)I_0 \exp[-\gamma(\omega)x] = -\gamma(\omega)I(x). \quad (2.5.19)$$

Но нужно еще учитывать, что, во-первых, темп генерации определяется все-таки не поглощенной энергией, а количеством поглощенных квантов. Во-вторых, не каждый поглощенный квант приведет к генерации электронно-дырочной пары. Поэтому нужно ввести дополнительный коэффициент  $\nu(\omega)$ , зависящий от  $\omega$  и называемый *коэффициент фотоэлектрического поглощения*. Если интенсивность света измеряется числом падающих квантов, то  $\nu(\omega) \leq 1$ <sup>1</sup>.

Итого, темп генерации электронно-дырочных пар:

$$g = \nu(\omega)\gamma(\omega)I(x). \quad (2.5.20)$$

Проводимость определяется не только концентрацией носителей, но и их подвижностью, которая зависит от энергии. До сих пор мы под  $\mu$  понимали некоторую среднюю величину для термически равновесных носителей. Но при оптическом поглощении  $\hbar\omega$  обычно превышает ширину зоны, так что носители приобретают дополнительную энергию, больше равновесной. И нередки ситуации, когда неравновесных носителей оказывается существенно больше, чем равновесных. Поэтому вопрос об их подвижности может быть существенен.

Ситуация облегчается тем, что в подавляющем большинстве случаев время жизни носителей в зоне  $\tau_n$ ,  $\tau_p$ , оказывается существенно больше времени  $\tau_E$ , за которое носители раздадут при соударениях с фононами или дефектами весь избыток энергии и

---

<sup>1</sup> Для того, чтобы выполнялось  $\nu(\omega) > 1$ , необходимо, чтобы энергия фотона в несколько раз превышала ширину зоны. Например, в кремнии  $\sim 3\Delta E$ . Такие ситуации возможны, но они выходят за рамки нашего рассмотрения.

придут к термически равновесному распределению (времени термолизации). Обычно здесь разница в три или более порядков.

Мы будем считать, что это действительно так, а потому подвижности неравновесных и равновесных носителей считаем одинаковыми. Тогда величина фотопроводимости может быть определена как:

$$\delta\sigma = e(\mu_p\delta p + \mu_n\delta n), \quad (2.5.21)$$

где  $\mu_n$ ,  $\mu_p$  - подвижности равновесных электронов и дырок.

Умножив первое из уравнений (2.5.9) на  $e\mu_p$ , второе – на  $e\mu_n$  и сложив их почленно, получим:

$$\frac{d\delta\sigma}{dt} = e(\mu_p + \mu_n)g - \frac{\delta\sigma}{\tau_{\text{фп}}}, \quad (2.5.22)$$

где новая введенная величина,

$$\tau_{\text{фп}} = \frac{\mu_p\delta p + \mu_n\delta n}{\frac{\mu_p}{\tau_p}\delta p + \frac{\mu_n}{\tau_n}\delta n}, \quad (2.5.23)$$

– ни что иное, как время затухания, или время релаксации фотопроводимости. В общем случае оно зависит от концентраций даже если  $\tau_n$  и  $\tau_p$  не зависят.

Стационарная величина фотопроводимости  $(\delta\sigma)_s$  может быть получена отсюда при подстановке  $\frac{d\delta\sigma}{dt} = 0$ :

$$(\delta\sigma)_s = e(\mu_p + \mu_n)g \cdot \tau_{\text{фп}}. \quad (2.5.24)$$

Решение уравнения (2.5.22), как и (2.5.5), есть:

$$\begin{aligned}\delta\sigma|_{on} &= (\delta\sigma)_s [1 - \exp(-t/\tau_{\text{фп}})], \\ \delta\sigma|_{off} &= (\delta\sigma)_1 \exp(-t/\tau_{\text{фп}})\end{aligned}\tag{2.5.25}$$

для включения и выключения генерации соответственно.

Это – не чисто экспоненциальная зависимость, так как  $\tau_{\delta i}$  – величина не постоянная. На рис. 2.5.3 показан примерный вид временной зависимости  $\delta\sigma(t)$  при прерывистом освещении.

Здесь принципиально интересны две характеристики – величина стационарной фотопроводимости  $(\delta\sigma)_s$  и наклоны в начальные моменты генерации и релаксации фотопроводимости.

По  $(\delta\sigma)_s$  можно определить стационарное время жизни фотопроводимости. На этом участке справедливо:

$$\delta p = g\tau_p, \quad \delta n = g\tau_n.$$

Подставив это в (2.5.23), получим:

$$\tau_{\text{фп}} = \frac{\mu_p \tau_p + \mu_n \tau_n}{\mu_p + \mu_n}.\tag{2.5.26}$$

Величины подвижности обычно можно определить из независимых измерений и, если у нас фотопроводимость монополярна, например,  $\tau_p \gg \tau_n$ , а подвижности одного порядка, то (2.5.26) превращается в:

$$\tau_{\text{фп}} = \frac{\mu_p}{\mu_p + \mu_n} \tau_p,\tag{2.5.27}$$

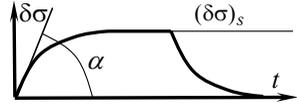


Рис. 2.5.3. Фотопроводимость как функция времени при прерывистом освещении.

отсюда можно определить время жизни долгоживущих носителей.

Наклон зависимости  $\delta\sigma(t)$  в начальный момент времени, при включении освещения, получим из (2.5.22), подставив в него  $\delta\sigma = 0$  :

$$tg\alpha = e(\mu_p + \mu_n)g = e(\mu_p + \mu_n)v\gamma I. \quad (2.5.28)$$

Отсюда можно определить величину  $v(\omega)$  при известных прочих параметрах.

#### 2.5.4. Квазиуровни Ферми

Если распределение по энергиям у фотогенерированных носителей то же, что у равновесных, т.е. определяется только температурой и плотностью состояний, то и концентрация их – величина равновесная. К обычному, тепловому механизму генерации добавился новый канал (фотогенерация), но это привело только к увеличению их концентрации и, возможно, более серьезному проявлению принципа Паули. Иными словами, их распределение по энергиям описывается статистикой Ферми, в рамках которой изменение концентрации при неизменной температуре возможно только при изменении энергии Ферми.

Введем формально понятие *квазиуровней Ферми*  $F_n, F_p$  таким образом, чтобы адекватно описать изменившиеся концентрации. Вдали от вырождения их взаимосвязь описывается экспоненциальной зависимостью:

$$n = n_0 + \delta n = N_c \exp \frac{F_n - E_c}{kT}, \quad (2.5.29a)$$

$$p = p_0 + \delta p = N_v \exp \frac{E_v - F_p}{kT}. \quad (2.5.29б)$$

Положения квазиуровней Ферми можно отсюда определить, зная  $\delta n$ ,  $\delta p$ .

В этих обозначениях закон действующих масс:

$$np = n_i^2 \exp \frac{F_n - F_p}{kT}, \quad (2.5.30)$$

т.е. увеличение концентрации носителей определяется величиной расщепления квазиуровней Ферми.

На рис. 2.5.4 показан примерный ход квазиуровней Ферми в длинном образце, в котором облучается только некоторая область в середине. К образцу приложена разность потенциалов  $u$ .

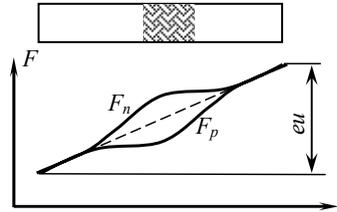


Рис. 2.5.4. Расщепление квазиуровней Ферми в освещенном образце, к которому приложена внешняя разность потенциалов  $u$ .

Ранее, см. соотношения (2.4.12а,б), мы видели, что плотность тока равна произведению концентрации, подвижности и градиента уровня Ферми. В нашем случае нужно отдельно определять плотности дырочного и электронного токов, а они определяются градиентом квазиуровней Ферми:

$$j_n = \mu_n n(\mathbf{r}) \nabla F_n = \sigma_n \frac{\nabla F_n}{e}, \quad j_p = \mu_p p(\mathbf{r}) \nabla F_p = \sigma_p \frac{\nabla F_p}{e}, \quad (2.5.31)$$

где введены величины электронной и дырочной проводимости,

$$\sigma_n = e \mu_n n(\mathbf{r}), \quad \sigma_p = e \mu_p p(\mathbf{r}). \quad (2.5.32)$$

С другой стороны, согласно закону Ома:

$$j_{nx} = \sigma_n (\mathcal{E}_x + \mathcal{E}_{nx}^*), \quad j_{px} = \sigma_p (\mathcal{E}_x + \mathcal{E}_{px}^*), \quad (2.5.33)$$

где  $E_x$  – приложенное внешнее кулоновское поле,  $E_{nx}^*$ ,  $E_{px}^*$  – поля сторонних сил. Для электронов и дырок они могут различаться.

В примере, приведенном на рис. 2.5.4, есть области, в которых  $\nabla F_n \approx 0$  или  $\nabla F_p \approx 0$ . Следовательно, в этих областях проводимость становится монополярной, для второй компоненты поле сторонних сил компенсирует внешнее кулоновское поле.

### 2.5.5. Инжекция в $p$ - $n$ -переходе, биполярный транзистор

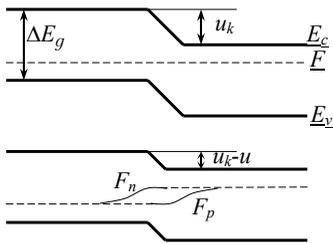


Рис. 2.5.5. Инжекция неосновных носителей при прямом смещении  $p$ - $n$ -перехода.

Мы уже рассматривали инжекцию носителей – в токах, ограниченных пространственным зарядом, см. разд. 2.4.3.1. Но это – инжекция основных носителей, на антизапорном контакте полупроводника с металлом. Аналогичная инжекция, но уже неосновных носителей, происходит и при прямом смещении  $p$ - $n$ -перехода, рис. 2.5.5.

Приложенное внешнее напряжение целиком падает на самом переходе, т.к. это – область, обедненная носителями. При этом, поскольку в переходе стал велик  $\nabla F$ , возникнут потоки носителей в соседние области. На рис. 2.5.5 электроны – влево, дырки – вправо. Эти потоки направлены на уменьшение тянущей силы, градиента уровня Ферми. И оба типа носителей в какой-то мере с этим справляются. Но в результате в областях близ перехода произведение  $np$  возрастает (см. (2.5.30)), возникает расщепление уровня Ферми на квазиуровни.

Увеличение концентраций ведет и к увеличению темпа рекомбинации, так что на некотором расстоянии от перехода концентрации снова станут термически равновесными.

Детальное рассмотрение возникающих токов и эффективности инжекции (см., например, [15], §VII.1) показывают, что в данном случае вольтамперная зависимость – та же, что мы получили для запорного контакта (2.4.88):

$$j = j_s [\exp(-\alpha u) - 1], \quad (2.5.34)$$

а ток насыщения при обратном смещении,

$$j_s = \frac{eD_p p_n}{L_p} + \frac{eD_n n_p}{L_n} = e \left( \frac{L_p p_n}{\tau_p} + \frac{L_n n_p}{\tau_n} \right), \quad (2.5.35)$$

определяется диффузией неосновных носителей, дырок из  $n$ - в  $p$ -область и наоборот, электронов из  $p$ - области в  $n$ -.

Для обеих областей эта инжекция может рассматриваться как генерация неравновесных основных носителей. Темп этой генерации равен темпу их термической генерации в приконтактных областях:

$$g_{pT} = \frac{p_n}{\tau_p}, \quad g_{nT} = \frac{n_p}{\tau_n} \quad (2.5.36)$$

Инжекция в  $p$ - $n$ -переходе используется в обычных биполярных транзисторах, рис. 2.5.6. Это – два включенных последовательно  $p$ - $n$ -перехода. В данном случае изображен так называемый  $n$ - $p$ - $n$ -транзистор.

Эмиттер в схемах обычно заземляется. При разомкнутой базе ток между эмиттером и коллектором течь не может, т.к. при любой полярности  $U_2$  один из переходов будет заперт. Если на коллектор подать положительный потенциал, то в переходе база-коллектор возникнет только ток насыщения, который стараются

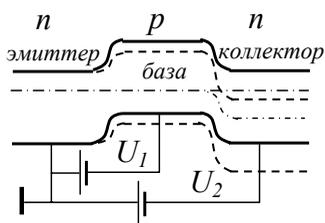


Рис. 2.5.6. Биполярный транзистор.

сделать малым. Из базы будут вытягиваться неосновные носители. Но если и на базу подать положительное смещение относительно эмиттера, то в нее будут инжектироваться именно неосновные носители, которые могут быть затянуты и в коллекторную цепь. На рис. штриховой линией показаны зоны

коллектора и базы при таком смещении.

Ток эмиттера  $I_e$  будет поделен между коллекторной  $I_c$  и базовой  $I_b$  цепями. В каком отношении он будет поделен, определяется тонкостями технологии изготовления транзистора и технология постоянно совершенствуется. Отношение этих токов,  $\beta = I_c/I_b$  — один из основных параметров транзистора. У первых промышленных образцов было  $\beta < 10$ , у современных — до нескольких сотен, а у составных — до тысяч.

Принципиально эмиттерный и коллекторных переходы одинаковы, различаются только размерами и формой. Коллекторный переход больше. К базе одно из основных требований — малая толщина, меньше диффузионной длины пробега.

Транзистор  $p-n-p$ -типа устроен полностью аналогично, но работает не при положительных, а при отрицательных смещениях.

## 2.5.6. Нейтральные пакеты неравновесных носителей

### 2.5.6.1. Амбиполярная диффузия

Один из важных вопросов, который мы пока не затрагивали — закон движения нейтрального облака неравновесных носителей. Оно может быть создано освещением.

Представим себе, что поверхность облучается сильно поглощающимся светом из области собственного поглощения. В тонком приповерхностном слое свет генерирует равные концентрации обоих типов неравновесных носителей. Зарядовое равновесие здесь не нарушается. Но подвижности и коэффициенты диффузии носителей различны.

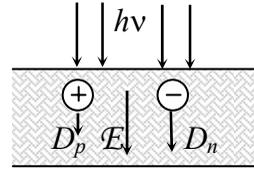


Рис. 2.5.7. Амбиполярная диффузия.

Предположим, что  $D_n > D_p$ . Тогда электроны быстрее удаляются от поверхности и возникает поле, препятствующее дальнейшему расплыванию облака неравновесных носителей. Оно может перемещаться только как единое целое, с общим коэффициентом диффузии, называемым **коэффициентом амбиполярной диффузии**.

Его можно определить из уравнений для токов:

$$j = j_p + j_n = \sigma E + e(D_n \nabla n - D_p \nabla p), \quad (2.5.37)$$

откуда внутреннее поле равно:

$$E = \frac{j}{\sigma} - \frac{e}{\sigma}(D_n \nabla n - D_p \nabla p) = E_0 + E_a. \quad (2.5.38)$$

Здесь первый член – приложенное внешнее поле, возбуждающее ток  $j$ , второй – поле амбиполярной диффузии:

$$E_a = \frac{e}{\sigma}(D_p \nabla p - D_n \nabla n). \quad (2.5.39)$$

Если пренебречь перезарядкой примесей, которая может исказить начальное равенство концентраций неравновесных носителей, то можно принять, что  $\delta n = \delta p$ ,  $\nabla n = \nabla p$ .

Тогда, подставляя (2.5.38) в (2.4.1), перепишем полную плотность тока как сумму дрейфового и диффузионного:

$$j_n = \frac{\sigma_n}{\sigma} j + eD\nabla n, \quad j_p = \frac{\sigma_p}{\sigma} j - eD\nabla p, \quad (2.5.40)$$

где 
$$D = \frac{D_p\sigma_n + D_n\sigma_p}{\sigma}, \quad (2.5.41)$$

а  $\sigma = \sigma_p + \sigma_n = e\mu_p p + e\mu_n n$  – полная проводимость полупроводника.

Воспользовавшись еще соотношением Эйнштейна (2.4.6), получим:

$$D = \frac{p+n}{\frac{p}{D_n} + \frac{n}{D_p}} = \frac{(p+n)D_p D_n}{pD_p + nD_n}, \quad \frac{1}{D} = \frac{1}{p+n} \left[ \frac{p}{D_n} + \frac{n}{D_p} \right]. \quad (2.5.42)$$

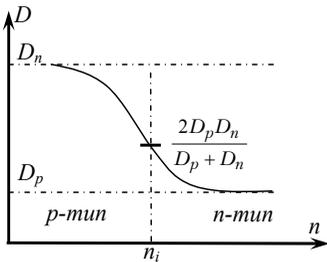


Рис. 2.5.8. Зависимость коэффициента амбиполярной диффузии от концентраций носителей.

Коэффициент амбиполярной диффузии – некоторая средняя взвешенная величина, зависящая не только от коэффициентов диффузии носителей, но и от их концентраций.

На рис. 2.5.8 принято, что  $D_n > D_p$ , а шкала концентраций электронов охватывает весь диапазон от  $p \gg n$  до  $n \gg p$ .

В собственном полупроводнике, при  $n = p \equiv n_i$ , величина,

обратная коэффициенту амбиполярной диффузии, равна полусумме обратных коэффициентов диффузии обоих носителей. При неравенстве концентраций скорость диффузии определяют неосновные носители.

### 2.5.6.2. Амбиполярный дрейф

Предположим, что вытягивающее поле однородно, направлено по оси  $x$  и столь велико, что диффузией можно пренебречь. Так же, как и раньше, пренебрежем рекомбинацией и будем считать, что перезарядка локальных центров не существенна. В таком случае  $\delta n = \delta p$ ,  $\frac{\partial p}{\partial t} = \frac{\partial n}{\partial t}$ ,  $\tau_n = \tau_p = \tau$ .

Уравнения для токов:

$$j_p = \sigma_p(x, t)\mathcal{E}(x, t), \quad j_n = \sigma_n(x, t)\mathcal{E}(x, t). \quad (2.5.43)$$

Уравнения непрерывности:

$$\begin{aligned} \frac{\partial p}{\partial t} &= -\frac{1}{e} \operatorname{div} j_p = -\frac{1}{e} \left( \sigma_p \frac{\partial \mathcal{E}}{\partial x} + \mathcal{E} \frac{\partial \sigma_p}{\partial x} \right), \\ \frac{\partial n}{\partial t} &= \frac{1}{e} \operatorname{div} j_n = \frac{1}{e} \left( \sigma_n \frac{\partial \mathcal{E}}{\partial x} + \mathcal{E} \frac{\partial \sigma_n}{\partial x} \right). \end{aligned} \quad (2.5.44)$$

$$\text{Здесь} \quad \sigma_p = e p \mu_p, \quad \sigma_n = e n \mu_n, \quad \sigma = \sigma_p + \sigma_n \quad (2.5.45)$$

и отсутствует внешняя генерация,  $g_n = g_p = 0$ . Мы рассматриваем дрейф уже образованного пакета носителей.

Домножая уравнения (2.5.44) на  $\sigma_n$  и  $\sigma_p$ , соответственно, и складывая их, избавимся от  $\frac{\partial \mathcal{E}}{\partial x}$  и получим:

$$-\frac{\partial p}{\partial t} = \frac{\mu_p \mu_n (n - p)}{\mu_p p + \mu_n n} \mathcal{E} \frac{\partial p}{\partial x} = v_{dr} \frac{\partial p}{\partial x}, \quad (2.5.46)$$

где стоящая справа величина

$$v_{dr} = \frac{\mu_p \mu_n (n - p)}{\mu_p p + \mu_n n} \mathcal{E} = \mu \mathcal{E} \quad (2.5.47)$$

есть скорость амбиполярного дрейфа, а

$$\mu = \frac{\mu_p \mu_n (n - p)}{\mu_p p + \mu_n n} - \quad (2.5.48)$$

подвижность пакета. Выражение для нее практически то же, что и для диффузии, но в числителе – разность концентраций. При  $n = p$  подвижность равна нулю. При разных концентрациях, во-первых, подвижность не равна нулю, хотя мы рассматривали нейтральное облако,  $\delta n = \delta p$ . Во-вторых, доминирующее значение, как и для дрейфа, имеет подвижность и направление (по/против поля) движения неосновных носителей.

### 2.5.6.3. *Общий случай*

Теперь попробуем объединить полученные результаты по дрейфу и подвижности. Учтем также генерацию и рекомбинацию, но генерацию по-прежнему считаем собственной,  $g_n = g_p = g$ .

$$\begin{aligned} \frac{\partial p}{\partial t} &= g - \frac{1}{e} \mathbf{E} \nabla \sigma_p - \frac{1}{e} \sigma_p \operatorname{div} \mathbf{E} + D_p \operatorname{div}(\nabla p) - \frac{\delta p}{\tau_p}, \\ \frac{\partial n}{\partial t} &= g + \frac{1}{e} \mathbf{E} \nabla \sigma_n + \frac{1}{e} \sigma_n \operatorname{div} \mathbf{E} + D_n \operatorname{div}(\nabla n) - \frac{\delta n}{\tau_n}. \end{aligned} \quad (2.5.49)$$

Исключив отсюда, как и раньше,  $\operatorname{div} \mathbf{E}$ , получим:

$$\begin{aligned} \frac{\partial p}{\partial t} &= g - \frac{1}{e \sigma} (\sigma_n \nabla \sigma_p - \sigma_p \nabla \sigma_n) \mathbf{E} + \\ &+ \frac{\sigma_n D_p + \sigma_p D_n}{\sigma} \operatorname{div}(\nabla p) - \frac{\delta p}{\tau_p}. \end{aligned} \quad (2.5.50)$$

Это выражение можно упростить, заметив, что второй член:

$$\begin{aligned} \frac{1}{e\sigma}(\sigma_n \nabla \sigma_p - \sigma_p \nabla \sigma_n) &= \frac{n-p}{\frac{p}{\mu_n} + \frac{n}{\mu_p}} \nabla p = \\ &= \frac{(n-p)\mu_p \mu_n}{p\mu_p + n\mu_n} \nabla p = \mu \nabla p, \end{aligned} \quad (2.5.51)$$

и далее:  $\mu \nabla p \mathbf{E} = v_{dr} \nabla p$ ,

а в третьем, используя соотношения Эйнштейна (2.4.6), получим:

$$\frac{\sigma_n D_p + \sigma_p D_n}{\sigma} = \frac{n\mu_n D_p + p\mu_p D_n}{p\mu_p + n\mu_n} = \frac{(n+p)D_p D_n}{pD_p + nD_n} = D \quad (2.5.52)$$

и далее преобразуем его к виду:

$$D \operatorname{div}(\nabla p) = \operatorname{div}(D \nabla p) - \nabla D \cdot \nabla p. \quad (2.5.53)$$

Итого получили:

$$\frac{\partial p}{\partial t} = g - v \nabla p + \operatorname{div}(D \nabla p) - \frac{\delta p}{\tau_p}, \quad (2.5.54)$$

где  $\mathbf{v} = \frac{n-p}{\frac{p}{\mu_n} + \frac{n}{\mu_p}} \mathbf{E} + \nabla D -$  (2.5.55)

– скорость перемещения пакета за счет дрейфа и диффузии.

Три последних члена в (2.5.54) – изменения концентрации за счет его перемещения, диффузного размытия и рекомбинации.

#### 2.5.6.4. Длины диффузии и дрейфа

Рассмотрим распространение неравновесных носителей из области генерации.

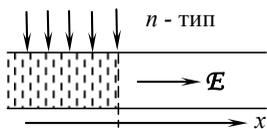


Рис. 2.5.9.

Пусть в тонком однородном образце, рис. 2.5.9, например,  $n$ -типа, генерация происходит в области  $x < 0$ . Концентрации  $\delta p$ ,  $\delta n$  - малы, так что  $\mu$ ,  $D$ ,  $E$  можно считать неизменными во всем интересующем нас объеме. В стационарном состоянии, при  $x \geq 0$

$\frac{\partial p}{\partial t} = 0$ ,  $g = 0$  и (2.5.54) можно переписать в виде:

$$(\delta p)'' - \frac{2}{\lambda} (\delta p)' - \frac{\delta p}{L^2} = 0, \quad (2.5.56)$$

где 
$$\lambda = \frac{2D}{v} = \frac{2D}{\mu E}, \quad L = \sqrt{D\tau}. \quad (2.5.57)$$

Решения (2.5.56) имеют вид:

$$\delta p = a \exp(k_1 x) + b \exp(k_2 x), \quad (2.5.58)$$

где

$$k_{1,2} = \frac{1}{\lambda} \pm \sqrt{\frac{1}{\lambda^2} + \frac{1}{L^2}} = \frac{1}{\lambda} \left( 1 \pm \sqrt{1 + \frac{\lambda^2}{L^2}} \right). \quad (2.5.59)$$

**При  $E > 0$  в образец затягиваются неосновные носители, дырки.** При этом  $\lambda > 0$ ,  $k_1 > 0$ ,  $k_2 < 0$ , так что из граничных условий:

$$x = 0, \quad \delta p(0) = \delta p_0, \quad \delta p(x \rightarrow \infty) \rightarrow 0 \quad (2.5.60)$$

следует:  $a = 0$ ,  $b = \delta p_0$ , так что концентрация убывает по экспоненте:

$$\delta p = \delta p_0 \exp\left(-\frac{x}{l}\right), \quad (2.5.61)$$

где

$$\frac{1}{l} = -k_2 = \frac{1}{\lambda} \left( \sqrt{1 + \left(\frac{\lambda}{L}\right)^2} - 1 \right). \quad (2.5.62)$$

При малых вытягивающих полях, когда  $\left(\frac{\lambda}{L}\right)^2 \gg 1$ ,

$$l = L = \sqrt{D\tau}, \quad (2.5.63)$$

так что проникновение носителей происходит только за счет диффузии.

Если же поле сильно, так что  $\left(\frac{\lambda}{L}\right)^2 \ll 1$ , получим:

$$l(E) \approx 2L^2/\lambda = \mu E \tau. \quad (2.5.64)$$

В этом случае все определяет дрейф, а диффузия оказывается не существенна.

**При  $E < 0$  в образец затягиваются основные носители, электроны.** В этом случае получим

$$\frac{1}{l} = -k_1 = \frac{1}{|\lambda|} \left( \sqrt{1 + \left(\frac{\lambda}{L}\right)^2} + 1 \right). \quad (2.5.65)$$

Поле, вытягивающее основные носители, только препятствует их проникновению в область  $x \geq 0$ . В малых полях

$$\frac{1}{l} \approx \frac{1}{|\lambda|} \left[ 1 + \frac{|\lambda|}{L} \right] = \frac{1}{|\lambda|} + \frac{1}{L} = \frac{\mu E}{2D} + \frac{1}{L}, \quad (2.5.66)$$

так что всегда  $l \leq L$ , а в сильных

$$\frac{1}{l} \approx \frac{1}{|\lambda|} \left[ 2 + \frac{1}{2} \left( \frac{\lambda}{L} \right)^2 \right] = \frac{4L^2 + \lambda^2}{2L^2 |\lambda|} \approx \frac{2}{|\lambda|}, \quad l \approx \frac{|\lambda|}{2} = \frac{D}{\mu E} \quad (2.5.67)$$

И при  $E \rightarrow \infty$   $l \rightarrow 0$ . Поле "запирает" диффузию носителей.

### 2.5.7. Рекомбинация

Рекомбинация электрона и дырки в любом случае – процесс случайный и, возможно, многостадийный. Вопрос о том, сколько стадий он включает, сколько частиц в них участвует и каков тип этих частиц – первый, с которым нужно разобраться прежде, чем мы сможем определить темп рекомбинации. Ясно, что чем меньше количество частиц, участвующих в данной стадии процесса, тем она более вероятна. Но это – при прочих равных условиях.

В принципе, можно утверждать, что в рекомбинации должно участвовать не менее трех частиц: сами рекомбинирующие электрон и дырка и еще минимум одна частица для сохранения импульса и энергии. Даже если сталкиваются носители с суммарным импульсом, равным нулю, так что закон сохранения импульса не будет нарушен, нужно куда-то передать довольно большую энергию, в среднем примерно на  $3kT$  больше ширины запрещенной зоны. Но требование равенства импульсов<sup>1</sup> рекомбинирующих носителей – слишком жестко, от таких процессов едва ли можно ожидать большой эффективности.

Возможна рекомбинация и через какие-либо промежуточные состояния, через дефекты. В этом случае она проходит минимум в два этапа: дефект должен сначала захватить один носитель, затем – другой. Здесь, правда, возникает еще один вопрос: если дефект захватил, например, электрон, то что более вероятно, захват

---

<sup>1</sup> Импульсы электрона и дырки в одном и том же квантовом состоянии различаются знаком. Здесь имеется в виду ситуация, в которой импульс электрона не изменяется при рекомбинации.

дырки (т.е. завершение акта рекомбинации) или возбуждение этого электрона обратно в с-зону? Ясно, что вероятности обоих этих процессов конечны и нужно рассматривать оба.

Если в рекомбинации участвует какой-либо локальный центр, то она может включать не две, а больше стадий. Например, захват электрона происходит на возбужденный уровень центра, затем следует релаксация этого возбуждения и в конце – захват дырки на нижний уровень.

Все эти типы процессов, способных привести к рекомбинации носителей, сильно различаются по эффективности хотя бы потому, что в них различны условия для сохранения импульса и энергии. Так, в процессах с участием дефектов сохранение импульса не сложно, поскольку массы носителя и дефекта сильно различны. Кроме того, разделение процесса рекомбинации на два или более этапов уменьшает энергию, выделяющуюся в каждом из них. Эти факторы очень существенны и, как мы увидим позднее, обычно наиболее вероятно именно рекомбинация через дефекты.

Итак, все эти процессы можно разделить на два типа:

- прямая рекомбинация зона-зона или
- рекомбинация через локальные состояния, связанные с дефектами.

При прямой межзонной рекомбинации, в принципе, есть две возможности для сохранения энергии: излучение кванта света, либо передача ее третьему носителю, электрону или дырке. Их называют, соответственно, *излучательная* и *Оже-* или *ударная рекомбинация*. Ясно, что второй процесс возможен только при большой концентрации носителей, в сильно легированных полупроводниках. Поэтому мы рассмотрим только излучательную рекомбинацию. С Оже-рекомбинацией можно ознакомиться, например, по [17], главы 11, 12.

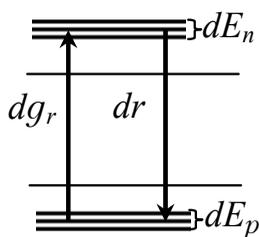


Рис. 2.5.10. Переходы между группами состояний в  $v$ - и  $c$ -зонах.

**Излучательная рекомбинация** возможна также через экситонные состояния и при участии дефектов. Излучение возникает либо при релаксации возбуждения в дефекте, либо при так называемой донорно-акцепторной рекомбинации, когда зарядами обмениваются два рядом расположенных разнотипных центра.

### 2.5.7.1. Рекомбинация зона-зона

Рассмотрим переходы между группами состояний в  $c$ - и  $v$ -зонах,  $dE_n$  и  $dE_p$ , рис. 2.5.10. Рекомбинация при переходах между этими состояниями пропорциональна концентрации электронов в верхних состояниях и концентрации дырок в нижних:

$$dr = W(E_n, E_p) N_c(E_n) N_v(E_p) f(E_n) f_p(E_p) dE_n dE_p. \quad (2.5.68)$$

Здесь  $W(E_n, E_p)$  – вероятность перехода в единицу времени,  $N_c(E_n)$ ,  $N_v(E_p)$  – плотности состояний в зонах. Полный темп рекомбинации – двойной интеграл по всем верхним и нижним состояниям.

По-прежнему будем считать, что неравновесные носители термализованы и их распределения по состояниям характеризуются квазиуровнями Ферми. Кроме того, и квазиуровни Ферми находятся достаточно далеко от краев зон, так что:

$$f(E_n) = \exp \frac{F_n - E_n}{kT}, \quad f_p(E_p) = \exp \frac{E_p - F_p}{kT}. \quad (2.5.69)$$

В этом случае концентрации носителей соотносятся с равновесными как  $\frac{np}{n_0p_0} = \exp\frac{F_n - F_p}{kT}$ . Тогда произведение функций заполнения:

$$\begin{aligned} f(E_n)f_p(E_p) &= \frac{np}{n_0p_0} \exp\frac{E_p - E_n}{kT} = \\ &= \frac{np}{N_c N_v} \exp\left(-\frac{E_n - E_c}{kT}\right) \exp\left(\frac{E_p - E_v}{kT}\right). \end{aligned} \quad (2.5.70)$$

Здесь  $N_c, N_v$  – эффективные плотности состояний в зонах, см. (2.2.11),  $E_c, E_v$  – энергии краев зон.

Подставив это соотношение в (2.5.68), получим выражение, которое удобно записать в виде:

$$r = \alpha np, \quad (2.5.71)$$

а величина

$$\begin{aligned} \alpha &= \frac{1}{N_c N_v} \int_{E_n=E_c}^{\infty} \int_{E_p=-\infty}^{E_n} W(E_n, E_p) N_c(E_n) N_v(E_p) \times \\ &\times \exp\left(-\frac{E_n - E_c}{kT}\right) \exp\left(\frac{E_p - E_v}{kT}\right) dE_n dE_p \end{aligned} \quad (2.5.72)$$

– коэффициент рекомбинации. Входящую сюда вероятность перехода  $W(E_n, E_p)$  можно вычислить, но нам сейчас нужно оценить, чем определяется и как найти величину  $\alpha$ . Она имеет размерность  $\left[ \text{см}^3/\text{с} \right]$ .

Полный темп рекомбинации зона-зона оказался, что вполне естественно, пропорционален произведению концентраций носителей. Реакции подобного типа в химии называются бимолеку-

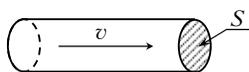


Рис. 2.5.11 Сечение рекомбинации.

лярными, а потому и рекомбинация с такой концентрационной зависимостью обычно называется бимолекулярной.

Темп рекомбинации избыточных носителей можно определить, вспомнив (см. разд. 2.5.1), что в термодинамическом равновесии темпы генерации и рекомбинации равны, так что:

$$R = r - g_T = \alpha np - g_T \quad \text{и} \quad g_T = \alpha n_0 p_0 = \alpha n_i^2.$$

Отсюда:

$$R = \alpha(np - n_0 p_0). \quad (2.5.73)$$

Сечение рекомбинации – очень удобный параметр. В основном, для расчета темпов захвата и рекомбинации на дефектах, но и в межзонной рекомбинации он также может быть использован. Представим себе, что электрон рекомбинирует, столкнувшись с некоторой площадкой  $S$ , рис. 2.5.11. Пусть это – сечение сферы, в пределах которой дырка захватит электрон. Число таких актов в единицу времени пропорционально числу электронов в "трубке", изображенной на рисунке, и концентрации дырок. Отсюда получаем  $r = (Sv_T n)p = (Sv_T)np = \alpha np$ , так что:

$$\alpha = Sv_T. \quad (2.5.74)$$

При этом мы приписали электрону среднюю тепловую скорость  $v_T$ . В расчетах часто используют скорость свободного электрона, имеющего массу  $m_0$ .

Пусть неравновесные носители генерируются оптическими перебросами зона-зона. Тогда  $\delta n = \delta p$ ,  $n = n_0 + \delta n$ ,  $p = p_0 + \delta n$  и

$$R = \alpha(np - n_0 p_0) \approx \alpha \cdot \delta n(n_0 + p_0 + \delta n). \quad (2.5.75).$$

Отсюда время жизни носителя в зоне равно:

$$\tau = \frac{\delta n}{R} = \frac{1}{\alpha(n_0 + p_0 + \delta n)} \quad (2.5.76)$$

При легировании, когда концентрации носителей одного типа существенно увеличивается, увеличивается и сумма концентраций, так что время жизни уменьшается. Оно может быть только меньше времени жизни в собственном полупроводнике:

$$\tau \leq \tau_i = \frac{1}{2\alpha n_i}. \quad (2.5.77)$$

### 2.5.7.2. *Излучательные межзонные переходы*

Время жизни по отношению к излучательной рекомбинации зона-зона можно оценить, зная спектры излучения и поглощения вещества в равновесном состоянии. Согласно принципу детального равновесия, см. разд. 2.6.2, в термодинамическом равновесии темпы оптической генерации и излучательной рекомбинации равны.

Оптическая генерация в отсутствие внешнего воздействия может происходить только под действием равновесного излучения, интенсивность которого описывается формулой Планка ( $n$  – показатель преломления):

$$\rho_0(\hbar\omega) = \frac{\omega^2 n^3}{\pi^2 c^3 \hbar} \cdot \frac{1}{\exp \frac{\hbar\omega}{kT} - 1}. \quad (2.5.78)$$

Объемная плотность фотонов с энергией в диапазоне от  $\hbar\omega$  до  $\hbar\omega + d(\hbar\omega)$  равна  $\rho_0(\hbar\omega)d(\hbar\omega)$ .

Обозначим  $\nu(\hbar\omega)$  – коэффициент фотоэлектрического поглощения (см. раздел 2.5.3), а  $P(\hbar\omega)$  – вероятность поглощения кванта в единицу времени. Тогда темп оптической генерации:

$$g_T = \int_0^{\infty} \rho_0(\hbar\omega) P(\hbar\omega) \nu(\hbar\omega) d(\hbar\omega). \quad (2.5.79)$$

Скорость затухания интенсивности в пространстве описывается коэффициентом поглощения  $\gamma$ :  $\frac{d\rho}{dx} = -\gamma\rho$ , и, поскольку свет движется со скоростью  $c/n$ , то:  $\frac{d\rho}{dt} = -\frac{\gamma c}{n}\rho$ . С другой стороны, согласно нашим обозначениям,  $\frac{d\rho}{dt} = -P(\hbar\omega)\rho$ . Следовательно,

$$P(\hbar\omega) = \gamma(\hbar\omega) \frac{c}{n(\hbar\omega)} \quad (2.5.80)$$

и отсюда, обозначив  $u = \frac{\hbar\omega}{kT}$ , получим:

$$g_T = \frac{c}{n(u)} \int_0^{\infty} \rho_0(u) \gamma(u) \nu(u) d(u) = \frac{(kT)^3}{\pi^2 c^2 \hbar^3} \int_0^{\infty} \frac{n^2(u) \gamma(u) \nu(u) u^2 du}{e^u - 1}. \quad (2.5.81)$$

Темп генерации – интеграл от произведения спектра поглощения на планковскую плотность излучения, которая в области энергий порядка ширин зон практически важных полупроводников экспоненциально спадает. Поэтому можно ожидать, что излучательная рекомбинация – процесс мало вероятный.

На рис. 2.5.12 показаны соответствующие спектры для Ge при комнатной температуре, а в таблице 2.5.1 ([13], гл.7) приведены излучательные времена жизни неосновных носителей в некоторых полупроводниках, собственных и легированных до концентрации основных носителей  $10^{17} \text{см}^{-3}$ .

Таблица 2.5.1

	Собств.	$n = 10^{17} \text{см}^{-3}$
Si	4,6 ч	2,5 мс
Ge	0,61 с	0,15 мс
GaAs	2,8 мкс	0,04 мкс
InAs	15 мкс	0,24 мкс
InSb	0,62 мкс	0,12 мкс

Действительно, спектр рекомбинационного излучения хорошо воспроизводит спектр  $\rho_0\gamma$ , показанный на рисунке. Правда, в наблюдаемом спектре, в спектре излучения, выходящего из кристалла, относительные интенсивности двух пиков искажены. Коротковолновое излучение сильнее поглощается, так что до поверхности доходит свет из более тонкого слоя, чем длинноволновый.

В таблице 2.5.1 следует отметить две основных закономерности. Во-первых, существенно различны времена жизни Ge, Si и двухкомпонентных соединений. Причина этого – в величине коэффициента поглощения. Ge и Si – многодолинные, не прямозонные полупроводники. Экстремумы с-зон у них находятся на краю (Ge) или близ края (Si) зоны Бриллюэна, так что в оптическом переходе должен участвовать еще и фонон. Это существенно уменьшает вероятность перехода. А двухкомпонентные соединения – прямозонные, имеют

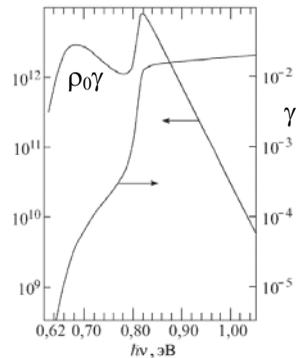


Рис. 2.5.12. Спектры поглощения Ge при 300 K и вероятности излучения, [13].

резкий край и большой коэффициент поглощения.

Во-вторых, легирование сильно увеличивает скорость рекомбинации, см. (2.5.77). В легированных прямозонных полупроводниках этот канал рекомбинации может быть определяющим.

### 2.5.7.3. Рекомбинация через локальные центры

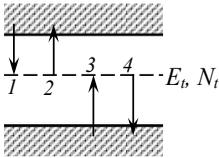


Рис. 2.5.13. Типы электронных переходов с участием локальных центров.

Рассмотрим одно какое-либо локальное состояние и не будем принимать во внимание его возможные возбуждения.

На рис. 2.5.13 показаны четыре типа переходов, которые могут происходить с участием такого центра. Это – захват электрона (1) или дырки (4) и обратный заброс в зону электрона (2) или дырки (3). Мы будем считать, что процессы 2 и 3 происходят только за счет тепловой энергии.

энергии.

Такой центр может работать и как центр рекомбинации, при последовательных процессах 1, 4, и как центр захвата, как ловушка, способная на некоторое время захватить носитель и потом вернуть его в зону. Основные его характеристики – энергия электрона на нем,  $E_t$ , и концентрация центров  $N_t$ .

Количество электронов на этих уровнях равно  $N_t f$ , количество дырок –  $N_t(1-f)$ , где  $f$  – пока не известная функция заполнения. В термодинамическом равновесии это  $f_0$ , функция Ферми, но при возникновении неравновесных носителей в зонах концентрация захваченных носителей уже не описывается статистикой Ферми.

Рассмотрим группу электронов в зоне проводимости, с энергией в интервале  $(E, E + dE)$ . Число переходов в единицу времени из зоны на этот уровень можно представить как

$$r_n = N_t [1 - f(E_t)] \int_{E_c}^{\infty} c_n(E) N_c(E) f(E) dE, \quad (2.5.82)$$

где  $c_n(E)$  – коэффициент, характеризующий вероятность захвата электрона свободной ловушкой.

Это выражение, опять-таки, удобно записать в виде:

$$r_n = \alpha_n N_t [1 - f(E_t)] n, \quad (2.5.83)$$

где  $\alpha$  – коэффициент рекомбинации, равный в данном случае

$$\alpha_n = \frac{\int_{E_c}^{\infty} c_n N_c(E) f(E) dE}{n} = \frac{\int_{E_c}^{\infty} c_n N_c(E) f(E) dE}{\int_{E_c}^{\infty} N_c(E) f(E) dE} = \langle c_n \rangle = S_n v_t. \quad (2.5.84)$$

Здесь, как и в (2.5.74), введено понятие сечения захвата  $S_n$ .

Обратный переброс, реэмиссия электрона в с-зону, может быть записан как  $g_{nT} = \beta_n N_t f(E_t)$ , где  $\beta_n$  – некоторый коэффициент, зависящий от температуры и свойств центра, определяющий вероятность такого возбуждения. Его величину можно сопоставить с  $\alpha_n$ , воспользовавшись принципом детального равновесия (см. раздел 2.6.2). В термодинамическом равновесии  $g_{nT} = r_n$ , т.е.

$$\alpha_n N_t (1 - f_0) = \beta_n N_t f_0, \quad (2.5.85)$$

и, поскольку  $f_0 = \frac{n_0}{n_0 + n_1}$ , где  $n_1 = \frac{g_0}{g_1} N_c \exp\left(-\frac{J}{kT}\right)$ , см. (2.2.25) и (2.2.26), так что  $g_{nT} = \alpha_n n_1 \cdot N_t f(E_t)$ , то отсюда получаем:

$$\beta_n = \alpha_n \frac{1-f_0}{f_0} n_0 = \alpha_n n_1. \quad (2.5.86)$$

Теперь можно определить темпы рекомбинации неравновесных электронов:

$$R_n = r_n - g_{nT} = \alpha_n N_t [n(1-f) - n_1 f] \quad (2.5.87)$$

и аналогично – для дырок:

$$R_p = r_p - g_{pT} = \alpha_p N_t [pf - p_1(1-f)]. \quad (2.5.88)$$

Если токи малы, так что изменение концентраций во времени определяются только процессами генерации и рекомбинации, то уравнения непрерывности получим в виде:

$$\begin{aligned} \frac{\partial p}{\partial t} &= g_p - R_p = g_p - \alpha_p N_t [pf - p_1(1-f)], \\ \frac{\partial n}{\partial t} &= g_n - R_n = g_n - \alpha_n N_t [n(1-f) - n_1 f]. \end{aligned} \quad (2.5.89)$$

Здесь совершенно не обязательно  $\frac{\partial p}{\partial t} = \frac{\partial n}{\partial t}$ , т.к. может изменяться и концентрация захваченных электронов. Очевидно, что при выполнении условия квазинейтральности  $(p - p_0) - (n - n_0) = N_t (f - f_0)$ , скорость изменения концентрации заполненных ловушек будет равна:

$$N_t \frac{\partial f}{\partial t} = \frac{\partial p}{\partial t} - \frac{\partial n}{\partial t} = (g_p - g_n) - (R_p - R_n). \quad (2.5.90)$$

Этих трех уравнений, (2.5.89) и (2.5.90), достаточно для определения трех неизвестных,  $p'_t, n'_t, f'_t$ , если известны темп генерации и коэффициенты рекомбинации или сечения захвата, см. (2.5.84).

Величины сечений захвата зависят от зарядового состояния ловушки. Для того, чтобы носитель был захвачен нейтральным

центром, он должен оказаться в области порядка радиуса электронной оболочки инородного атома, в этом случае сечения захвата имеют порядок  $S \approx 10^{-15} \text{ см}^2$ .

Центр, заряженный противоположно, т.е. притягивающий носитель, имеет существенно большее сечение. Можно считать, что носитель затаянется на него из всего пространства, в котором энергия притяжения превышает  $kT$ . Сечение зависит от концентрации носителей, поскольку она определяет радиус экранирования. Но в среднем можно считать, что  $S \approx 10^{-12} \text{ см}^2$ .

Сечение захвата отталкивающего центра меньше за счет двух факторов. Во-первых, носитель должен иметь энергию больше высоты потенциального барьера, что сразу отсекает часть электронов. Нас же интересует усредненная характеристика, – среднее сечение захвата для тепловых электронов. Во-вторых, должен быть мал момент количества движения в паре носитель/ловушка. Иными словами, должна быть достаточно велика кинетическая энергия носителя, связанная с радиальной составляющей его скорости. В этом случае для однозарядных центров ориентировочно  $S \approx 10^{-17} - 10^{-18} \text{ см}^2$ .

Этот вопрос более подробно рассмотрен в [17], гл. 3, 4, 5.

#### **2.5.7.4. *Время жизни при монополярном возбуждении***

Здесь мы предполагаем, что  $g_p = 0$ ,  $g_n \neq 0$ . Такая инжекция возможна на контакте, либо электроны генерируются возбуждением из локальных центров, мало участвующих в последующей рекомбинации, либо генерация биполярна, но дырки эффективно захватываются.

---

<sup>1</sup>  $10^{-18} \text{ см}^2 = 1 \text{ МБарн}$ , 1 МегаБарн.

В этом случае остается только одно из уравнений (2.5.89) и, подставляя в него  $n = n_0 + \delta n$ ,  $f = f_0 + \delta f$ , и учитывая, что  $n_0(1 - f_0) = n_1 f_0$  (см. (2.2.25), (2.2.26)), получим:

$$\begin{aligned} \frac{\partial \delta n}{\partial t} &= g_n - \alpha_n N_t \left[ n_0 + \delta n - (n_0 + \delta n + n_1)(f_0 + \delta f) \right] = \\ &= g_n - \alpha_n N_t \left[ \delta n(1 - f_0) + (n_0 + n_1 + \delta n)\delta f \right]. \end{aligned} \quad (2.5.91)$$

Используя уравнение квазинейтральности, которое в данном случае упрощается до  $\delta n + N_t \delta f = 0$ , исключим  $\delta f$ :

$$\frac{\partial n}{\partial t} = g_n - \alpha_n \left( N_t^0 + n_0 + n_1 + \delta n \right) \delta n = g_n - \frac{\delta n}{\tau}, \quad (2.5.92)$$

где  $N_t^0 = N_t(1 - f_0)$  – равновесная концентрация ловушек, не занятых электроном. Отсюда:

$$\tau_n = \frac{1}{\alpha_n \left( N_t^0 + n_0 + n_1 + \delta n \right)}. \quad (2.5.93a)$$

При малом уровне возбуждения, или на начальном этапе, пока  $\delta n \ll N_t^0 + n_0 + n_1$ , оказывается  $\tau_n = Const$ .

Полностью аналогично для дырок мы получили бы:

$$\tau_p = \frac{1}{\alpha_p \left( N_t^- + p_0 + p_1 + \delta p \right)}, \quad (2.5.93b)$$

где  $N_t^- = N_t f$  – равновесная концентрация заполненных ловушек.

### 2.5.7.5. Времена жизни при биполярном возбуждении

Пусть  $g_p = g_n = g \neq 0$ . Представим, как и прежде, отклонения от равновесия как:  $\delta n = n - n_0$ ,  $\delta p = p - p_0$ ,  $\delta f = f - f_0$ . Тогда темп рекомбинации электронов:

$$\begin{aligned} R_n &= \alpha_n N_t [n(1-f) - n_1 f] = \\ &= \alpha_n N_t [\delta n(1-f_0) - \delta f(n_0 + n_1 + \delta n)]. \end{aligned} \quad (2.5.94)$$

Здесь в явном виде записана разность темпов рекомбинации и реэмиссии электрона в с-зону.

С помощью уравнения квазинейтральности  $N_t \delta f = \delta p - \delta n$ , исключим отсюда  $\delta f$  и результат представим в виде:

$$R_n = \alpha_n N_t [n(1-f) - n_1 f] = a \delta n + b \delta p. \quad (2.5.95a)$$

Для дырок все аналогично:

$$R_p = \alpha_p N_t [pf - p_1(1-f)] = c \delta p + d \delta n. \quad (2.5.95b)$$

Здесь использованы обозначения:

$$a = \alpha_n (N_t^0 + n_0 + n_1 + \delta n), \quad b = -\alpha_n (n_0 + n_1 + \delta n), \quad (2.5.96a)$$

$$c = \alpha_p (N_t^- + p_0 + p_1 + \delta p), \quad d = -\alpha_p (p_0 + p_1 + \delta p), \quad (2.5.96b)$$

Уравнения непрерывности в этом случае оказываются нелинейны:

$$\frac{\partial \delta n}{\partial t} = g - a \delta n - b \delta p = g - \frac{\delta n}{\tau_n}, \quad (2.5.97a)$$

$$\frac{\partial \delta p}{\partial t} = g - c \delta p - d \delta n = g - \frac{\delta p}{\tau_p}. \quad (2.5.97b)$$

Как и при монополярном возбуждении, при  $\delta n \ll n_0 + n_1$  и  $\delta p \ll p_0 + p_1$ , получим:  $\tau_n, \tau_p - Const$ . В этом случае и в (2.5.97) исчезают нелинейные члены и уравнения для неравновесных концентраций могут быть записаны как:

$$\frac{\partial^2 (\delta n)}{\partial t^2} + (a + c) \frac{\partial \delta n}{\partial t} + (ac - bd)\delta n + g(b - c), \quad (2.5.98a)$$

$$\frac{\partial^2 (\delta p)}{\partial t^2} + (a + c) \frac{\partial \delta p}{\partial t} + (ac - bd)\delta p + g(d - a). \quad (2.5.98б)$$

В общем виде решения этих уравнений:

$$\delta n = A_n + B_n \exp\left(-\frac{t}{\tau_1}\right) + C_n \exp\left(-\frac{t}{\tau_2}\right) \quad (2.5.99a)$$

$$\delta p = A_p + B_p \exp\left(-\frac{t}{\tau_1}\right) + C_p \exp\left(-\frac{t}{\tau_2}\right), \quad (2.5.99б)$$

где  $A_n = g \frac{c - b}{ac - bd}$ ,  $A_p = g \frac{a - d}{ac - bd}$ , а постоянные  $B$  и  $C$  должны быть определены из начальных условий. Величины  $\tau_1$  и  $\tau_2$  – времена релаксации неравновесных концентраций. Они являются корнями характеристического уравнения:

$$\left(\frac{1}{\tau}\right)^2 - (a + c) \frac{1}{\tau} + (ac - bd) = 0, \quad (2.5.100)$$

решение которого дает:

$$\frac{1}{\tau_{1,2}} = \frac{1}{2}(a + c) \pm \sqrt{\frac{1}{4}(a + c)^2 - (ac - bd)}. \quad (2.5.101)$$

Оба эти времени относятся и к электронам, и к дыркам. Кинетика затухания – биэкспоненциальная. Смысл и природу возникновения этих двух времен можно представить себе рассмотр-

рев предел очень высокого легирования, когда  $N_t^0 \gg n_0 + n_1$  и  $N_t^- \gg p_0 + p_1$ . Тогда  $a \gg b$ ,  $c \gg d$  и получаем:

$$\frac{1}{\tau} = \frac{1}{2}(a + c) \pm \frac{1}{2}(a - c), \quad (2.5.102)$$

т.е. 
$$\frac{1}{\tau_1} = a \approx \alpha_n N_t^0, \quad \frac{1}{\tau_2} = c \approx \alpha_p N_t^-, \quad (2.5.103)$$

времена  $\tau_1$  и  $\tau_2$  определяются темпами захвата электронов и дырок, соответственно.

### 2.5.7.6. Стационарные состояния

Если возбуждение имеется и процесс установился, т.е.  $\frac{\partial p}{\partial t} = \frac{\partial n}{\partial t} = 0$ , то уравнения непрерывности упрощаются и при монополярном возбуждении мы получим те же времена жизни, что и в переходный период, см. (2.5.93).

При биполярном возбуждении, поскольку концентрации уже не меняются, то справедливо:

$$R_n = R_p = R \quad (2.5.104)$$

и времена жизни можно определять как:

$$\tau_n = \frac{\delta n}{R(\delta n)}, \quad \tau_p = \frac{\delta p}{R(\delta p)}. \quad (2.5.105)$$

Рассмотрение упрощается при очень большом уровне инжекции, когда в уравнениях (2.5.95а), (2.5.95б) можно пренебречь реэмиссией (она всегда ограничена, поскольку ограничена концентрация ловушек и  $f \leq 1$ ) и можно считать, что  $n = p$ . Тогда равенство

$$R_n = R_p = R = \alpha_n N_t [n(1-f) - n_1 f] = \alpha_p N_t [pf - p_1(1-f)]$$

упрощается до  $\alpha_n(1-f) = \alpha_p f$ , откуда получаем:

$$f = \frac{\alpha_n}{\alpha_n + \alpha_p}, \quad (1-f) = \frac{\alpha_p}{\alpha_n + \alpha_p} \quad (2.5.106)$$

и тогда:

$$R = \alpha_n N_t n(1-f) = \alpha_p N_t pf = N_t n \frac{\alpha_n \alpha_p}{\alpha_n + \alpha_p}, \quad (2.5.107)$$

Следовательно, время жизни при большом уровне биполярной инжекции:

$$\tau_\infty = \frac{n}{R} = \frac{\alpha_n + \alpha_p}{\alpha_n \alpha_p N_t} = \tau_{n0} + \tau_{p0}, \quad (2.5.108)$$

где мы ввели обозначения:

$$\tau_{n0} = \frac{1}{\alpha_n N_t}, \quad \tau_{p0} = \frac{1}{\alpha_p N_t}. \quad (2.5.109)$$

Время жизни при высоком уровне возбуждения одинаково для электронов и дырок и не зависит от концентрации неравновесных носителей заряда, а определяется наибольшим из времен  $\tau_{n0}$ ,  $\tau_{p0}$ , т. е. наименьшим коэффициентом захвата, наиболее "узким" каналом рекомбинации.

В более общем случае из равенства

$$R = \alpha_n N_t [n(1-f) - n_1 f] = \alpha_p N_t [pf - p_1(1-f)]$$

следует, что функции распределения:

$$f = \frac{\alpha_n n + \alpha_p p_1}{\alpha_n (n + n_1) + \alpha_p (p + p_1)}, \quad (2.5.110a)$$

$$1 - f = \frac{\alpha_n n_1 + \alpha_p p}{\alpha_n (n + n_1) + \alpha_p (p + p_1)}, \quad (2.5.110б)$$

так что темп рекомбинации

$$R = \frac{np - n_0 p_0}{\tau_{p0} (n + n_1) + \tau_{n0} (p + p_1)}. \quad (2.5.111)$$

Если концентрация ловушек мала и их перезарядка не влияет существенно на концентрации, то  $\delta n = \delta p$ , так что их времена жизни одинаковы и темп рекомбинации равен:

$$R = \delta n \frac{n_0 + p_0 + \delta n}{\tau_{p0} (n_0 + n_1 + \delta n) + \tau_{n0} (p_0 + p_1 + \delta n)}, \quad (2.5.112)$$

т.е. времена жизни:

$$\tau_n = \tau_p = \tau = \frac{\tau_{p0} (n_0 + n_1 + \delta n) + \tau_{n0} (p_0 + p_1 + \delta n)}{n_0 + p_0 + \delta n}. \quad (2.5.113)$$

Они и в этом случае оказались равны потому, что мы пренебрегли перезарядкой ловушек. Но при относительно малом уровне инжекции становятся существенными и равновесные концентрации носителей. Если  $\delta n \ll n_0, p_0$ , то (2.5.113) упрощается до:

$$\tau_0 = \frac{\tau_{p0} (n_0 + n_1) + \tau_{n0} (p_0 + p_1)}{n_0 + p_0}. \quad (2.5.114)$$

Это выражение называется формулой Шокли — Рида — Холла. Из него видно, что при высоком уровне легирования, когда, на-

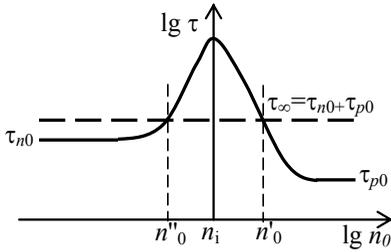


Рис. 2.5.14. Зависимость времени жизни носителей от концентрации электронов при малом (сплошная линия) и большом уровне биполярной инжекции.

малом уровне инжекции в собственном полупроводнике мы получим:

$$\tau_0 = \frac{\tau_{p0} + \tau_{n0}}{2} + \frac{\tau_{p0}n_1 + \tau_{n0}p_1}{2n_i} > \tau_{p0} + \tau_{n0} = \tau_\infty,$$

т.е. при  $n = n_i$  время жизни максимально из возможных для любого уровня инжекции. Следовательно, существуют такие равновесные концентрации, при которых время жизни от уровня инжекции не зависит. На рис. 2.5.14 они обозначены  $n'_0$  и  $n''_0$ .

пример,  $n_0 \gg p_0, n_1$ ,  $\tau_0 \rightarrow \tau_{p0}$ , время жизни определяется рекомбинацией неосновных носителей.

На рис. 2.5.14 показана зависимость времени жизни носителей при биполярной инжекции от концентрации электронов.

Из (2.5.114) видно, что в любом случае при

## 2.6. КИНЕТИЧЕСКАЯ ТЕОРИЯ БОЛЬЦМАНА

*Кинетическое уравнение Больцмана* – интегродифференциальное уравнение, которому удовлетворяют неравновесные одночастичные функции распределения системы из большого числа частиц, например, функции распределения электронов и фононов в кристалле и т. п. Кинетическое уравнение Больцмана – основное уравнение микроскопической теории неравновесных процессов (кинетики физической).

Кинетические уравнения Больцмана для квазичастиц в кристаллах, например, для электронов, называют также просто кинетическими уравнениями или уравнениями переноса.

Любой процесс переноса энергии или заряда – результат перемещения в пространстве отдельных частиц. Нас здесь будут интересовать электроны и фононы.

Суммарный эффект переноса – интеграл по вкладам всех частиц ансамбля, причем вклад каждой из них зависит от ее импульса. Плотность электронных состояний с импульсом в интервале от  $p$  до  $p + dp$  равна:

$$dN(p) = \frac{2}{(2\pi\hbar)^3} dp. \quad (2.6.1)$$

Количество электронов пропорционально вероятности заполнения этих состояний  $f(\mathbf{p}, \mathbf{r})$ , которая может зависеть и от координат. Скорость переноса пропорциональна скорости движения  $\mathbf{v}(p) = \nabla_p E(p)$ . Таким образом, плотность тока:

$$\mathbf{j} = -\frac{2e}{(2\pi\hbar)^3} \int \mathbf{v}(p) f(p, r) dp, \quad (2.6.2)$$

а поток переносимой электронами энергии:

$$\mathbf{I} = \frac{2}{(2\pi\hbar)^3} \int \mathbf{v}(p)[E(p) - e\varphi(r)]f(p, r)dp. \quad (2.6.3)$$

Здесь учтено, что энергия электрона равна  $E(p) - e\varphi(r)$ , она определяется не только внутрикристаллическими взаимодействиями, но и внешними полями.

В термодинамическом равновесии функция распределения  $f(\mathbf{p}, \mathbf{r})$  – функция Ферми. Она зависит только от энергии  $E(p)$ , четной функции импульса. Скорость – функция нечетная. Поэтому подынтегральные выражения в (2.6.2) и (2.6.3) – функции нечетные и интеграл по всему пространству равен нулю. Это естественно, в равновесии нет переноса. Но отсюда следует, что перенос возникнет только в случае, если  $f(\mathbf{p}, \mathbf{r})$  станет отлична от равновесной, станет нечетной функцией. Это может быть только при внешнем воздействии и для решения уравнений (2.6.2) и (2.6.3) нужно определить эту взаимосвязь неравновесной функции  $f(\mathbf{p}, \mathbf{r})$  и внешних воздействий.

Кинетическое уравнение Больцмана – уравнение, определяющее эту взаимосвязь.

### 2.6.1. Кинетическое уравнение

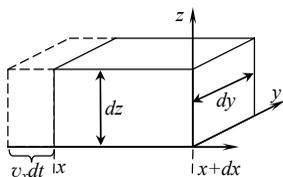


Рис. 2.6.1. Элемент пространства.

Интересующая нас неравновесная функция распределения определена в шестимерном пространстве, – координат и импульсов. Кроме того, она, очевидно не постоянна во времени.

Рассмотрим элемент пространства  $dr = dx dy dz$ , изображенный на рис. 2.6.1. Таким же образом можно выделить элемент простран-

ства импульсов  $dp = dp_x dp_y dp_z$ . Количество электронов (с одинаковым спином) в элементе пространства  $dp \cdot dr$  в момент времени  $t$  равно:

$$dn = f(p, r, t) \frac{dp \cdot dr}{(2\pi\hbar)^3}. \quad (2.6.4)$$

В следующий момент времени оно будет уже иным:

$$dn' = f(p, r, t + dt) \frac{dp \cdot dr}{(2\pi\hbar)^3} = \left[ f(p, r, t) + \frac{\partial f(p, r, t)}{\partial t} dt \right] \frac{dp \cdot dr}{(2\pi\hbar)^3},$$

так что изменение составит:

$$dn' - dn = \frac{\partial f}{\partial t} \frac{dp \cdot dr \cdot dt}{(2\pi\hbar)^3}. \quad (2.6.5)$$

Вычислив эту разность, мы сможем найти и  $f(p, r, t)$ . Для этого нужно рассмотреть все возможные причины изменения  $f(p, r, t)$ . Это:

- смещение их в пространстве, транспорт;
- смещение их в пространстве импульсов под действием внешней силы;
- рассеяние в результате столкновений с дефектами или друг с другом;
- генерация или рекомбинация, изменение плотности частиц.

Но последним фактором мы пренебрежем, остановимся только на стационарных состояниях. Итого, рассматриваем три основные причины изменения функции распределения.

### ***Транспорт.***

За счет движения в координатном пространстве за время  $dt$  какое-то количество частиц попадет в элемент объема, изображенный на рис. 2.6.1, какое-то из него выйдет. Рассмотрим дви-

жение электронов с фиксированным импульсом по одной из координат,  $x$ . Концентрации частиц в точках  $x$  и  $x+dx$  могут быть различны, но скорости одинаковы (мы сейчас не рассматриваем транспорт в пространстве импульсов). Изменение числа частиц в объеме  $dr$  за счет движения по оси  $x$  составит:

$$\begin{aligned}
 & f(p, x, y, z, t) v_x dt \frac{dp \cdot dy dz}{(2\pi\hbar)^3} - \\
 & - f(p, x + dx, y, z, t) v_x dt \frac{dp \cdot dy dz}{(2\pi\hbar)^3} = \quad (2.6.6) \\
 & = - \frac{\partial f}{\partial x} dx \cdot v_x dt \frac{dp \cdot dy dz}{(2\pi\hbar)^3} = - \frac{\partial f}{\partial x} v_x dt \frac{dp \cdot dr}{(2\pi\hbar)^3}.
 \end{aligned}$$

По двум другим координатам мы получим, очевидно, такие же выражения, так что полное изменение за счет движения равно сумме трех выражений типа (2.6.6) и составит:

$$(dn' - dn)_{\text{трансл}} = -(v, \nabla f) \frac{dp \cdot dr \cdot dt}{(2\pi\hbar)^3}. \quad (2.6.7)$$

### *Действие внешних сил.*

Движение в пространстве импульсов можно рассмотреть точно так же. Скорость изменения импульса во времени равна  $\frac{d\mathbf{p}}{dt} = \mathbf{F}$ , градиент функции распределения в данном случае тоже нужно брать в пространстве импульсов, итого, получим:

$$\begin{aligned}
 (dn' - dn)_{\text{сил}} & = - \left( \frac{d\mathbf{p}}{dt}, \nabla_p f \right) \frac{dp \cdot dr \cdot dt}{(2\pi\hbar)^3} = \\
 & = - (\mathbf{F}, \nabla_p f) \frac{dp \cdot dr \cdot dt}{(2\pi\hbar)^3}. \quad (2.6.8)
 \end{aligned}$$

### *Столкновения, рассеяние импульса.*

При столкновениях координаты частиц не изменяются. Они остаются в том же элементе объема, но мгновенно изменяется их импульс.

Рассмотрим сначала рассеяние на дефектах. В этом случае нас интересует только один из участников столкновения, электрон. Будем обозначать как  $\mathbf{p}$  состояния, соответствующие "нашему" элементу фазового пространства,  $\mathbf{p}'$  – произвольные иные.

В результате рассеяния возможен "приход" частиц, т.е. процесс  $\mathbf{p}' \rightarrow \mathbf{p}$ , но возможен и "уход", переход  $\mathbf{p} \rightarrow \mathbf{p}'$ . Вероятность того, что электрон попадет в "наш" элемент пространства, пропорциональна количеству электронов в состоянии  $\mathbf{p}'$  и концентрации свободных мест в состоянии  $\mathbf{p}$ , т.е. равна:

$$\mathcal{P}_1(\mathbf{p}', \mathbf{p}) f(\mathbf{p}', \mathbf{r}) [1 - f(\mathbf{p}, \mathbf{r})] \frac{d\mathbf{p} \cdot d\mathbf{p}' \cdot d\mathbf{r}}{(2\pi\hbar)^6} \quad (2.6.9)$$

Здесь  $\mathcal{P}_1(\mathbf{p}', \mathbf{p})$  – некоторый не известный нам пока коэффициент пропорциональности. В принципе, он может зависеть и от координат, и от внешних полей, но мы будем рассматривать пространственно однородные структуры. Зависимость же от полей может возникнуть лишь в том случае, если по величине они сравнимы с внутрикристаллическими. Это – очень сильные поля, мы таких рассматривать не будем.

Полная вероятность того, что электрон придет в "наш" объем, равна интегралу от (2.6.9) по всем "внешним" состояниям  $\mathbf{p}'$ :

$$\begin{aligned} \frac{d\mathbf{p} \cdot d\mathbf{r} \cdot dt}{(2\pi\hbar)^3} \int \mathcal{P}_1(\mathbf{p}', \mathbf{p}) f(\mathbf{p}', \mathbf{r}) [1 - f(\mathbf{p}, \mathbf{r})] \frac{d\mathbf{p}'}{(2\pi\hbar)^3} = \\ = A \frac{d\mathbf{p} \cdot d\mathbf{r} \cdot dt}{(2\pi\hbar)^3}. \end{aligned} \quad (2.6.10)$$

Это – так называемый *интеграл прихода*. Полностью аналогично можно найти и *интеграл ухода*:

$$\begin{aligned} \frac{d\mathbf{p} \cdot d\mathbf{r} \cdot dt}{(2\pi\hbar)^3} \int \mathcal{P}_2(\mathbf{p}, \mathbf{p}') f(\mathbf{p}, \mathbf{r}) [1 - f(\mathbf{p}', \mathbf{r})] \frac{d\mathbf{p}'}{(2\pi\hbar)^3} = \\ = B \frac{d\mathbf{p} \cdot d\mathbf{r} \cdot dt}{(2\pi\hbar)^3}, \end{aligned} \quad (2.6.11)$$

где  $\mathcal{P}_2(\mathbf{p}, \mathbf{p}')$ , вообще говоря, может отличаться от  $\mathcal{P}_1(\mathbf{p}', \mathbf{p})$ .

Итого, в сумме изменение числа частиц в "нашем" объеме за счет столкновений с решеткой можно записать как:

$$(dn' - dn)_{\text{столкн}} = J[f] \frac{dp \cdot dr \cdot dt}{(2\pi\hbar)^3}, \quad (2.6.12)$$

где величина

$$\begin{aligned} J[f] = A - B = \frac{1}{(2\pi\hbar)^3} \int \{ \mathcal{P}_1(\mathbf{p}', \mathbf{p}) f(\mathbf{p}', \mathbf{r}) [1 - f(\mathbf{p}, \mathbf{r})] - \\ - \mathcal{P}_2(\mathbf{p}, \mathbf{p}') f(\mathbf{p}, \mathbf{r}) [1 - f(\mathbf{p}', \mathbf{r})] \} d\mathbf{p}' \end{aligned} \quad (2.6.13)$$

– называется *интеграл столкновений*.

Подобным способом можно рассмотреть и электрон-электронные столкновения. В этом случае для вероятности прихода нужно рассмотреть пару электронов в состояниях  $\mathbf{p}'_1$  и  $\mathbf{p}'_2$ , которые при столкновении перейдут в состояния  $\mathbf{p}_1$  и  $\mathbf{p}_2$ . Нас интересует состояние, которое мы обозначили как  $\mathbf{p}_1$ , так что полная вероятность прихода или ухода – тройной интеграл по  $\mathbf{p}_2$ ,  $\mathbf{p}'_1$  и  $\mathbf{p}'_2$ . Для интеграла столкновений получим тогда:

$$\begin{aligned}
J[f] = & \int \frac{d\mathbf{p}'_1 d\mathbf{p}'_2 d\mathbf{p}_2}{(2\pi\hbar)^9} \times \\
& \times \{ \mathcal{P}_1(\mathbf{p}'_1, \mathbf{p}'_2, \mathbf{p}_1, \mathbf{p}_2) f(\mathbf{p}'_1, \mathbf{r}) f(\mathbf{p}'_2, \mathbf{r}) [1 - f(\mathbf{p}_1, \mathbf{r})] \cdot [1 - f(\mathbf{p}_2, \mathbf{r})] - \\
& - \mathcal{P}_2(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}'_1, \mathbf{p}'_2) f(\mathbf{p}_1, \mathbf{r}) f(\mathbf{p}_2, \mathbf{r}) [1 - f(\mathbf{p}'_1, \mathbf{r})] \cdot [1 - f(\mathbf{p}'_2, \mathbf{r})] \}.
\end{aligned} \tag{2.6.14}$$

**Итого**, просуммировав все каналы изменения функции распределения, из (2.6.5) получим:

$$\frac{\partial f}{\partial t} = -(\mathbf{v}, \nabla f) - (\mathbf{F}, \nabla_p f) + J[f]. \tag{2.6.15}$$

Это интегродифференциальное уравнение может быть решено при заданных начальных условиях, но прежде нужно вычислить функции  $\mathcal{P}_1$  и  $\mathcal{P}_2$ , для чего нужно рассмотреть механизмы рассеяния.

## 2.6.2. Состояние термодинамического равновесия.

### Принцип детального равновесия.

В равновесии:

$$\frac{\partial f}{\partial t} = 0 = -(\mathbf{v}, \nabla f_0) - (\mathbf{F}, \nabla_p f_0) + J[f_0], \tag{2.6.16}$$

где равновесная функция распределения

$$f_0 = \frac{1}{\exp \frac{E(p) - e\varphi - F}{kT} - 1}. \tag{2.6.17}$$

Подставим в (2.6.16) производные от нее:

$$\nabla f_0 = -e f_0' \nabla \varphi, \quad \nabla_p f_0 = (\nabla_p E) f_0' = v f_0'. \tag{2.6.18}$$

Здесь мы ввели новое обозначение, производную по энергии от равновесной функции распределения:

$$f_0' \equiv \frac{df_0}{dE} \quad (2.6.19)$$

откуда следует, что при  $\mathbf{F} = -e\mathbf{E} - \frac{e}{c}[\mathbf{v}\mathbf{B}]$  и  $\mathbf{E} = -\nabla\varphi$

$$(\mathbf{F}, \nabla_p f_0) = ef_0' \left\{ (\mathbf{v}, \nabla\varphi) - \frac{1}{c}(\mathbf{v}, [\mathbf{v}\mathbf{B}]) \right\} = ef_0'(\mathbf{v}, \nabla\varphi), \quad (2.6.20)$$

получим:

$$ef_0'(v, \nabla\varphi) - ef_0'(v, \nabla\varphi) + J[f] = 0 \Rightarrow \underline{J[f] = 0} ! \quad (2.6.21)$$

Интеграл столкновений в равновесии равен нулю. Для столкновений с решеткой это означает:

$$\int \{ \mathcal{P}_1(\mathbf{p}', \mathbf{p}) f_0(\mathbf{p}', \mathbf{r}) [1 - f_0(\mathbf{p}, \mathbf{r})] - \mathcal{P}_2(\mathbf{p}, \mathbf{p}') f_0(\mathbf{p}, \mathbf{r}) [1 - f_0(\mathbf{p}', \mathbf{r})] \} d\mathbf{p}' = 0. \quad (2.6.22)$$

И это равенство должно быть справедливо для всех  $\mathbf{p}$ . Поэтому разумно предположить, что интеграл (2.6.22) равен нулю только в случае, если равно нулю подынтегральное выражение.

Приняв это и подставив в (2.6.22) выражение для  $f_0$ , получим:

$$\frac{\mathcal{P}_1(\mathbf{p}', \mathbf{p})}{\mathcal{P}_2(\mathbf{p}, \mathbf{p}')} = \exp \frac{E(\mathbf{p}') - E(\mathbf{p})}{kT}. \quad (2.6.23)$$

Это равенство, не доказанное строго, и называется *принцип детального равновесия*.

Из него следует, в частности, что при упругом рассеянии, без изменения энергии,  $\mathcal{P}_1(\mathbf{p}', \mathbf{p}) = \mathcal{P}_2(\mathbf{p}, \mathbf{p}')$ , потоки частиц из состояния  $\mathbf{p}$  в  $\mathbf{p}'$  и из  $\mathbf{p}'$  в  $\mathbf{p}$  равны. При неупругом рассеянии нижнее состояние в  $\exp \frac{\Delta E}{kT}$  раз больше заселено, но вероятность их

возбуждения во столько же раз меньше. Потоки в оба направления и в этом случае равны.

Для электрон-электронного рассеяния мы получили бы то же самое, но в этом случае состояния системы, начальное и конечное, определяются уже не двумя, а четырьмя точками в  $\mathbf{p}$ -пространстве и, соответственно, четырьмя энергиями. Принцип детального равновесия будет по сути, тем же самым, а его запись несколько длиннее:

$$\frac{\mathcal{P}_1(\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}'_1, \mathbf{p}'_2)}{\mathcal{P}_2(\mathbf{p}, \mathbf{p}')} = \exp \frac{E(\mathbf{p}'_1) + E(\mathbf{p}'_2) - E(\mathbf{p}_1) - E(\mathbf{p}_2)}{kT}. \quad (2.6.24)$$

### 2.6.3. Малые отклонения от равновесия

Определившись с равновесным состоянием системы, рассмотрим отклонения от равновесия. Но для того, чтобы облегчить задачу, примем, что отклонение не велико. В неравновесном состоянии функцию распределения запишем как сумму равновесной и неравновесной:

$$f(\mathbf{p}, \mathbf{r}) = f_0(\mathbf{p}, \mathbf{r}) + f_1(\mathbf{p}, \mathbf{r}), \quad (2.6.25)$$

где

$$f_0 = \frac{1}{\exp \frac{E(p) - e\varphi(r) - F(r)}{kT(r)} - 1}. \quad (2.6.26)$$

Во всяком случае, было бы более справедливо вместо (2.6.25) записать разложение  $f(\mathbf{p}, \mathbf{r})$  в ряд близ точки  $f_0(\mathbf{p}, \mathbf{r})$ . Но учет высоких членов разложения затруднителен, а заключения, сделанные в пренебрежении производными второго и более высоких порядков, могут быть справедливы только в случае

$$|f_1(p, r)| \ll f_0(p, r). \quad (2.6.27)$$

Малыми отклонениями от равновесия будем называть ситуации, в которых (2.6.27) справедливо.

Запишем кинетическое уравнение в виде:

$$\frac{\partial f}{\partial t} + (\mathbf{v}, \nabla f) - (e\mathbf{E}, \nabla_p f) - \frac{e}{c}([\mathbf{v}\mathbf{B}], \nabla_p f) = J[f], \quad (2.6.28)$$

разделив силовой член на действие электрического и магнитного полей.

В общем случае в пространстве не постоянны ни потенциал, ни температура, ни энергия Ферми. Потому:

$$\nabla f_0 = \left( -e\nabla\varphi - \nabla F - \frac{E(p) - e\varphi - F}{kT} \nabla(kT) \right) f_0'. \quad (2.6.29)$$

Градиент  $f_0(p, r)$  по импульсу всегда есть:

$$\nabla_p f_0 = \mathbf{v} f_0'. \quad (2.6.30)$$

Тогда во втором и третьем членах (2.6.28):

$$\begin{aligned} & (\mathbf{v}, \nabla f_0) - (e\mathbf{E}, \nabla_p f_0) = \\ & = \left\{ -e(\mathbf{v}, \nabla\varphi) - (\mathbf{v}, \nabla F) - \frac{E(p) - e\varphi - F}{T} (\mathbf{v}, \nabla T) - e(\mathbf{v}, \mathbf{E}) \right\} f_0' \end{aligned}$$

Это равенство несколько упрощается, поскольку  $\nabla\varphi = -\mathbf{E}$ :

$$(\mathbf{v}, \nabla f_0) - (e\mathbf{E}, \nabla_p f_0) = \left\{ -(\mathbf{v}, \nabla F) - \frac{E(p) - e\varphi - F}{T} (\mathbf{v}, \nabla T) \right\} f_0' \quad (2.6.31)$$

Электрическое поле ушло из кинетического уравнения, его роль играет величина  $\mathbf{E}' = \frac{1}{e} \nabla F$ .

При условии (2.6.27) в этих членах можно пренебречь неравновесной частью функции распределения. Все равно  $f_0(p, r)$  дает основной вклад.

Но последний член левой части (2.6.28) в равновесии равен нулю, т.к. векторно-скалярное произведение с двумя одинаковыми векторами равно нулю, а  $\nabla_p f_0$  пропорционален скорости, см. (2.6.30).

$$\frac{e}{c}([\mathbf{v}\mathcal{B}], \nabla_p f) = \frac{e}{c} \left\{ ([\mathbf{v}\mathcal{B}], \mathbf{v}f'_0) + ([\mathbf{v}\mathcal{B}], \nabla_p f_1) \right\} = \frac{e}{c} ([\mathbf{v}\mathcal{B}], \nabla_p f_1).$$

Значит, здесь остается только неравновесная часть.

Производная по времени от равновесной части функции распределения тоже равна нулю. Итого, кинетическое уравнение при малом отклонении от равновесия можно записать в виде:

$$\begin{aligned} \frac{\partial f_1}{\partial t} - e(\mathbf{v}, e\mathcal{E}')f'_0 - \frac{E(p) - e\varphi - F}{T}(\mathbf{v}, \nabla T)f'_0 - \\ - \frac{e}{c}([\mathbf{v}\mathcal{B}], \nabla_p f_1) = J[f]. \end{aligned} \quad (2.6.32)$$

## 2.6.4. Кинетические эффекты

Теперь нужно уточнить общий вид неравновесной части функции распределения. Для этого рассмотрим несколько конкретных типов внешних воздействий. При этом нужно помнить, что все поля – векторы ( $\mathcal{B}$  – псевдовектор), а  $f_1$  – скаляр.

### 2.6.4.1. Носители в слабом постоянном и однородном электрическом поле.

Пусть магнитное поле отсутствует и температура всюду постоянна. Наиболее общее линейное соотношение, связывающее векторы  $j$  и  $\mathcal{E}$ , имеет вид обобщенного закона Ома (см. раздел 2.3.1):

$$\begin{aligned}
 j_x &= \sigma_{xx} \mathcal{E}_x + \sigma_{xy} \mathcal{E}_y + \sigma_{xz} \mathcal{E}_z \\
 j_y &= \sigma_{yx} \mathcal{E}_x + \sigma_{yy} \mathcal{E}_y + \sigma_{yz} \mathcal{E}_z \\
 j_z &= \sigma_{zx} \mathcal{E}_x + \sigma_{zy} \mathcal{E}_y + \sigma_{zz} \mathcal{E}_z
 \end{aligned}
 \tag{2.6.33}$$

или, сокращенно,  $j_\alpha = \sigma_{\alpha\beta} \mathcal{E}_\beta$ ,  $\alpha, \beta = x, y, z$ .

В отсутствие магнитного поля тензор электропроводности оказывается симметричным,  $\sigma_{\alpha\beta} = \sigma_{\beta\alpha}$ . При этом выбором системы координат можно диагонализировать матрицу проводимости, так что  $\sigma_{\alpha\beta} = 0$  при  $\alpha \neq \beta$ .

В кубических кристаллах к тому же  $\sigma_{xx} = \sigma_{yy} = \sigma_{zz} = \sigma$ . Мы остановимся на этом приближении,  $j = \sigma \mathcal{E}$ .

В неоднородных кристаллах и при наличии иных полей, например, температурных, существенно возникновение полей сторонних сил и тогда нужно учитывать полное поле,

$$\mathcal{E}' = \mathcal{E} + \mathcal{E}^* = -\nabla\phi + \mathcal{E}^* = \frac{1}{e} \nabla F.
 \tag{2.6.34}$$

Поскольку мы рассматриваем токи в отсутствие иных полей,  $\nabla T, \mathcal{B} = 0$ , то в кинетическом уравнении (2.6.32) останется :

$$\frac{\partial f_1}{\partial t} = J[f] + e(\mathbf{v}, e\mathcal{E}') f_0'.
 \tag{2.6.35}$$

Неравновесная функция распределения зависит только от импульса и напряженности поля, причем – от скалярной функции этих векторов, их скалярного произведения и квадрата импульса или энергии. Квадрат электрического поля будем считать пренебрежимо малым, поскольку мы рассматриваем малые отклонения от равновесия. Итого,  $f_1$  должна иметь вид:

$$f_1 = (\mathbf{p}, \mathcal{E}) \psi(E).
 \tag{2.6.36}$$

### 2.6.4.2. Носители заряда в слабом постоянном и однородном температурном поле.

При наличии градиента температуры появляется физически выделенное направление. Это приводит к возникновению целой группы термоэлектрических явлений. В однородных полупроводниках важнейшие из них – *термоэлектродвижущая сила (термоэдс) и эффекты Томсона и Пельтье*.

Первый из них заключается в том, что между концами разомкнутого проводника, которые имеют различную температуру, возникает разность потенциалов, а, значит, внутри проводника возникает электродвижущая сила.

**Возникновение термоэдс** связано с тем, что концентрация носителей быстро растет с температурой, поток диффузии носителей от нагретого конца к холодному больше, чем в обратном направлении. Поэтому на концах проводника (и на его поверхности) появляются электрические заряды, а внутри проводника — электрическое поле. В стационарном состоянии разомкнутого проводника это поле таково, что вызываемый им ток дрейфа как раз компенсирует результирующий поток диффузии.

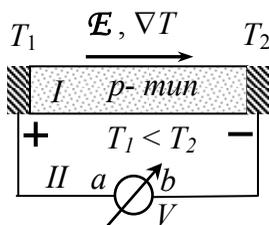


Рис. 2.6.2. К объяснению термоэдс и эффекта Томсона.

Однако полная величина напряжения, регистрируемого вольтметром между металлическими электродами *a* и *b* (рис. 1.3), складывается не только из падения напряжения внутри проводника. На границах полупроводник/металл возникает дополнительная разность потенциалов (контактная разность потенциалов), которая зависит от температуры. Если оба металлических электрода одинаковы и разности температур нет, то сумма обоих

скачков потенциала равна нулю. Если же температура концов различна, то и величина обоих скачков потенциала уже неодинакова. Это вносит дополнительный вклад в термоэдс.

Формально описывать термоэдс принято величиной дифференциальной термоэдс:

$$\alpha = \frac{dV_0}{dT}. \quad (2.6.37)$$

Таким образом, для плотности тока при наличии градиента температуры следует написать сумму двух членов:

$$\mathbf{j} = a \cdot \nabla T + \sigma \mathbf{E}'. \quad (2.6.38)$$

При разомкнутых концах  $j = 0 \Rightarrow \sigma \mathbf{E}' = -\frac{a}{\sigma} \nabla T$ . Следовательно, разность потенциалов на концах разомкнутого полупроводника

$$V_0 = -\int \frac{a}{\sigma} (\nabla T \cdot d\mathbf{l}). \quad (2.6.39)$$

Если градиент температуры направлен по длине образца, то:

$$V_0 = -\int_{T_1}^{T_2} \frac{a}{\sigma} dT = -\frac{a}{\sigma} (T_2 - T_1). \quad (2.6.40)$$

Но, согласно (2.6.37),  $V_0 = \alpha(T_2 - T_1)$ , т.е.,

$$\alpha = -\frac{a}{\sigma} = \frac{1}{e} \frac{|\nabla F|}{|\nabla T|}. \quad (2.6.41)$$

И тогда

$$\mathbf{j} = \sigma \mathbf{E}' - \alpha \sigma \cdot \nabla T = \sigma (\mathbf{E}' - \alpha \cdot \nabla T). \quad (2.6.42)$$

Отсюда можно определить величину дифференциальной термоэдс. Вычислив плотность тока, мы получим явное выражение коэффициента у  $\nabla T$ .

### 2.6.4.3. Выделение тепла. Эффект Томсона

**Эффект Томсона** – выделение дополнительной энергии, кроме джоулева тепла, выделяемой в объеме полупроводника при наличии электрического тока и градиента температуры.

$$\Delta Q = -\alpha_T j \nabla T .$$

Здесь  $\alpha_T$  – **коэффициент Томсона**. Поскольку это тепло зависит от первой степени тока, то при изменении направления тока меняет знак и  $\Delta Q$ , т.е. тепло может либо выделяться, либо поглощаться.

Наряду с этим существует и обычная теплопроводность. Поток тепла за счет теплопроводности равен:

$$I_{\text{теплопров}} = -K \frac{dT}{dx} ,$$

где  $K$  – коэффициент теплопроводности.

Если этот поток не постоянен в пространстве, то это означает выделение (поглощение) тепла. Итого, плотность выделяемого тепла при наличии тока равна:

$$Q_V = \frac{j^2}{\sigma} - \alpha_T (\mathbf{j} \cdot \nabla T) + \text{div}(K \nabla T) . \quad (2.6.43)$$

Нас интересует поток тепла, переносимого электронами, так что в этом выражении под  $K$  нужно понимать электронный вклад в теплопроводность.

#### 2.6.4.4. Выделение тепла на контакте. Эффект Пельтье

Обратимое выделение тепла наблюдается и на границе, на контакте двух различных проводников. Количество тепла, выделяемое на единичной площади контакта в единицу времени,  $Q_S$ , пропорционально проходящему току,

$$Q_S = \Pi_{12} j. \quad (2.6.44)$$

Здесь  $\Pi_{12}$  – коэффициент Пельтье, зависящий от свойств контактирующих тел. При изменении направления тока изменяется и знак  $Q_S$ . При одном направлении тепло выделяется, при другом – поглощается, так что  $\Pi_{12} = -\Pi_{21}$ .

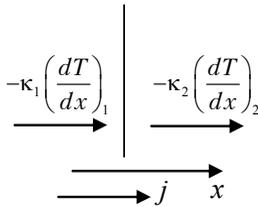


Рис. 2.6.3. К выводу формулы эффекта Пельтье.

Рассмотрим контакт двух различных тел, рис. 2.6.3, через который протекает некоторый ток. Очевидно, в стационарном состоянии потоки тепла по обе стороны контакта равны,  $I_1 = I_2$ . Они складываются из теплопроводности, пропорциональной градиенту температуры, и энергии, переносимой электронами. В первом приближении она определяется энергией

Ферми, но при таком обмене электронами переносимая энергия определяется не средней энергией электронного ансамбля, а средней энергией электронов, участвующих в переносе. Если контактируют полупроводники, то в идеальном случае это – максвелловское распределение. Но в реальных случаях оно может быть иным и из-за отклонений закона дисперсии от параболического, и из-за принципа Паули, отклонений статистики Ферми от больцмановского распределения.

В общем, энергия, переносимая электронами, несколько отличается от энергии Ферми, так что баланс потоков на границе можно записать как:

$$-K_1 \left( \frac{dT}{dx} \right)_1 - j_1 \left( \frac{F_1}{e} - \Pi_1 \right) = -K_2 \left( \frac{dT}{dx} \right)_2 - j_2 \left( \frac{F_2}{e} - \Pi_2 \right), \quad (2.6.45)$$

где  $\Pi_1, \Pi_2$  – не известные пока скалярные параметры.

Поскольку и токи, и уровни Ферми с обеих сторон равны,  $j_1 = j_2 = j$ ,  $F_1 = F_2$ , то:

$$K_2 \left( \frac{dT}{dx} \right)_2 - K_1 \left( \frac{dT}{dx} \right)_1 = j(\Pi_2 - \Pi_1) = j\Pi_{21}. \quad (2.6.46)$$

При учете этой дополнительной энергии поток энергии равен:

$$\mathbf{I} = -K \cdot \nabla T + \Pi \mathbf{j} - \frac{F}{e} \mathbf{j}, \quad (2.6.47)$$

А количество выделяемого тепла:

$$Q_V = -\text{div} \mathbf{I} = \text{div}(K \nabla T) - (\mathbf{j}, \nabla \Pi) + (\mathbf{j}, \mathbf{E}'). \quad (2.6.48)$$

Из (2.6.42) имеем:  $\mathbf{E}' = \frac{j}{\sigma} + \alpha \cdot \nabla T$ , так что (2.6.48) можно переписать в виде:

$$Q_V = -\text{div} \mathbf{I} = \text{div}(\kappa \nabla T) + \frac{j^2}{\sigma} - \left( \frac{d\Pi}{dT} - \alpha \right) (\mathbf{j}, \nabla \Pi). \quad (2.6.49)$$

Сравнивая это с (2.6.43), получим взаимосвязь дифференциальной термоэдс и коэффициентов Томсона и Пельтье:

$$\alpha_T = \frac{d\Pi}{dT} - \alpha. \quad (2.6.50)$$

Эти коэффициенты предстоит вычислить с помощью кинетического уравнения Больцмана.

#### 2.6.4.5. Полупроводник в электрическом и магнитном полях

При наличии обоих полей возникают три независимых вектора:

$$\mathbf{E}, [\mathbf{E}\mathbf{B}], \mathbf{B}(\mathbf{B}\mathbf{E}).$$

Поэтому плотность тока в общем случае – сумма трех компонент:

$$\mathbf{j} = a_1\mathbf{E} + a_2[\mathbf{E}\mathbf{B}] + a_3\mathbf{B}(\mathbf{B}\mathbf{E}), \quad (2.6.51)$$

где  $a_1, a_2, a_3$  – скалярные коэффициенты.

В то же время всегда справедлива запись:

$$j_\alpha = \sigma_{\alpha\beta}(\mathbf{B})E_\beta. \quad (2.6.52)$$

Если магнитное поле направить по оси  $z$ , то:

$$\begin{aligned} \sigma_{xx} = \sigma_{yy} = a_1, & \quad \sigma_{xy} = -\sigma_{yx} = a_2\mathcal{B}, \\ \sigma_{zz} = a_1 + a_3\mathcal{B}^2, & \quad \sigma_{xz} = \sigma_{zx} = \sigma_{yz} = \sigma_{zy} = 0. \end{aligned} \quad (2.6.53)$$

Для постоянной Холла и поперечного магнитосопротивления (см. (2.6.33) и (2.6.37) получим тогда:

$$R = \frac{a_2}{a_1^2 + a_2^2\mathcal{B}^2}, \quad (2.6.54)$$

$$\frac{\rho_\perp - \rho_0}{\rho_0} = \frac{a_1(1 - \rho_0 a_1) - \rho_0 a_2^2 \mathcal{B}^2}{\rho_0 (a_1^2 + a_2^2 \mathcal{B}^2)}. \quad (2.6.55)$$

Следовательно, рассчитав плотность тока для этого случая, т.е. получив явные выражения для коэффициентов (2.6.51), мы

рассчитаем и постоянную Холла, и поперечное магнитосопротивление.

### 2.6.5. *Общая форма неравновесной функции распределения*

Для всех рассмотренных эффектов мы получили, что суммарный эффект, поток заряда или энергии, определяется векторными внешними воздействиями. В то же время, величины этих потоков определяются скалярной величиной – неравновесной функцией распределения, см. (2.6.2) и (2.6.3), для которой получено выражение, связывающее ее с воздействующими факторами, (2.6.32). Для этого нужно уточнить общий вид неравновесной части функции распределения.

Очевидно, что она зависит от импульса частицы, ее энергии (квадрата импульса) и внешнего воздействия. Логично поэтому искать неравновесную функцию распределения в виде скалярного произведения импульса на вектор воздействия, умноженного на некоторый скалярный коэффициент, зависящий от энергии.

Для электропроводности она будет тогда иметь вид:

$$f_1 = (\mathbf{p}, \mathbf{E})\psi(E). \quad (2.6.56)$$

При наличии температурного поля имеются два вектора:

$$f_1 = (\mathbf{p}, \nabla T)\chi_1(E) + (\mathbf{p}, \mathbf{E}')\chi_2(E). \quad (2.6.57)$$

В электрическом и магнитном полях – три:

$$f_1 = (\mathbf{p}, \mathbf{E})\psi_1(E) + (\mathbf{p}, [\mathbf{E}\mathbf{B}])\psi_2(E) + (\mathbf{p}, \mathbf{B})(\mathbf{B}\mathbf{E})\psi_3(E). \quad (2.6.58)$$

Множители, зависящие от энергии, предстоят определению.

Но можно заметить, что все эти три функции – скалярное произведение импульса на некоторый вектор воздействия. Поэтому их можно записать в общем виде:

$$f_1 = (\mathbf{p}, \xi(E)). \quad (2.6.59)$$

Легко видеть, что это – нечетная функция импульса. Поэтому прибавление ее к равновесной функции,  $f = f_0 + f_1$ , не изменит средней энергии носителей:

$$\langle E \rangle = \frac{2e}{(2\pi\hbar)^3} \int (f_0 + f_1) \cdot E(p) dp.$$

Энергия – четная функция импульса, произведение  $f_1 \cdot E(p)$  – нечетная, а интеграл здесь берется по симметричному промежутку.

Напротив, в интегралах (2.6.2) и (2.6.3) функция распределения умножается на нечетную функцию – скорость. Поэтому они могут быть отличны от нуля только в случае, если к равновесной функции распределения добавится нечетная часть.

### **2.6.6. Интеграл столкновений для упругого рассеяния в изотропном кристалле**

При решении этого уравнения наибольшую сложность составляет интеграл столкновений. В общем случае нахождение его – задача крайне сложная, так что приходится строить упрощающие модели. Одна из них – пренебрежение изменением энергии при столкновении.

Этим изменением действительно можно пренебречь при рассеянии носителей на атомах и заряженных или нейтральных примесях, дислокациях или иных структурных дефектах. Все эти дефекты обладают значительно большей массой, нежели электрон или дырка. При столкновении легкой частицы с тяжелой может сильно измениться импульс каждой из них, но обмен энергией весьма затруднен: изменение энергии легкой частицы при столкновении оказывается малым по сравнению с самой этой энергией. Можно показать (см. [15], глава XIV), что так же обстоит дело и

при рассеянии электронов на акустических колебаниях решетки. Изменение энергии при рассеянии пропорционально  $\sqrt{m/M}$ , где  $m$  — эффективная масса электрона, а  $M$  — масса атома решетки.

Итак, пренебрежем обменом энергией при соударениях. Все соударения будем считать упругими. Это серьезно упростит нашу задачу, но нужно помнить, что при этом мы ограничиваем и область применимости ее результатов. В частности, отсюда нельзя будет рассчитать выделение джоулева тепла и вообще, вероятно, возникнут ошибки во всех случаях, когда энергия, приобретаемая в поле за время свободного пробега сравнима с равновесной энергией.

Для того, чтобы в (2.6.13) и (2.6.32) вписать явно упругость рассеяния, вероятности рассеяния запишем как:

$$\begin{aligned} \mathcal{P}_1(\mathbf{p}', \mathbf{p}) &= \delta[E(\mathbf{p}') - E(\mathbf{p})] \cdot S(\mathbf{p}', \mathbf{p}), \\ \mathcal{P}_2(\mathbf{p}, \mathbf{p}') &= \delta[E(\mathbf{p}') - E(\mathbf{p})] \cdot S(\mathbf{p}, \mathbf{p}'). \end{aligned} \quad (2.6.60)$$

Согласно принципу детального равновесия,

$$\mathcal{P}_1(\mathbf{p}', \mathbf{p}) = \mathcal{P}_2(\mathbf{p}, \mathbf{p}') \text{ и } S(\mathbf{p}', \mathbf{p}) = S(\mathbf{p}, \mathbf{p}'), \quad (2.6.61)$$

так что интеграл столкновений (2.6.13) (учитываем только столкновения с решеткой) несколько упростится:

$$J[f] = \frac{1}{(2\pi\hbar)^3} \int \delta[E(\mathbf{p}') - E(\mathbf{p})] \cdot S(\mathbf{p}, \mathbf{p}') [f(\mathbf{p}', \mathbf{r}) - f(\mathbf{p}, \mathbf{r})] d\mathbf{p}'.$$

В изотропном кристалле сохранение энергии означает и сохранение импульса (по модулю). Не нужно забывать также, что  $S(\mathbf{p}, \mathbf{p}')$  — скаляр, он может зависеть только от скалярных комбинаций векторов  $\mathbf{p}$  и  $\mathbf{p}'$ . Квадраты импульсов равны и пропорциональны энергии частицы, скалярное произведение  $\mathbf{p}$  и  $\mathbf{p}'$  пропорционально  $\cos\theta$ . Энергия и угол рассеяния — величины

независимые. Поэтому вероятность рассеяния может зависеть только от величины энергии и от угла рассеяния  $\theta$ :

$$S(\mathbf{p}, \mathbf{p}') = S(E, \cos \theta).$$

В дальнейшем будем рассматривать интеграл столкновений в форме:

$$J[f] = \frac{1}{(2\pi\hbar)^3} \int \delta[E(\mathbf{p}') - E(\mathbf{p})] \cdot S(E(\mathbf{p}), \cos \theta) \times \quad (2.6.62) \\ \times [f(\mathbf{p}', \mathbf{r}) - f(\mathbf{p}, \mathbf{r})] d\mathbf{p}'.$$

Подставим сюда  $f = f_0 + f_1$ , учтем, что при упругом рассеянии  $f(\mathbf{p}', \mathbf{r}) = f_0(\mathbf{p}, \mathbf{r})$ , а неравновесную часть запишем в виде (2.6.59):

$$J[f] = \frac{1}{(2\pi\hbar)^3} \int \delta[E(\mathbf{p}') - E(\mathbf{p})] \cdot S(E(p), \cos \theta) \times \quad (2.6.63) \\ \times [f(\mathbf{p}', \boldsymbol{\xi}) - f(\mathbf{p}, \boldsymbol{\xi})] d\mathbf{p}'.$$

Этот интеграл удобнее рассматривать в полярных координатах. Полярную ось совместим с  $\mathbf{p}$ , направляющие углы векторов  $\mathbf{p}'$  и  $\boldsymbol{\xi}$  обозначим соответственно  $(\alpha, \beta)$  и  $(\theta, \varphi)$ . Величину  $d\mathbf{p}'$  определим через энергию этого состояния, введем обозначение  $E' = E(p')$ :

$$d\mathbf{p}' = \frac{p'^2(E')}{v(E')} dE' \sin \theta \cdot d\theta d\varphi. \quad (2.6.64)$$

Здесь скорость  $v(E') = \left| \frac{dE'}{dp'} \right|$ .

Проинтегрируем (2.6.63) по  $E'$  с помощью  $\delta$ -функции по энергиям состояний и введем сюда плотность состояний, которую для произвольного закона дисперсии можно определить как интеграл по изоэнергетической поверхности:

$$N(E) = \frac{2}{(2\pi\hbar)^3} \int_{\Sigma} \frac{d\Sigma}{|\nabla_p E(p)|} = \frac{2}{(2\pi\hbar)^3} \int_{\Sigma} \frac{d\Sigma}{v(p)}. \quad (2.6.65)$$

Останутся интегралы по углам:

$$J[f] = \frac{N(E)\xi(E)}{8\pi} \int_0^{\pi} d\theta \cdot \sin\theta \cdot S(E, \cos\theta) \times \\ \times \int_0^{2\pi} d\varphi (p \cos(\mathbf{p}', \xi) - p \cos\alpha). \quad (2.6.66)$$

Из тригонометрии известно:

$$\cos(p', \xi) = \cos\theta \cos\alpha + \sin\theta \sin\alpha \cdot \cos(\varphi - \beta).$$

Подставив это в (2.6.66), получим:

$$J[f] = -\xi(E)p(E) \cos\alpha \frac{N(E)}{4} \int_0^{\pi} d\theta \cdot \sin\theta \times \\ \times (1 - \cos\theta) S(E, \cos\theta) \quad (2.6.67)$$

или 
$$J[f] = -\frac{f_1}{\tau(E)}, \quad (2.6.68)$$

где введенная новая величина

$$\frac{1}{\tau(E)} = \frac{N(E)}{4} \int_0^{\pi} d\theta \cdot \sin\theta \cdot (1 - \cos\theta) S(E, \cos\theta) \quad (2.6.69)$$

действительно имеет размерность обратного времени. Это – скорость релаксации возбуждения, описываемого функцией  $f_1$ . Обратная ей функция,  $\tau(E)$ , – **время релаксации импульса**.

Действительно, рассмотрим релаксацию возбуждения в однородном образце, сразу после снятия возбуждения. Например,

выключения электрического поля. Поскольку поля отсутствуют, а  $f_1 \neq 0$ , кинетическое уравнение (2.6.32) упростится:

$$\frac{\partial f}{\partial t} + \frac{f_1}{\tau(E)} = \frac{\partial f}{\partial t} + \frac{f - f_0}{\tau(E)} = 0.$$

Его решение:  $f_1 = f - f_0 = \text{Const} \cdot e^{-t/\tau}$ .

### 2.6.7. Стационарные решения кинетического уравнения для электронного ансамбля

Мы будем пользоваться уравнением (2.6.32), в котором для интеграла столкновений примем его запись в виде (2.6.68). Это – так называемое  **$\tau$ -приближение**.

#### 2.6.7.1. Статическая электропроводность

$$\mathbf{E} = \mathcal{E}, \quad \mathbf{B} = 0, \quad \nabla T = 0.$$

Поскольку для данного случая мы получили  $f_1 = (\mathbf{p}, \mathcal{E})\psi(E)$ , кинетическое уравнение (2.6.32) запишется в виде:

$$e(\mathbf{v}, \mathcal{E})f_0' - (\mathbf{p}, \mathcal{E})\frac{\Psi}{\tau} = 0.$$

В изотропном кристалле скорость параллельна импульсу, так что  $p = m(E)v$ , где  $m(E)$  – величина размерности массы, но зависящая от энергии. Тогда:

$$\frac{e}{m}(\mathbf{p}, \mathcal{E})f_0' - (\mathbf{p}, \mathcal{E})\frac{\Psi}{\tau} = \left(\frac{e}{m}f_0' - \frac{\Psi}{\tau}\right)(\mathbf{p}, \mathcal{E}) = 0. \quad (2.6.70)$$

В общем случае  $(\mathbf{p}, \mathcal{E}) \neq 0$ , так что

$$\psi = \frac{e}{m} \tau f_0', \quad (2.6.71)$$

А функция распределения

$$f(\mathbf{p}) = f_0(E) + \frac{e}{m} \tau(\mathbf{p}, \mathbf{E}) f_0'. \quad (2.6.72)$$

Нужно определить условия, при которых выполняется требование малости  $f_1$ . При  $f_0' = -\frac{1}{kT} f_0$  оно может быть записано как  $\frac{e p \tau \mathbf{E}}{m k T} \ll 1$ . Обозначив средние равновесные скорость и импульс как  $v_T, p_T$ , это соотношение приведем к соотношению скоростей дрейфа и средней тепловой:  $\frac{e \tau \mathbf{E}}{m} \equiv v_{dr} \ll \frac{kT}{p_T} \approx v_T$ .

Скорость дрейфа зависит от приложенного поля, которое должно быть много меньше некоторого критического:

$$\mathbf{E} \ll \mathbf{E}_{\text{крит}} = \frac{m v_T}{e \tau}. \quad (2.6.73)$$

Это – условие применимости линейного по  $\mathbf{E}$  приближения. В вырожденном полупроводнике здесь должна фигурировать не средняя тепловая скорость, а скорость Ферми.

Найдя неравновесную функцию распределения, можно теперь найти, как зависит плотность тока от поля. Для этого нужно подставить (2.6.73) в (2.6.2):

$$\mathbf{j} = -\frac{2e^2}{(2\pi\hbar)^3} \int \frac{(\mathbf{p}, \mathbf{E})}{m^2(E)} \mathbf{p} \tau f_0' d\mathbf{p}.$$

Этот интеграл так же, как и (2.6.63), удобнее брать в полярных координатах, совместив полярную ось с  $\mathbf{E}$ . Получим:

$$\mathbf{j} = -\frac{e^2 \mathbf{E}}{3} \int_0^\infty N(E) \tau(E) v^2(E) f_0' dE, \quad (2.6.74)$$

откуда проводимость равна

$$\sigma = -\frac{e^2}{3} \int_0^\infty N(E) \tau(E) v^2(E) f_0' dE. \quad (2.6.75)$$

Здесь можно формально ввести подвижность,  $\sigma = e n \mu$ , и определить ее, зная концентрацию носителей (считаем, что имеем дело с  $n$ -типом проводимости):  $n = \int_0^\infty N f_0 dE$ . Тогда величина подвижности, усредненная по ансамблю:

$$\mu = -\frac{e}{3} \cdot \frac{\int_0^\infty N \tau v^2 f_0' dE}{\int_0^\infty N f_0 dE}. \quad (2.6.76)$$

Она, в свою очередь, связана с временем рассеяния импульса и массой как  $\mu = \frac{e \tau_p}{m_\sigma}$ , где  $m_\sigma$ , эффективную массу, определенную из законов проводимости, и называют *эффективной массой электропроводности*. Выражение для нее можно получить, вынеся усредненную величину  $\tau$  из-под интеграла в числителе (2.6.76), получим:

$$m_\sigma = -\frac{3 \int_0^\infty N f_0 dE}{\int_0^\infty N v^2 f_0' dE} \equiv m_{opt}. \quad (2.6.77)$$

Эта же величина фигурирует и в оптических переходах, потому она же называется *оптической эффективной массой*.

При этом закон усреднения времени рассеяния:

$$\tau_p \equiv \langle \tau \rangle = - \frac{\int_0^{\infty} N \tau v^2 f_0' dE}{\int_0^{\infty} N v^2 f_0' dE}. \quad (2.6.78)$$

В вырожденном полупроводнике, в котором  $\xi \gg kT$ , тогда  $f_0' \rightarrow \delta(E - \xi)$ , интегралы в (2.6.78) легко берутся и получится  $\tau_p \rightarrow \tau(\xi)$ .

В вырожденном полупроводнике  $f_0 \approx \exp\left(\frac{\xi - E}{kT}\right)$  и подстановка ее в (2.6.76) дает:

$$\mu = - \frac{e}{3} \cdot \frac{\int_0^{\infty} N \tau v^2 f_0' dE}{\int_0^{\infty} N f_0 dE} = - \frac{e}{3kT} \cdot \frac{\int_0^{\infty} N \tau v^2 \exp\left(-\frac{E}{kT}\right) dE}{\int_0^{\infty} N \exp\left(-\frac{E}{kT}\right) dE}. \quad (2.6.79)$$

В параболической зоне плотность состояний  $N(E) = \frac{(2m^*)^{3/2}}{2\pi^2 \hbar^3} \sqrt{E}$ .

Интеграл в знаменателе (2.6.79) равен:

$$\int_0^{\infty} N \exp\left(-\frac{E}{kT}\right) dE = \frac{1}{\sqrt{2}} \left(\frac{mkT}{\pi \hbar^2}\right)^{3/2}$$

и, подставляя в числителе  $v^2 = \frac{2E}{m}$ , для подвижности получим:

$$\mu = \frac{4e}{3m(kT)^{5/2}\sqrt{\pi}} \cdot \int_0^{\infty} E^{3/2} \tau \cdot \exp\left(-\frac{E}{kT}\right) dE. \quad (2.6.80)$$

Рассмотрение механической части задачи о рассеянии дает (см. [15], гл. XIV) для времени рассеяния показательную зависимость от энергии,  $\tau(E) = CE^r$ . Подстановка ее в (2.6.80) позволяет упростить выражение для подвижности, используя  $\Gamma$ -функции Эйлера,  $\Gamma(x) = \int_0^{\infty} y^{x-1} \exp(-y) dy$ :

$$\mu = \frac{4e\Gamma\left(r + \frac{5}{2}\right)C}{3m\sqrt{\pi}} \cdot (kT)^r. \quad (2.6.81)$$

### 2.6.7.2. Действие температурного поля

Неравновесная функция распределения ищется в виде (2.6.57):

$$f_1 = (\mathbf{p}, \nabla T) \chi_1(E) + (\mathbf{p}, \mathbf{E}') \chi_2(E).$$

В стационарном состоянии  $\frac{\partial f_1}{\partial t} = 0$ , магнитное поле  $\mathbf{B} = 0$ , так что кинетическое уравнение в этом случае:

$$(\mathbf{v}, e\mathbf{E}') f_0' - \frac{E(p) - \zeta}{T} (\mathbf{v}, \nabla T) f_0' = \frac{f_1}{\tau}. \quad (2.6.82)$$

Здесь  $\zeta$  - химический потенциал для электронов. Мы по-прежнему для определенности считаем, что имеем дело с  $n$ -полупроводником.

Отсюда сразу получаем, принимая для изотропного полупроводника соотношение  $p = m(E)v$ ,

$$\chi_1 = \tau \frac{E(p) - \zeta}{mT} f_0', \quad \chi_2 = \frac{e\tau}{m} f_0', \quad (2.6.83)$$

и тогда функция распределения равна:

$$f = f_0 + (\mathbf{p}, \nabla T) \frac{E(p) - \zeta}{mT} \tau f_0' + (\mathbf{p}, \mathbf{E}') \frac{e\tau}{m} f_0'. \quad (2.6.84)$$

### *Дифференциальная термоэдс.*

Плотность тока можно найти, опять-таки, подставив ее в (2.6.2):

$$j = \frac{e}{3} \nabla T \int_0^{\infty} N v^2 \tau \frac{\zeta - E}{T} f_0' dE - \frac{e^2 \mathbf{E}'}{3} \int_0^{\infty} N v^2 \tau f_0' dE. \quad (2.6.85)$$

Здесь мы опустили явное указание на зависимость подынтегральных членов от энергии. В этой сумме двух членов второй – ток, вызванный наличием электрического поля. От (2.6.74) он отличается только тем, что здесь фигурирует не внешнее поле  $\mathbf{E}$ , а эффективное  $\mathbf{E}'$ .

При феноменологическом описании тока, возникающего в температурном поле, мы получили (2.6.42):

$$j = \sigma \mathbf{E}' - \alpha \sigma \cdot \nabla T = \sigma (\mathbf{E}' - \alpha \cdot \nabla T),$$

где  $\alpha$  - дифференциальная термоэдс. Из (2.6.85) получаем для нее:

$$\alpha = -\frac{k}{e} \left\{ -\frac{\zeta}{kT} + \frac{1}{kT} \frac{\int_0^{\infty} EN v^2 \tau f_0' dE}{\int_0^{\infty} N v^2 \tau f_0' dE} \right\}. \quad (2.6.86)$$

Для дырочного полупроводника мы получили бы то же самое, с заменой  $\zeta \rightarrow -(E_g + \zeta)$  и выражений для скорости, времени релаксации и плотности состояний на дырочные.

В параболической зоне  $N = \text{Const} \cdot \sqrt{E}$ ,  $v^2 = 2E/m$ , так что (2.6.86) можно несколько упростить:

$$\alpha = -\frac{k}{e} \left\{ -\frac{\zeta}{kT} + \frac{1}{kT} \frac{\int_0^{\infty} E^{5/2} \tau f_0' dE}{\int_0^{\infty} E^{3/2} \tau f_0' dE} \right\}. \quad (2.6.87)$$

Эта формула называется формулой Писаренко.

Для невырожденного электронного газа :

$$\alpha = -\frac{k}{e} \left\{ -\frac{\zeta}{kT} + \frac{1}{kT} \frac{\int_0^{\infty} E^{5/2} \tau \exp(-E/kT) dE}{\int_0^{\infty} E^{3/2} \tau \exp(-E/kT) dE} \right\}. \quad (2.6.88)$$

Здесь, опять-таки, можно принять температурную зависимость времени рассеяния  $\tau(E) = CE^r$ , перейти к  $\Gamma$ -функциям и, воспользовавшись их свойством:  $\Gamma(x+1) = x\Gamma(x)$ , получим:

$$\alpha_n = -\frac{k}{e} \left\{ -\frac{\zeta}{kT} + \left( r + \frac{5}{2} \right) \right\}, \quad \alpha_p = \frac{k}{e} \left\{ \frac{\zeta + E_g}{kT} + \left( r + \frac{5}{2} \right) \right\}. \quad (2.6.89)$$

В вырожденном полупроводнике можно положить  $f_0' = -\delta(E - \zeta)$ , но тогда получим тождественный ноль:

$$\alpha = -\frac{k}{e} \left\{ -\frac{\zeta}{kT} + \frac{1}{kT} \frac{\zeta^{5/2} \tau(\zeta)}{\zeta^{3/2} \tau(\zeta)} \right\} \equiv 0.$$

Более точное вычисление интегралов в (2.6.87) для вырожденной системы дает (см. [15], Приложение XV):

$$\alpha = -\frac{k \pi^2 kT}{e 3 \zeta} \left[ \frac{3}{2} + \frac{d \ln \tau(E)}{d \ln E} \Big|_{E=\zeta} \right]. \quad (2.6.90)$$

При  $\tau(E) = CE^r$  член в квадратных скобках равен  $(1,5 + r)$ .

Легко видеть, что величины дифференциальной термоэдс, полученные по формулам (2.6.89), (2.6.90) различаются принципиально. И для того, и для другого случая можно считать, что  $\zeta > kT$ . В металлах эта величина имеет порядок  $10^2 \dots 10^3$ , в невырожденных полупроводниках – больше, чем  $(3 \dots 5)$ . Но в (2.6.90) она стоит в знаменателе, а остальные члены в этих формулах имеют одинаковые порядки. Таким образом, дифференциальную термоэдс в металлах можно считать пренебрежимо малой.

### ***Поток энергии***

Воспользовавшись функцией распределения (2.6.84) и формулой (2.6.3), получим:

$$\begin{aligned} \mathbf{I} = & \frac{2}{(2\pi\hbar)^3} \int \mathbf{v}(\mathbf{p}) \tau \frac{E(\mathbf{p}) - \zeta}{mT} (\mathbf{p}, \nabla T) (E(\mathbf{p}) - e\varphi) f_0' d\mathbf{p} + \\ & + \frac{2e}{(2\pi\hbar)^3} \int \mathbf{v}(\mathbf{p}) \cdot \boldsymbol{\tau} \cdot (\mathbf{p}, \boldsymbol{\mathcal{E}}') (E(\mathbf{p}) - e\varphi) f_0' d\mathbf{p}. \end{aligned} \quad (2.6.91)$$

В этой сумме можно перейти, как и раньше, к интегрированию по энергии. Тогда окажется, что слагаемые, содержащие  $\varphi$  точно равны произведению плотности тока (2.6.85) на  $\varphi$  и полный поток энергии можно представить как

$$\mathbf{I} = \mathbf{j}\varphi + \mathbf{I}', \quad \text{где} \quad (2.6.92)$$

$$\mathbf{I}' = \frac{4nkC(kT)^{r+1}\Gamma\left(r + \frac{7}{2}\right)}{3m\sqrt{\pi}} \times \left\{ -\nabla T - \frac{e}{k}\mathbf{E}' + \left[ \frac{\zeta}{kT} - \left(r + \frac{5}{2}\right) \right] \nabla T \right\} \quad (2.6.93)$$

Мы использовали зависимость  $\tau(E) = CE^r$ .

Сравнивая третий член в фигурных скобках с (2.6.89) видим, что он равен  $\left[ \frac{\zeta}{kT} - \left(r + \frac{5}{2}\right) \right] \nabla T = \frac{e\alpha}{k} \nabla T$ . Сумма второго и третьего, как следует из (2.6.42), равна  $-\frac{e}{k} \frac{\mathbf{j}}{\sigma} = -\frac{\mathbf{j}}{kn\mu}$ .

Используя выражение (2.6.81) для  $\mu$ , коэффициент перед фигурной скобкой в (2.6.93) можно переписать в виде:

$$\frac{k}{e} \mu nkT \left(r + \frac{5}{2}\right)$$

и тогда сумма двух последних членов (2.6.93) равна:

$$-\mathbf{j} \frac{kT}{e} \left(r + \frac{5}{2}\right) = -\mathbf{j} \frac{kT}{e} \left[ \left(r + \frac{5}{2}\right) - \frac{\zeta}{kT} \right] - \frac{\zeta}{e} \mathbf{j} = T\alpha \mathbf{j} - \frac{\zeta}{e} \mathbf{j}.$$

Первый член в (2.6.93) – поток тепла, не связанный с протеканием тока, т.е. за этот член ответственна теплопроводность. Если, опять-таки, воспользоваться выражением (2.6.42) для  $\mu$  и  $\sigma = en\mu$ , то для коэффициента электронной теплопроводности получим:

$$K = \frac{4\pi C(kT)^{r+1}\Gamma\left(r + \frac{7}{2}\right)}{3m\sqrt{\pi}} = \frac{k^2 T \sigma}{e^2} \left(r + \frac{5}{2}\right). \quad (2.6.94)$$

Окончательная формула для плотности потока энергии электронов:

$$\mathbf{I} = -K\nabla T + T\alpha\mathbf{j} - \frac{\zeta - e\varphi}{e}\mathbf{j} = -K\nabla T + T\alpha\mathbf{j} - \frac{F}{e}\mathbf{j}. \quad (2.6.95)$$

Мы видим, что это выражение совпадает с (2.6.47) при  $\Pi = \alpha T$ . Теперь коэффициент Томсона (2.6.50) можно записать как функцию только  $\alpha$ :

$$\alpha_T = \frac{d\Pi}{dT} - \alpha = T \frac{d\alpha}{dT}. \quad (2.6.96)$$

Условие применимости полученных выражений,  $|f_1| \ll f_0$ , можно записать (см. (2.6.83), (2.6.84)) как:

$$v_T \tau \frac{|\nabla T|}{T} \left| \frac{\zeta}{kT} - 1 \right| \ll 1. \quad (2.6.97)$$

Фигурирующий здесь модуль разности больше единицы, для невырожденных полупроводников он от нескольких единиц до нескольких десятков. Если им пренебречь, из (2.6.97) получим неравенство:

$$v_T \tau |\nabla T| = l_p |\nabla T| \ll T. \quad (2.6.98)$$

Здесь  $l_p$  – длина свободного пробега. Иными словами, изменение температуры на длине свободного пробега должно быть малым по сравнению с самой температурой. Это – довольно естественное требование. В противном случае трудно вообще говорить о температуре. В комнатных условиях это требование не слишком жестко, но с понижением температуры одновременно растет и  $l_p$ .

### 2.6.7.3. Совместное действие электрического и магнитного полей

Природу эффекта Холла и поперечного магнетосопротивления мы рассматривали в разделе 2.3.2, см. (2.3.41) (2.3.42).

Для стационарного состояния, при  $\nabla T = 0$  кинетическое уравнение принимает вид:

$$e(\mathbf{v}, e\mathbf{E}')f_0' + \frac{e}{c}([\mathbf{v}\mathcal{B}], \nabla_p f_1) = \frac{f_1}{\tau}. \quad (2.6.99)$$

При неравновесной функции распределения (2.6.58) второй член слева равен:

$$([\mathbf{v}\mathcal{B}], \nabla_p f_1) = (\mathbf{v}, [\mathcal{B}\mathbf{E}])\psi_1 + (\mathbf{v}\mathbf{E})\mathcal{B}^2\psi_2 - (\mathbf{p}, \mathcal{B})(\mathcal{B}\mathbf{E})\psi_3. \quad (2.6.100)$$

Тогда кинетическое уравнение приводится к виду:

$$\begin{aligned} \left\{ \frac{e}{m}f_0' - \frac{\psi_1}{\tau} + \frac{e\mathcal{B}^2}{mc}\psi_2 \right\} (\mathbf{p}\mathbf{E}) - \left\{ \frac{e}{mc}\psi_1 + \frac{\psi_2}{\tau} \right\} (\mathbf{p}, [\mathcal{E}\mathcal{B}]) - \\ - \left\{ \frac{e}{mc}\psi_2 + \frac{\psi_3}{\tau} \right\} (\mathbf{p}, \mathcal{B})(\mathcal{B}\mathbf{E}) = 0. \end{aligned} \quad (2.6.101)$$

Но углы между векторами в скалярных произведениях  $(\mathbf{p}\mathbf{E})$ ,  $(\mathbf{p}, [\mathcal{E}\mathcal{B}])$ ,  $(\mathbf{p}, \mathcal{B})(\mathcal{B}\mathbf{E})$  произвольны, так что равенство нулю здесь возможно только при равенстве нулю всех трех слагаемых. Отсюда сразу получаем:

$$\psi_1 = \frac{e\tau}{m(1 + \omega_c^2\tau^2)}f_0', \quad (2.6.102)$$

$$\psi_2 = \frac{1}{c} \frac{(e\tau)^2}{m^2(1 + \omega_c^2\tau^2)}f_0', \quad (2.6.103)$$

$$\psi_3 = \frac{1}{c^2} \frac{(e\tau)^3}{m^3(1 + \omega_c^2\tau^2)}f_0', \quad (2.6.104)$$

где  $\omega(E) = \frac{e}{m(E)c}\mathcal{B}$  – циклотронная частота, см. пункт 2.3.2.1.

Точное соответствие получится только для параболического за-

кона дисперсии, когда  $m$  – постоянная величина, не зависящая от энергии.

Теперь мы знаем неравновесную функцию распределения и можем определить плотность тока, подставив ее в (2.6.2). Получим выражение вида (2.6.51), но с явно выписанными коэффициентами. Для параболического закона дисперсии и фиксированной массы получим:

$$\begin{aligned} a_1 &= \frac{2e^2\sqrt{2m}}{3\pi^2\hbar^3} \mathcal{J}_1, & a_2 &= \frac{2e^2\sqrt{2m}}{3\pi^2\hbar^3} \frac{e}{mc} \mathcal{J}_2, \\ a_3 &= \frac{2e^2\sqrt{2m}}{3\pi^2\hbar^3} \left(\frac{e}{mc}\right)^2 \mathcal{J}_3, \end{aligned} \quad (2.6.105)$$

где использованы обозначения интегралов:

$$\mathcal{J}_k = -\int_0^\infty E^{3/2} \frac{\tau^k}{1 + \omega_c^2 \tau^2} f_0' dE. \quad (2.6.106)$$

В вырожденном случае, когда  $f_0' \rightarrow -\delta(E - \zeta)$ , получим:

$$\mathcal{J}_k = \zeta^{3/2} \frac{\tau^k(\zeta)}{1 + \omega_c^2(\zeta)\tau^2(\zeta)}. \quad (2.6.107)$$

Но и это – сложная функция магнитного поля. Выражения несколько упрощаются в крайних случаях,

при  $\omega_c^2 \tau^2 \ll 1$  получим  $\mathcal{J}_k \approx -\int_0^\infty E^{3/2} \tau^k (1 - \omega_c^2 \tau^2) f_0' dE$ , (2.6.108)

при  $\omega_c^2 \tau^2 \gg 1$  –  $\mathcal{J}_k = -\frac{1}{\omega_c^2} \int_0^\infty E^{3/2} \tau^{k-2} f_0' dE$ . (2.6.109)

### Постоянная Холла

В поперечном поле,  $\mathcal{B} \perp \mathbf{j}$ , подставив (2.6.105) в (2.6.54), получим:

$$R = -\frac{3\pi^2 \hbar^3}{2ec\sqrt{2m^3}} \frac{\mathcal{J}_2}{\mathcal{J}_1^2 + \omega_c^2 \mathcal{J}_2^2} = -\frac{\gamma}{nec}, \quad (2.6.110)$$

где мы воспользовались для  $n$  выражением:

$$n = 2 \left( \frac{2\pi m_n kT}{(2\pi\hbar)^2} \right)^{3/2} \Phi_{1/2} \left( \frac{\zeta}{kT} \right) \equiv N_c \Phi_{1/2} \left( \frac{\zeta}{kT} \right)$$

и обозначили числитель, равный отношению подвижностей Холла и электропроводности (2.6.80), (2.6.81), символом  $\gamma$ . Для этого отношения из (2.6.110) следует:

$$\gamma = -\frac{3\sqrt{\pi}}{4} (kT)^{3/2} \Phi_{1/2} \left( \frac{\zeta}{kT} \right) \frac{\mathcal{J}_2}{\mathcal{J}_1^2 + \omega_c^2 \mathcal{J}_2^2}. \quad (2.6.111)$$

Можно показать, что в вырожденных полупроводниках и в сильных (по критерию (2.6.109)) полях  $\gamma \approx 1$ .

В слабых полях

$$\gamma_{\text{сл.поле}} = -\frac{3\sqrt{\pi}}{4} (kT)^{3/2} \Phi_{1/2} \left( \frac{\zeta}{kT} \right) \frac{\mathcal{J}_2^{(0)}}{(\mathcal{J}_1^{(0)})^2}, \quad (2.6.112)$$

где

$$\mathcal{J}_k^{(0)} = -\int_0^\infty E^{3/2} \tau^k f_0' dE. \quad (2.6.113)$$

Вдали от вырождения в слабых полях получим:

$$\gamma_{\text{сл. поле, не выр.}} = -\frac{3\sqrt{\pi}}{4} (kT)^{5/2} \frac{\int_0^{\infty} E^{3/2} \tau^2 \exp(-E/kT) dE}{\left[ \int_0^{\infty} E^{3/2} \tau \exp(-E/kT) dE \right]^2} \quad (2.6.114)$$

или, в случае зависимости типа  $\tau(E) = CE^r$ , это выражение можно преобразовать к виду:

$$\gamma_{\text{сл. поле, не выр.}} = \frac{3\sqrt{\pi}\Gamma\left(2r + \frac{5}{2}\right)}{4\Gamma^2\left(r + \frac{5}{2}\right)}. \quad (2.6.115)$$

Разделение по критериям (2.6.108), (2.6.109) полей на "сильное" и "слабое", приводит к некоторому "пороговому" полю, при котором (2.6.108), (2.6.109) могут быть переписаны в форме:

$$\mathcal{B} \ll \mathcal{B}_{\text{порог}} \text{ или } \mathcal{B} \gg \mathcal{B}_{\text{порог}}. \text{ Здесь } \mathcal{B}_{\text{порог}} = \frac{mc}{e\tau}.$$

рогового поля для эффективных масс порядка  $0,1m_0$  (характерных для "классических" полупроводников) и некоторых средних времен рассеяния при комнатной температуре дают величину порядка  $\mathcal{B}_{\text{порог}} \approx 0,01 \div 0,1 \text{ Т}$ . Это много меньше напряженностей квантующих полей, для которых должно быть выполнено  $\hbar\omega_c = \hbar \frac{e\mathcal{B}}{mc} > \langle E \rangle$ . В невырожденных полупроводниках при комнатной температуре  $\langle E \rangle \approx 0.025 \text{ эВ}$ , (у вырожденных  $\langle E \rangle = \zeta$ , что может быть много больше). Квантующие поля получаются порядка нескольких десятков Тесла. Потому для наблюдения квантовых эффектов необходима работа при низких температурах.

### ***Поперечное магнетосопротивление***

Подставив в (2.6.55) значения коэффициентов из (2.6.105), получим:

$$\frac{\rho_{\perp} - \rho_0}{\rho_0} = \frac{\mathcal{J}_1 \mathcal{J}_1^2}{\mathcal{J}_1^2 + \omega_c^2 \mathcal{J}_2^2} - 1. \quad (2.6.116)$$

Можно показать, что в вырожденном газе, в изотропном кристалле, эта величина равна нулю. Поперечное магнетосопротивление отсутствует. В невырожденном – для слабого и сильного полей получим соответственно:

$$\begin{aligned} \left. \frac{\rho_{\perp} - \rho_0}{\rho_0} \right|_{\text{сл. поле}} &= \\ &= \omega_c^2 C^2 (kT)^{2r} \frac{\Gamma(3r + 5/2) \Gamma(r + 5/2) - \Gamma^2(2r + 5/2)}{\Gamma^2(r + 5/2)} \end{aligned} \quad (2.6.117)$$

и

$$\left. \frac{\rho_{\perp} - \rho_0}{\rho_0} \right|_{\text{сильн. поле}} = \frac{16}{9\pi} \Gamma(5/2 - r) \Gamma(5/2 + r) - 1. \quad (2.6.118)$$

#### ***2.6.7.4. Действие слабого переменного электрического поля***

Кинетическое уравнение при наличии только электрического поля:

$$\frac{\partial f}{\partial t} - e(\mathbf{E}, \nabla_p f) + \frac{f - f_0}{\tau(E)} = 0. \quad (2.6.119)$$

Предположим, что в данном случае действует монохроматическое поле:

$$\mathbf{E} = \mathbf{E}_m e^{-i\omega t}. \quad (2.6.120)$$

Неравновесная функция распределения имеет вид:

$$f_1 = (\mathbf{p}, \mathbf{E}) \psi(E). \quad (2.6.121)$$

Подставив ее в (2.6.119), получим:

$$\psi(E) = \frac{e}{m} \frac{\tau}{(1 - i\omega\tau)} f_0', \quad (2.6.122)$$

т.е. 
$$f = f_0 + \frac{e}{m} (\mathbf{p}, \mathbf{E}_m) e^{-i\omega\tau} \frac{\tau}{1 - i\omega\tau} f_0'. \quad (2.6.123)$$

Носители заряда в слабом электрическом поле обладают своеобразной "инерцией". Функция распределения отстает по фазе на угол  $\text{arctg}(\omega\tau)$ . Существенно здесь же отметить, что этот эффект связан не с массой частиц, а со статистическими свойствами их системы. Поле ускоряет все электроны (дырки) одинаковым образом, но на виде функции распределения это сказывается лишь в результате столкновений.

Если в (2.6.2) подставить эту функцию распределения, то плотность тока получится комплексной величиной, так что приходится и проводимость считать величиной комплексной,

$$\sigma = \sigma_1 + i\sigma_2. \quad (2.6.124)$$

Для ее компонент получим:

$$\sigma_1 = -\frac{e^2}{3} \int_0^\infty N(E) v^2(E) \frac{\tau(E)}{1 + \omega^2(E)\tau^2(E)} f_0' dE, \quad (2.6.125)$$

$$\sigma_2 = -\omega \frac{e^2}{3} \int_0^\infty N(E) v^2(E) \frac{\tau^2(E)}{1 + \omega^2(E)\tau^2(E)} f_0' dE. \quad (2.6.126)$$

Смысл вещественной и мнимой частей проводимости понятен. Мнимая часть описывает реакцию, не связанную с рассеянием энергии. Если за полпериода колебаний внешнего поля электрон не испытает рассеяния, то он будет "хранить память" о

предшествующем времени и двигаться будет колебательно. В среднем его координаты не будут изменяться, так что он и не переносит заряда.

При малых частотах,  $\omega^2\tau^2 \ll 1$ , имеем:

$$\sigma_1|_{\omega \rightarrow 0} = -\frac{e^2}{3} \int_0^{\infty} N\tau v^2 f_0' dE, \quad (2.6.127)$$

$$\sigma_2|_{\omega \rightarrow 0} = -\omega \frac{e^2}{3} \int_0^{\infty} N\tau^2 v^2 f_0' dE \rightarrow 0. \quad (2.6.128)$$

Вещественная часть растет с  $\tau$ . Это понятно, с  $\tau$  растет подвижность носителей, средняя скорость их движения в единичном поле.

Мнимая часть растет с  $\tau$  еще быстрее, растет вероятность того, что за период колебаний носитель не успеет столкнуться с решеткой и при этом, как было отмечено, не примет участия в переносе заряда. Но мнимая часть к тому же пропорциональна  $\omega$  и на низких частотах исчезает. Остается только вещественная, равная статической электропроводности (2.6.75).

На высоких частотах,  $\omega^2\tau^2 \gg 1$ , напротив, имеем:

$$\sigma_1 = -\frac{e^2}{3\omega^2} \int_0^{\infty} Nv^2 \frac{1}{\tau} f_0' dE, \quad (2.6.129)$$

$$\sigma_2 = -\frac{e^2}{3\omega} \int_0^{\infty} Nv^2 f_0' dE. \quad (2.6.130)$$

Вещественная часть очень быстро устремляется к нулю и уменьшается при росте  $\tau$ . Мнимая от  $\tau$  вообще не зависит. Следовательно, ее уменьшение с ростом  $\omega$  – чисто инерциальный эффект, связанный с массой. Действительно,  $\sigma_2$  описывает диэлек-

трическую реакцию электронного газа, колебания, а не дрейф носителей. А амплитуда вынужденных колебаний падает с ростом частоты.

### **Связь $\sigma_2$ с диэлектрической проницаемостью**

Действительно, рассмотрим уравнение Максвелла:

$$\operatorname{rot}\mathcal{H} = \frac{4\pi}{c}\mathbf{j} + \frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}. \quad (2.6.131)$$

$$\text{Здесь } \mathbf{j} = (\sigma_1 + i\sigma_2)\mathbf{E}, \quad \mathbf{D} = \varepsilon_0\mathbf{E}.$$

Мы записываем эти соотношения в системе СГСЭ и здесь  $\varepsilon_0$  – относительная диэлектрическая проницаемость, обусловленная решеткой и той частью электронного ансамбля, которую мы считаем неподвижной.

Считая, что  $\varepsilon_0 = \varepsilon_1 + i\varepsilon_2$  и подставив  $\mathbf{j}$  и  $\mathbf{D}$  в (2.6.129), разделим вещественную и мнимую части. Получим:

$$\operatorname{rot}\mathcal{H} = \frac{4\pi}{c}\left(\omega\frac{\varepsilon_2}{4\pi} + \sigma_1\right)\mathbf{E} - i\frac{\omega}{c}\left(\varepsilon_1 - \frac{4\pi}{\omega}\sigma_2\right)\mathbf{E}. \quad (2.6.132)$$

Видим, что вещественная часть диэлектрической проницаемости изменилась на величину:

$$\Delta\varepsilon = -\frac{4\pi}{\omega}\sigma_2 = \frac{4\pi e^2}{3}\int_0^\infty N(E)v^2\frac{\tau^2}{(1+\omega^2\tau^2)}f_0'dE. \quad (2.6.133)$$

Эта величина отрицательна, поскольку  $f_0' < 0$  и, в принципе, по модулю может превышать  $\varepsilon_1$ . Тогда становится отрицательной полная диэлектрическая проницаемость, волновой вектор волны в кристалле становится мнимым. Это означает, что волны соответствующей частоты не могут распространяться в данном веществе, испытывая полное отражение от его поверхности. Такая ситуация

иногда реализуется в плазме. В том числе – в ионосфере и в электронной плазме твердых тел, см. разд. 1.6.4.6 в [первая часть].

На высоких частотах,  $\omega^2 \tau^2 \gg 1$ , получим:

$$\Delta \varepsilon = -\frac{4\pi n e^2}{m_{opt} \omega^2}, \quad (2.6.134)$$

где 
$$\frac{1}{m_{opt}} = \frac{1}{3} \left( \frac{1}{m_x} + \frac{1}{m_y} + \frac{1}{m_z} \right). \quad (2.6.135)$$

Следует отметить также, что все приведенные соображения и результаты относятся к пространственно однородному полю. Простые соображения показывают, что это приближение может быть справедливо в случае, если длина свободного пробега существенно меньше длины волны.

### ***Плазменные волны***

Ранее, в разд. 1.6.4.6, в [1], мы видели, что при  $\varepsilon = 0$  в веществе могут существовать только продольные волны, так называемые плазменные колебания.

Согласно (2.6.134), это возможно при

$$\varepsilon_0 = -\Delta \varepsilon = \frac{4\pi n e^2}{m_{opt} \omega^2} \equiv \frac{\omega_p^2}{\omega^2}, \quad (2.6.136)$$

где

$$\omega_p^2 = \frac{4\pi n e^2}{m_{opt}} \quad (\text{в СГСЭ}) \quad \text{или} \quad \omega_p \equiv \sqrt{\frac{n e^2}{m \varepsilon_0}} \quad (\text{в СИ}), \quad (2.6.137)$$

в полном соответствии с (1.6.71).

# Литература

---

- 1 Заглубский А.А., Чернова А.П., Пособие "Структура кристаллов и электронные состояния", СПб, Соло, 2007.
- 2 Жданов Г.С. Физика твердого тела. М., Изд-во МГУ, 1962
- 3 Ансельм А.И. Введение в теорию полупроводников, М., Наука, 1978
- 4 Киттель Ч., Введение в физику твердого тела, М., Наука, 1978
- 5 Ашкрофт Н., Мермин Н., Физика твердого тела, т.1, 2, М., Мир, 1979
- 6 de Haas W. J., van Alphen P. M., Leiden Comm., 208d, 212a (1930); 220d (1932).
- 7 Лифшиц И. М., Косевич А. М., ДАН СССР, 1953, т. 91, с. 795; ЖЭТФ, 1955, т. 29, с. 730.
- 8 Ландау Л.Д., Лифшиц Е.М. Краткий курс теоретической физики. Квантовая механика. М., Наука, 1972.
- 9 Гирвин С. Квантовый эффект Холла. ISBN 5-93972-288-1, ИКИ, 2003 г.
- 10 Драгунов В.П., Неизвестный И.Г., Гридчин В.А. Основы наноэлектроники, М., Логос, 2006.
- 11 Von Klitzing K., Dorda G., Pepper M. New method for high accuracy determination of the fine-structure constant based on quantum Hall resistance// Phys. Rev. Lett. 1980. V.45. p. 494-497.
- 12 Ebert G., von Klitzing K., Probst C., Ploog K. // Sol. St. Comm. 44, 95, 1982
- 13 Кардона М. Основы физики полупроводников. М., ФИЗМАТЛИТ, 2002.
- 14 Роуз А., основы теории фотопроводимости. М., Мир, 1966.
- 15 Бонч-Бруевич В.Л., Калашников С.Г., Физика полупроводников, изд. 2, М., Наука, 1990.
- 16 Анималу А. Квантовая теория кристаллических твердых тел, М., Мир, 1981
- 17 Абакумов В.Н., Перель В.И., Ясиевич И.Н., Безызлучательная рекомбинация, СПб, 1997